frontiers in NEUROLOGY

Brain Imaging Methods

Subject-Motion Correction in HARDI Acquisitions: Choices and Consequences

Shireen Y. Elhabian, Yaniv Gur, Clement Vachet, Joseph Piven, Martin Andreas Styner, Ilana R Leppert, Bruce Pike and Guido Gerig

Journal Name:	Frontiers in Neurology
ISSN:	1664-2295
Article type:	Original Research Article
First received on:	15 Jun 2014
Revised on:	04 Nov 2014
Frontiers website link:	www.frontiersin.org

Frontiers in Neurology

Research Article
20144 November 2014

Subject-Motion Correction in HARDI Acquisitions: Choices and Consequences

Shireen Elhabian [1,6,*], **Yaniv Gur** [1], **Clement Vachet** [1], **Joseph Piven** [2] **for IBIS, Martin Styner** [2,3], **Ilana R Leppert** [4], **G. Bruce Pike** [4,5], **Guido Gerig** [1]

[1] *Scientific Computing and Imaging Institute, Salt Lake City, UT, USA.*
[2] *Dept. of Psychiatry and* [3] *Dept. of Computer Science, University of North Carolina, NC, USA.*
[4] *Dept. of Neurology and Neurosurgery, Montreal Neurological Institute, Montreal, Quebec, Canada.*
[5] *Dept. of Radiology, University of Calgary, Calgary, Canada.*
[6] *Faculty of Computers and Information, Cairo University, Cairo, Egypt.*

Correspondence*:
Shireen Elhabian
Scientific Computing and Imaging Institute, 72 Central Campus Drive, Salt Lake City, UT, 84112, USA., shireen@sci.utah.edu

ABSTRACT

Diffusion-weighted imaging (DWI) is known to be prone to artifacts related to motion originating from subject movement, cardiac pulsation and breathing, but also to mechanical issues such as table vibrations. Given the necessity for rigorous quality control and motion correction, users are often left to use simple heuristics to select correction schemes, which involves simple qualitative viewing of the set of DWI data, or the selection of transformation parameter thresholds for detection of motion outliers. The scientific community offers strong theoretical and experimental work on noise reduction and orientation distribution function (ODF) reconstruction techniques for HARDI data, where postacquisition motion correction is widely performed, e.g., using the open-source DTIprep software (**Oguz et al.**, 2014), FSL (the FMRIB Software Library) (**Jenkinson et al.**, 2012) or TORTOISE (**Pierpaoli et al.**, 2010). Nonetheless, effects and consequences of the selection of motion correction schemes on the final analysis, and the eventual risk of introducing confounding factors when comparing populations, are much less known and far beyond simple intuitive guessing. Hence, standard users lack clear guidelines and recommendations in practical settings. This paper reports a comprehensive evaluation framework to systematically assess the outcome of different motion correction choices commonly used by the scientific community on different DWI-derived measures. We make use of human brain HARDI data from a well-controlled motion experiment to simulate various degrees of motion corruption and noise contamination. Choices for correction include exclusion/scrubbing or registration of motion corrupted directions with different choices of

interpolation, as well as the option of interpolation of all directions. The comparative evaluation is based on a study of the impact of motion correction using four metrics that quantify (1) similarity of fiber orientation distribution functions (fODFs), (2) deviation of local fiber orientations, (3) global brain connectivity via Graph Diffusion Distance (GDD) and (4) the reproducibility of prominent and anatomically defined fiber tracts. Effects of various motion correction choices are systematically explored and illustrated, leading to a general conclusion of discouraging users from setting ad-hoc thresholds on the estimated motion parameters beyond which volumes are claimed to be corrupted.

Keywords: HARDI, subject motion, motion correction, fiber orientations, orientation distribution functions, tractography comparison, impact quantification

1 INTRODUCTION

Diffusion-weighted (DW)-MRI enables probing the fiber architecture of biological tissues - *in vivo* - by encoding the microscopic direction and speed of the diffusion of water molecules (**Yendiki et al.**, 2014), while reflecting the amount of hindrance experienced by such molecules along the axis of the applied diffusion gradient due to barriers and obstacles imposed by micro-structures (**Jones et al.**, 2013). Today, diffusion tensor imaging (DTI) is the method of choice for most neuroimaging studies, e.g., autism (**Wolff et al.**, 2012), schizophrenia (**Gilmore et al.**, 2010) and Huntington's disease (**Dumas et al.**, 2012). Nonetheless, DTI assumes a homogeneous axon population inside a single voxel (**Le Bihan et al.**, 2006) and fails at modeling more realistic heterogeneous populations. High angular resolution diffusion imaging (HARDI) (**Tuch et al.**, 2002), on the other hand, allows the diffusion acquisition to focus on the angular component of the DW signal using strong gradients and long diffusion times (**Jones et al.**, 2013)), while revealing the intra-voxel orientational heterogeneity, such as crossing and merging fiber bundles. The promising potential of HARDI-based DW-MRI in describing fiber tracts within the human brain comes with a price tag of a wide variety of artifacts related to the gradient system hardware, pulse sequence, acquisition strategy and subject motion (**Soares et al.**, 2013). Such artifacts renders the quality of diffusion imaging questionable and reduces the accuracy of findings when left uncorrected (**Oguz et al.**, 2014).

1.1 MOTION ARTIFACTS

In today's clinical DW-MRI acquisitions, the presence of the long and strong gradient pulses have made diffusion MRI more sensitive to the detrimental effects of subject motion than other MRI techniques (**Pierpaoli**, 2010; **Le Bihan et al.**, 2006; **Gumus et al.**, 2013). During a scanning session, the degree of a patient's cooperation may vary: elderly people who may become uncomfortable during large scanning sessions, patients in pain who become restless and agitated during a scan and unsedated pediatric subjects who will not cooperate long enough to be imaged without motion artifacts. Hence, it is safe to assume that there are always motion artifacts in any given DW-MRI acquisition due to the increased likelihood of involuntary subject motion; especially with HARDI acquisitions, which use a large number of gradient directions resulting in longer scan times. A proof-of-concept of this hypothesis is presented in 2.1.

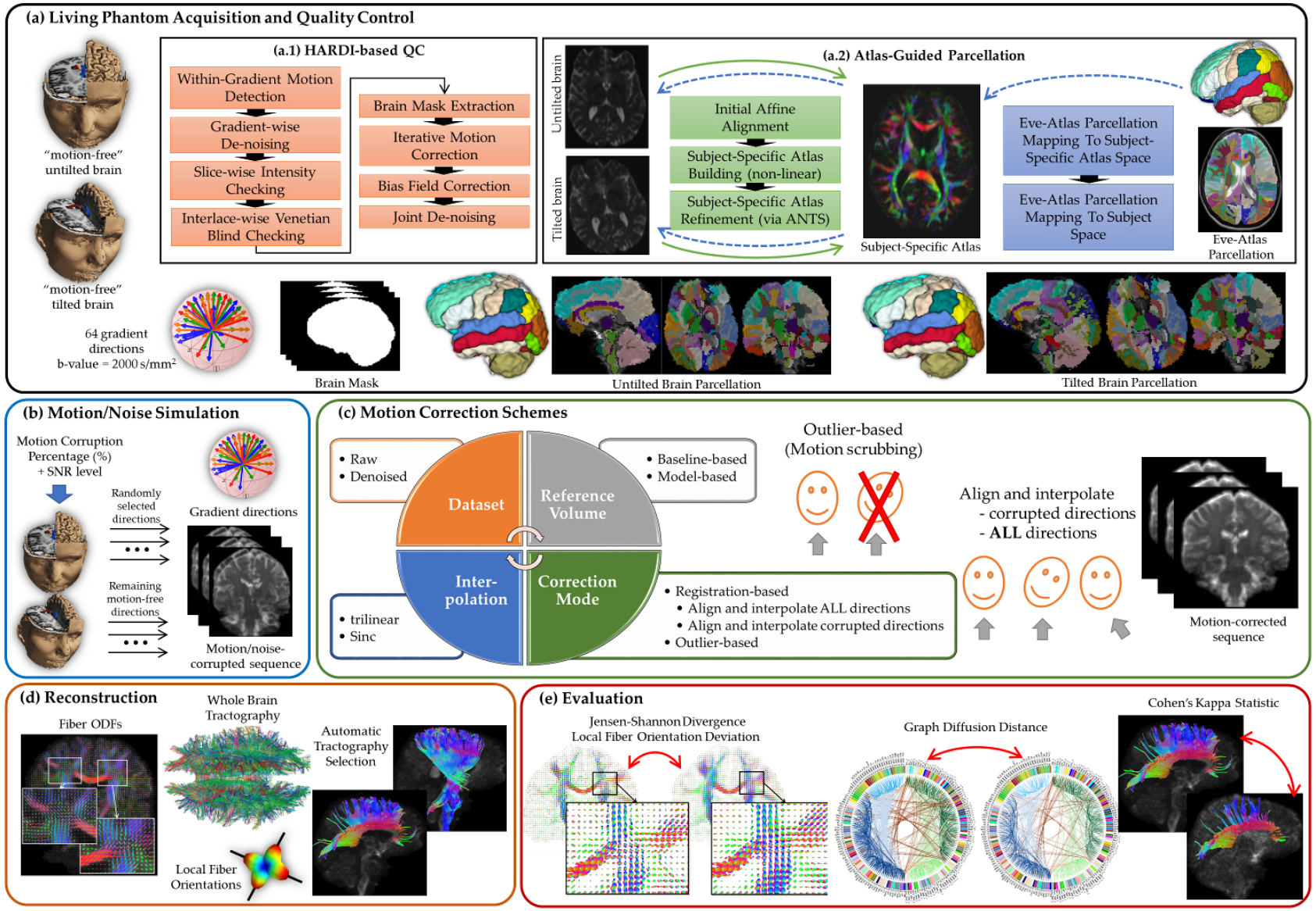


Figure 1. A comprehensive experimental framework for subject motion simulation to systematically evaluate the outcome of different motion correction choices commonly used by the scientific community on HARDI-based reconstructions and tractography.

Motion artifacts range from physiological motion (e.g., cardiac pulsation and respiration), to physical (voluntary or involuntary) *bulk movement* by the patient (**Benner et al.**, 2011). Physiological motion can be controlled by gating or in the sequence design (**Nunes et al.**, 2005), but the patient bulk movement during the diffusion-encoding gradient pulses leads to severe signal perturbation (**Rohde et al.**, 2004; **Chang et al.**, 2012; **Mohammadi et al.**, 2010), which results in a significant signal phase shift or signal loss (**Tournier et al.**, 2011). The effects of bulk motion are two-fold: *slow bulk motion* can cause misalignment of diffusion data between subsequent gradient applications (i.e., DWI-volumes), resulting in an underestimation of diffusion anisotropy (**Yendiki et al.**, 2014), whereas *fast bulk motion* during the application of a single diffusion gradient causes inhomogeneous signal dropout/attenuation artifacts in the diffusion-weighted images. This dropout effect arises due to signal dephasing within the voxels (**Benner et al.**, 2011; **Gumus et al.**, 2013), which is the very phenomenon that gives rise to the DW-MRI contrast, leading to an overestimation of diffusion anisotropy (**Yendiki et al.**, 2014). Although misalignment can be tackled by registration-based correction methods (**Sakaie and Lowe**, 2010), the signal dropout due to intragradient motion will persist (**Yendiki et al.**, 2014), where such images are identified and excluded from further processing and/or scheduled for reacquisition during the same scan (**Shi et al.**, 2009; **Porter and Heidemann**, 2009; **Benner et al.**, 2011; **Aksoy et al.**, 2011; **Gumus et al.**, 2013). Left uncorrected, motion-corrupted datasets introduce bias in the subsequent findings due to the induced variability of diffusion MRI measurements, while affecting the statistical properties of diffusion derived measures in heterogeneous brain regions.

1.2 MOTION CORRECTION CHOICES

The identification and elimination of slow bulk motion artifacts in HARDI data, which is characterized by a high b-value and low signal-to-noise (SNR) ratio, still remains a challenge. In order to allow correction approaches to proceed with reasonable accuracy, motion occurring between diffusion gradients can be treated as if it occurred all at once (**Oakes et al.**, 2005).

Motion effects can be reduced by real-time motion control during the acquisition (a.k.a. *prospective* motion correction) (**Herbst et al.**, 2012; **Kober et al.**, 2012; **Caruyer et al.**, 2013), where the acquisition and the source of motion are synchronized, so that the data is never corrupted. In addition, the development of accelerated acquisition methods (e.g., **Feinberg and Setsompop** (2013)) can reduce the duration of a scan to minimize the susceptibility of subject motion. A comfortable padding can also be used to minimize head motion while urging the participant to remain without movement (**Soares et al.**, 2013). Nonetheless padding is not always effective in studies involving infants (e.g., autism diagnosis **Alexander et al.** (2007)), where remaining still in the scanner may be more challenging. Nevertheless, prospective methods for motion correction might affect the acquisition time due to the reacquisition of motion-corrupted gradients (**Benner et al.**, 2011). Such methods might also require external optical tracking systems (**Aksoy et al.**, 2011), free-induction decay navigators (**Kober et al.**, 2012) or volumetric navigators (**Alhamud et al.**, 2012), which are not always available on current scanners (**Caruyer et al.**, 2013), coupled with the need of time-consuming calibration steps prior to their use (**Benner et al.**, 2011). Furthermore, rapid modification of diffusion gradients may induce eddy current artifacts (**Gumus et al.**, 2013), and there is no guarantee that the head will move back to the original position.

Motion compensation can also be performed as a postprocessing step after acquisition, i.e., *retrospective*, to guarantee voxel-wise correspondence between different DWIs referring to the same anatomical structure. A common practice is to heuristically select transformation parameter thresholds for detection of motion outliers, where registration and interpolation are applied to gradient directions that are claimed to be corrupted. Software packages for image-based registration of DWIs are becoming readily available, e.g., FSL-MCFLIRT (**Jenkinson et al.**, 2002, 2012), the Advanced Normalization Tools (ANTS) (**Avants et al.**, 2008), TORTOISE (**Pierpaoli et al.**, 2010) and BRAINSFit (**Johnson et al.**, 2007) employed in DTIPrep (**Oguz et al.**, 2014).

A typical retrospective motion correction algorithm involves two stages (**Sakaie and Lowe**, 2010): first, finding the global transformation parameters that would transform all DWIs to the same coordinate frame, and then, applying the estimated transformations to the diffusion data. Solving for the transformation parameters usually involves rigidly registering the DWIs to a reference volume representing the same anatomical structure, but without being contaminated by motion artifacts. Examples of such a reference include a T2-weighted image (**Rohde et al.**, 2004), or a nondiffusion-weighted image (a.k.a baseline with b-value = 0) due to its high SNR and lesser vulnerability to eddy current distortion (**Netsch and van Muiswinkel**, 2004), where the difference in intensity profiles is compensated for using normalized mutual information similarity measure. Another alternative is a model-based reference volume computed for each diffusion-weighted image based on tensor fitting (**Bai and Alexander**, 2008; **Ben-Amitay et al.**, 2012). Model-based motion correction implicitly assumes that the original position defined by the baseline volume is the reference position to be aligned to (**Sakaie and Lowe**, 2010). Recently, it has been shown that model-based motion correction becomes a more powerful choice for correcting higher b-value diffusion imaging, which does not contain enough anatomical features to be registered accurately (**Ben-Amitay et al.**, 2012).

Applying the estimated transformation parameters is performed using *interpolation*, which computes intensities at transformed voxel coordinates as a weighted sum of the scaled intensities at surrounding voxels. The diffusion gradient vectors are also reoriented to incorporate the rotational component of subject motion (**Leemans and Jones**, 2009). Interpolation is usually carried out by an exact fit of a continuously defined model to discrete data samples. Nonetheless, this exact fit is less appropriate when data is noise-corrupted, since the model is forced to fit the noise too. Although using regularized interpolation can tackle noisy data, it is only preferable to applying denoising followed by standard interpolation under the assumption that the signal is a *stationary Gaussian process* (**Ramani et al.**, 2010); a situation that is not applicable for diffusion-weighted images, which are contaminated by Rician noise. Based on the central limit theorem, the (weighted) average of a large set of i.i.d. samples tends to follow a normal distribution. Thus, interpolation between Rician distributed samples might change the distribution towards a Gaussian PDF (**Veraart et al.**, 2013). We can, therefore, argue that the denoising process decreases the effect of standard interpolation on altering the underlying data distribution.

Another retrospective approach is to cast motion correction as an outlier rejection process, ranging from simply excluding one or more gradients bearing strong motion artifacts beyond acceptable levels of motion (**Benner et al.**, 2011; **Liu et al.**, 2010; **Soares et al.**, 2013), to statistical methods for detecting and discarding voxel-wise diffusion measurements as outliers (**Chang et al.**, 2005, 2012; **Pannek et al.**, 2012). Usually discarding entire scans (a.k.a *motion scrubbing* in functional MRI) either can be performed

by visual inspection or based on predefined thresholds on estimated motion parameters (**Yendiki et al.**, 2014). Nevertheless, removing gradients limits the ability to reconstruct crossing fibers, especially at small separation angles, due to the decreased number of distinct gradient directions needed for diffusion reconstruction. Moreover, scrubbing would introduce intersubject SNR and bias differences that would in turn affect subsequent statistical analysis (**Oguz et al.**, 2014). On the other hand, local exclusion of corrupted voxels for robust diffusion reconstruction in the presence of outliers is based on the deviation of the observed measurements (usually after motion correction) from the assumed diffusion model. Using these approaches for motion correction itself would mingle the effect of being an outlier to an assumed model with that of being corrupted due to motion. Further, local exclusion would lead to a different number of DWIs locally available for each voxel, complicating subsequent analysis to avoid bias due to different SNR values for different brain regions (**Oguz et al.**, 2014).

A common concern with retrospective methods in clinical studies, whether registration-based and/or outlier-based, is that data with different levels of motion will be subject to different schemes of motion correction. For instance, patients may show more motion than controls, or sedated subjects may be different from nonsedated. Applying different motion correction schemes could introduce a confounding factor for statistical analysis of populations that show different motion patterns. Nonetheless, eyeballing the acquired/preprocessed DWIs prior to proceeding to further analysis is highly recommended.

1.3 OBJECTIVE AND CONTRIBUTIONS

The lack of a comprehensive/rigorous quality control (QC) for HARDI datasets can result in considerable error and bias in subsequent analyses, which may affect research studies using these datasets. Most current software packages such as DTIPrep (**Oguz et al.**, 2014), TORTOISE (**Pierpaoli et al.**, 2010) and FSL (**Jenkinson et al.**, 2012), which offer various tools for processing and analysis of diffusion-weighted images, are mostly limited to DTI datasets, which are characterized by low b-values (i.e., higher SNR) and fewer gradients (i.e., shorter acquisition times). Nonetheless, special care is needed for HARDI datasets due to their low SNR and longer acquisition times, which increase the likelihood of subject motion. As a part of a thorough pipeline for HARDI-QC, this paper addresses the motion correction aspect for *slow bulk motion* where users often do not fully understand the consequences of different types of correction schemes on the final analysis, and whether those choices may introduce confounding factors when comparing populations. Therefore, the presented work is directed towards clear guidelines and recommendations to the standard users in practical settings.

The optimal preprocessing pipeline for HARDI sequences remains an open question and a challenge for real data. Questions that might arise include: Is there a threshold that would identify a motion-corrupted volume? How sensitive are HARDI reconstructions to such a predefined threshold? What is the impact of various motion correction schemes on subsequent HARDI-based reconstructions and tractography? So far, these questions have received, surprisingly, little attention in various DW-MRI studies of clinical populations. This study, then, focuses on the effect of preprocessing schemes, in particular motion correction, commonly deployed as a postacquisition step, on succeeding steps. We propose a *comprehensive* experimental framework (see Figure 1) that enables making use of human brain HARDI data from a well-controlled motion experiment to simulate various degrees of motion/noise corruption. The comprehensiveness is related to the systematic evaluation of the outcome of different motion

correction choices commonly used by the scientific community on different DWI-derived measures. To our knowledge, this evaluation does not exist in the literature and has not been discussed in detail.

Choices for correction include exclusion or registration of motion corrupted directions, with different choices of interpolation, as well as the option of registration/interpolation of all directions versus corrupted directions only. The effect of denoising as a preprocessing step applied prior to motion correction is also investigated. Further, the choice of the reference volume used in the registration framework is also discussed. The comparative evaluation covers four metrics: (1) the similarity of fiber orientation distribution functions (fODFs) via Jensen-Shannon divergence (JSD), (2) the deviation of multiple fiber orientations at each voxel, (3) the global brain connectivity via Graph Diffusion Distance (GDD) and (4) the reproducibility of seven anatomically-defined fiber pathways via Cohen's Kappa statistics. On the basis of our findings, we recommend assuming that motion is inevitable, even subtle, in the acquired scans. Motion correction, therefore, needs to be applied to all gradient directions without relying heuristically on a threshold that determines a gradient direction to be claimed as motion corrupted.

2 MATERIALS AND METHODS

2.1 MOTION IS INEVITABLE: PROOF-OF-CONCEPT

To back up our assumption that motion is omnipresent, we analyzed data from three healthy human phantoms (males 30-40 years old) visiting each of the four clinical sites (Chapel Hill, Philadelphia, St. Louis and Seattle) as a part of the ACE-IBIS study (Autism Centers for Excellence, Infant Brain Imaging study (**Wolff et al.**, 2012)), using a total of six MRI systems (two sites using both research and hospital scanners). All study procedures were approved by the institutional review board at each clinical site, and informed, written consent was obtained for all participants. In addition, the traveling phantoms sign consent forms at each of the sites, as per their own institutional IRBs. The sites include the University of Washington, Seattle, the Washington University in St. Louis, the Childrens hospital of Philadelphia, and the University of North Carolina at Chapel Hill. Each subject was scanned twice on a 3T Siemens Tim Trio scanner[1] with a strict calibration of image acquisition parameters. Test-retest reliability at each site was established with two scans within 24 hours. The scans were acquired within one week to guarantee that there were no major brain changes over time. The scanning environment was well controlled. Comfortable padding was used to minimize head motion and patients were urged to remain without movement. Eddy current was compensated for using a Twice Refocused Spin Echo (TRSE) protocol[2] , with FoV = $209mm$,

[1] The protocol used a GRAPPA parallel imaging factor of 2 and a partial Fourier factor of 3/4, which does indeed result in non-Rician noise distributions. However, the effect of the noise distribution is expected to be relatively small at a b values of 2000 s/mm2 (e.g., **Jones and Basser** (2004)) and we do not expect the difference in noise profile to affect our conclusions in terms of the motion correction schemes.

[2] In our analysis, we opt to using a prospective approach (a Twice-Refocused Spin Echo (TRSE) sequence) for eddy current compensation in order not to introduce any alignment-based preprocessing before running motion correction that could already have a confounding effect (otherwise we would work on resampled and interpolated images before motion detection). Further, we have eye-balled the FA map of the acquired sequences (prior to motion correction) where the prominent edge artifact (regions of anomalous diffusion contrast resulted from misregistration of dissimilar materials) that should be visible in case of Eddy current distortion was almost entirely absent in the TRSE images. To further support our decision, we used FSL-MCFLIRT (**Jenkinson et al.**, 2002) to provide the affine transformation matrix (i.e, 12 degrees of freedom) to detect the scaling and skewing parameters which might occur due to induced eddy current where the affine transformation matrix can be written as:

$$M = R \begin{pmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & s_z \end{pmatrix} \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}$$

76 transversal slices, thickness = $2mm$, $(2mm)^3$ voxel resolution, matrix size = 106×106, TR = $11100ms$, TE = $103ms$, one baseline image with zero b-value and 64 DWI with b-value at $2000\ s/mm^2$, with a total scan time of 12.5 minutes.

Initially, we ran automated Quality Control on the DWIs via DTIPrep (**Oguz et al.**, 2014), which includes among other steps interlaced correlation analysis for detection and removal of fast bulk motion within a single DWI volume, where no quantitative within-gradient motion was detected. Inspired by **Sakaie and Lowe** (2010), FSL-MCFLIRT (**Jenkinson et al.**, 2002) was then used to provide the rigid transformation matrix (i.e., 6 degrees of freedom) for each volume having the baseline image as the reference for motion correction and normalized mutual information as the cost function. It is worth noting that MCFLIRT employs a global-local hybrid optimization method for robust affine registration that is specifically tailored to brain images. Within a multiresolution framework, four scales were used (8, 4, 2 and 1 mm, i.e., supervoxel vs. subvoxel). At each scale, volumes were resampled after initial filtering to reduce the effect of noise. Further, we tested motion correction based on denoised HARDI sequences using the Joint Rician LMMSE filter (**Tristán-Vega and Aja-Fernández**, 2010) implemented as part of 3D Slicer (www.slicer.org), and found that the quantified motion with and without noise reduction was very similar.

To quantify motion, we used the magnitude of the translation vector (in mm) as well as the axis-angle rotation representation (in degrees) (**Yendiki et al.**, 2014). The boxplots in Figure 2 show the rotational and translational components of the motion being detected from a total of 24 DWI datasets, showing an average of 0.39^o rotation and $0.61mm$ translation. The graphs in Figure 2 illustrate the arbitrariness of a common calculation of percentage of motion correction to determine the number of affected scans, here shown as a function of thresholding on the estimated motion parameters. While this experiment attributes the estimated rotation and translation parameters to actual subject motion, a part of the experimentally obtained parameters may be due to some imaging/image-processing uncertainty and also to image differences due to anatomical properties of the object (e.g. tissue orientation) that make the images ”look” different even if they were perfectly aligned. To backup our analysis, we conducted another experiment where we contaminated a single DWI dataset with two independent realizations of rician noise such that the two generated DWI images were perfectly aligned because they were the exact same image. Then, we ran motion correction where all DWI images were aligned to the same baseline, we obtain similar motion parameters although we are registering two independent acquisitions of the same subject. We therefore conclude that the transformation parameter estimates from FSL-MCFLIRT (**Jenkinson et al.**, 2002) are resilient to noise and may primarily caused by subject motion during a DWI scan, or eventually also by relative motion between subject and scans if considering atifacts due to pulse sequence and scanner technology.

2.2 LIVING PHANTOM: ACQUISITION AND GOLD STANDARD GENERATION

Unlike conventional MRI, where realistic phantoms exist (**Collins et al.**, 1998), there is no widely acceptable realistic DWI phantom for the assessment of different processing tasks (**Tristán-Vega and**

where R is the rotation matrix, $s_x, s_y, s_z \in \mathbb{R}$ are the scaling parameters, $a, b, c \in \mathbb{R}$ are the skewing parameters and $t_x, t_y, t_z \in \mathbb{R}$ are the translation parameters. Table 4 reports the average and standard deviation (along all gradient directions per dataset) of the estimated transformation parameters where all datasets tend to have a unit scale with minimal skewing values. These values confirm the decision of bypassing eddy current compensation in our analysis.

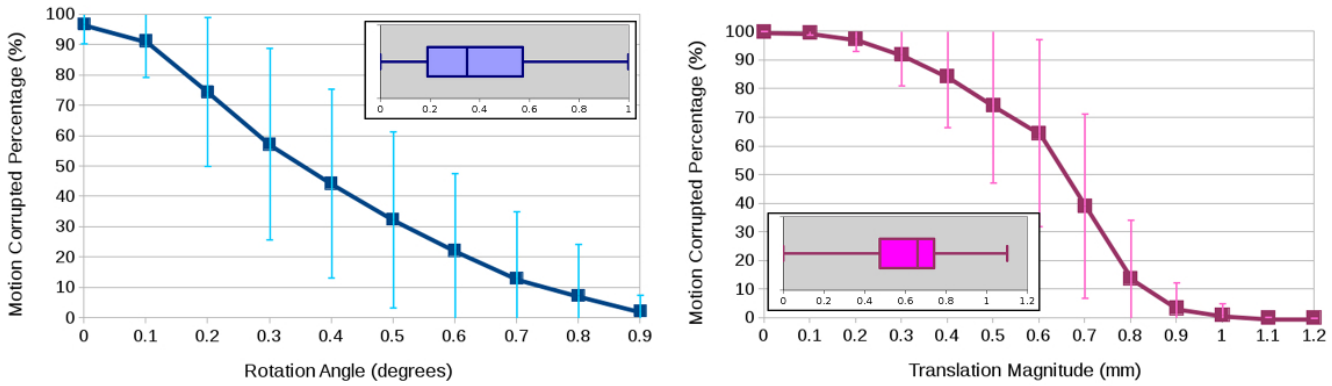


Figure 2. Average and standard deviation of the percentage of motion-corrupted gradient directions as a function of thresholding on the estimated rotation angle in degrees (left) and the estimated translation magnitude in mm (right) for three human phantoms scanned twice at four clinical sites. The boxplots show the overall statistics of estimated motion parameters.

Aja-Fernández, 2009, 2010). Existing phantoms simulate crossing sections in two and three dimensions, but they are not representative of white matter complex architecture with multiple fiber crossing, bending and branching. The lack of realistic phantoms motivates us to base our analysis on living (human) phantoms being scanned under well-controlled environments and propose a HARDI-based QC to yield motion- and noise-free datasets. Acquired DWIs were preprocessed (refer to Figure 1(a)) to obtain nearly *noise-free* and *motion-free* datasets according to the following pipeline, and therefore to be used as a *gold standard* for reconstruction and tractography.

2.2.1 HARDI-based Quality Control (QC): The QC process starts with identifying individual volumes having fast bulk (intra/within-gradient) motion using the signal dropout score proposed in **Benner et al.** (2011). The score was computed for each slice in each volume, where slices with a score greater than 1 were considered to have suspect signal dropout. Based on a zero-tolerance strategy, any volume having at least one slice with signal dropout was excluded from further analysis. It is worth noting that no within-gradient motion was detected in our phantom acquisitions. Each gradient was then independently denoised to reduce noise using the Rician LMMSE estimator with an 11×11 neighborhood (**Aja-Fernández et al.**, 2008) implemented in 3D Slicer (www.slicer.org) where the noise parameter is automatically estimated. Using DTIPrep (**Oguz et al.**, 2014), interslice brightness artifacts were detected via normalized correlation analysis between successive slices within a single DWI volume, where corrupted gradients were excluded before being streamed into the next steps. Further, interlaced correlation analysis (**Oguz et al.**, 2014) was used for detection and removal of venetian blind artifacts (seen when motion occurs between the interleaved parts of an individual gradient volume) and fast bulk motion within a single DWI volume, where no quantitative within-gradient motion was detected.

For each DW-MRI scan, iterative FSL-MCFLIRT (**Sakaie and Lowe**, 2010) was used to correct for intergradient subtle motion ($< 1^o$ rotation and $< 0.8mm$ translation), with the baseline volume as the reference for rigid alignment (i.e., six degrees of freedom with normalized mutual information as the cost function). The corresponding diffusion-weighting gradient vectors were reoriented accordingly (**Leemans and Jones**, 2009). To palliate the effect of spatial intensity inhomogeneities, N4 correction (**Tustison**

et al., 2010) was performed where the bias field was computed from the baseline volume and subsequently applied to all diffusion-weighted images. For further noise reduction, the Joint LMMSE (**Tristán-Vega and Aja-Fernández**, 2010) (www.slicer.org) was used to exploit the joint information from neighboring gradients from motion-corrected sequences. To avoid over-blurring, we used a $2 \times 2 \times 2$ neighborhood with six neighboring gradients.

2.2.2 Atlas-Guided Parcellation: For automated tractography selection and the quantification of whole brain connectivity, we defined a subject-specific unbiased atlas via DTI-derived data from HARDI sequences belonging to the same subject/phantom. This results in a tensor atlas, where we can define a detailed parcellation of neuroanatomical structures, and map it back to each raw scan. This reduces registration variability between each phantom data when defining the parcellation in subject spaces. The full process entails atlas creation and parcellation definition, as detailed in the following.

(a) Co-registration and Atlas Building:

To define a common reference space, our framework is centered around the creation of a DTI atlas, generated as an unbiased average atlas from the study dataset via a deformable atlas building strategy. Unbiased atlas building is used to provide one-to-one mapping between the image data and the template atlas, wherein the atlas is built from the population of data as the centered image with the smallest deformation distances. The overall registration framework, similar to what has been presented in **Verde et al.** (2013), proceeds in four steps: (1) image preprocessing via skull-stripping and tensor estimation, (2) affine alignment, (3) unbiased diffeomorphic atlas computation via GreedyAtlas module in AtlasWerks[3] software (**SCIInstitute**, 2014) and (4) a refinement step via symmetric diffeomorphic registration using the Advanced Normalization Tools - ANTS (**Avants et al.**, 2008).

Image preprocessing: A brain masking is first performed on the baseline images using FSL-BET2 (Brain Extraction Tool) (**Smith**, 2002) to remove all nonbrain parts of the image. BET2 uses a surface model approach to robustly and accurately carry out the segmentation. We then model tensors using the brain masks from the initial DWI datasets by using weighted least squares estimation, and then extract related scalar maps such as fractional anisotropy (FA) images.

Affine alignment: The second step applies affine registration of baseline images to a previously defined baseline template. A multithreaded, coarse-to-fine registration scheme using mattes mutual information metric is employed in that regard (**Johnson et al.**, 2007). The transformations are applied to curvature FA maps. The use of curvature FA as feature to derive registration has initially been presented by **Goodlett et al.** (2009). It is defined as the maximum eigenvalue of the Hessian of the FA image, therefore measuring image intensity curvature (second derivative) in the direction of largest curvature which acts like a 3D ridge detector. It is computed by convolution of the FA image with a set of Gaussian second derivatives with a fixed aperture, proportional to the size of the white matter structures. The curvature feature image proved to be an efficient detector of the 3D manifold skeleton of major fiber bundles which occur as tubular or sheet-like thin structures (similarly to the TBSS software), with the strongest response at their center.

[3] http://www.sci.utah.edu/software/atlaswerks.html

It is thus commonly used by our group when building population atlases to optimize correspondence of fiber tract geometries, and integrated into our freely distributed software package (**Verde et al.**, 2013). The curvature FA maps are thus mapped to this template space, and then the intensity is rescaled via histogram matching.

Atlas building: We then use an unbiased deformable atlas-building procedure (**Joshi et al.**, 2004) that applies large deformation diffeomorphic metric mapping transformations to these intensity rescaled mapped curvature FA images. The procedure relates individual datasets to the subject-specific atlas template space by means of nonlinear, invertible transformation. Tensor maps are transformed into the atlas space with tensor reorientation by the finite strain approach (**Alexander et al.**, 2001), taking into account both affine transformation and nonlinear deformation. The transformed tensor images are finally averaged using the Riemannian framework proposed in **Fletcher and Joshi** (2007), resulting in a three dimensional average tensor atlas.

Atlas refinement: An additional step is performed by direct symmetric diffeomorphic registration of initial FA images to the previously created DTI-FA atlas via the Advanced Normalization Tools - ANTS (**Avants et al.**, 2008). In our experience, this dual stage procedure has been shown to produce a sharper atlas with improved registration accuracy, most likely attributable to the use of local normalized cross-correlation as the image similarity metric. Final affine transformation and deformation fields are then available from subject space to atlas space.

(b) White Matter Parcellation:

We used the publicly available JHU-DTI-SS (a.k.a. "Eve") atlas described in **Oishi et al.** (2009). Defined as a single subject template, it includes both structural (T1w,T2w) and DTI images with white matter map parcellations, defining 176 hand-segmented core and peripheral regions of interest (ROIs). A multithreaded, coarse-to-fine diffeomorphic registration scheme using the cross-correlation metric via ANTS is employed on FA images between the Eve atlas and the subject-specific atlas. The computed deformation field is then applied to the Eve white matter label map. We can then map the parcellation, now defined in our subject atlas space, back to raw data in the initial image space, via the use of previously computed displacement fields. On a specific note, we concatenated the transformations from Eve atlas space to our initial images in order to directly map the parcellation and avoid the use of multiple interpolations. The white matter parcellation map is then defined both in the subject-specifc atlas space and in each individual subject space.

2.3 SUBJECT MOTION: BETWEEN SIMULATION AND CORRECTION

2.3.1 Human Motion Simulation As a pilot study, one human phantom was asked to be rescanned with his head tilted to simulate noticeable motion. The two datasets, after being QCed (see 2.2.1), were then used to construct motion-corrupted sequences (see Figure 1(b)). Based on the alignment of the baseline images of the two scans (original and tilted) using FSL-MCFLIRT, about 12^o of rotation and $7\ mm$ of translation were detected, whereas less than 1^o of rotation and $0.8mm$ of translation were found when aligning individual DWIs to their corresponding baseline image. It is worth noting that the

quantified motion between the acquired datasets (i.e., untilted versus tilted brains) can be classified as severe subject motion (**Ben-Amitay et al.**, 2012). We then arbitrarily considered the first out of the two scans as the *"motion-free"* sequence and used it as a reference for performance evaluation of different motion correction schemes. A random percentage of DW images ($10, 30, 50, 70$ and 90%, each with five distinct random sets of gradient directions) drawn from the second scan (tilted brain) were mixed with the first scan to construct 25 motion-corrupted datasets. Noisy sequences were generated by simulating Rician noise based on seven levels of SNRs from 4 to 20 (**Burdette et al.**, 2001), yielding 175 (5 experiments $\times$ 5 corruption percentages $\times$ 7 SNR levels) sequences.

2.3.2 Motion Correction Schemes Correction for subject motion involves four main decision variables (see Figure 1(c)), where each distinct combination of choices defines a motion correction scheme. The first variable is which reference volume is to be used in the alignment process. Two options are available (**Sakaie and Lowe**, 2010): *baseline-based* (e.g., **Rohde et al.** (2004)) and *model-based* (e.g., **Bai and Alexander** (2008); **Ben-Amitay et al.** (2012)). In this context, we use the FMAM (Fit Model to All Measurements) method (**Bai and Alexander**, 2008) where target images for registration were generated by first fitting the diffusion tensor to the DWIs, followed by diffusion simulation to provide target images of similar contrast to the DWIs. Notice that with $> 50\%$ motion corrupted, model-based reconstruction infers the spatial position/orientation from the gradients corresponding to the tilted brain due to its majority (i.e., gradients of the untilted brain are considered the motion-corrupted directions). Therefore, with model-based correction for sequences having more than 50% corrupted directions, the tilted brain was used as a reference for performance evaluation.

The second variable denotes whether the correction is performed based on *raw* or *denoised* DWIs, where the denoising process should not take into account joint information between diffusion gradients due to motion corruption. In our experiments, we denoised motion-corrupted sequences using the Rician LMMSE estimator (**Aja-Fernández et al.**, 2008), where each gradient was independently denoised.

The third variable entails the mode of correction, i.e., *registration-based* versus *outlier-based*. The first choice explores two options: (1) only aligning and interpolating the corrupted gradient directions to mimic the situation where a predefined motion parameter threshold is used to claim whether a DWI volume is motion-corrupted, (2) assuming there is always motion, forcing the alignment and interpolation of all DWI volumes. Note that both options involve the reorientation of the diffusion gradient vectors corresponding to the corrupted volumes (**Leemans and Jones**, 2009) to incorporate the rotational component of subject motion. In the second choice, i.e., outlier-based, we mimic the motion scrubbing approach, where we exclude the affected gradient directions from subsequent computations (i.e., diffusion profile reconstruction and tractography). Eventually, the interpolation step in the registration-based choices introduces the fourth variable where we study the impact of using trilinear and sinc interpolants.

It is important to stress that, in our motion simulation paradigm (i.e., randomly mixing DW volumes from a tilted-brain dataset), the identity of the motion-corrupted directions is known apriori without any use of parameters. This prior information is used via the outlier-based correction, as well as the interpolate corrupted directions choices. Nonetheless, in practice, this apriori information corresponds to heuristically set thresholds on the estimated motion parameters beyond which volumes are claimed to be

corrupted/outliers. For example, a rotation threshold of 0.5^o and a translation threshold of about one voxel spacing are set by default in DTIprep (**Oguz et al.**, 2014).

2.4 RECONSTRUCTION AND TRACTOGRAPHY

The reconstruction and whole brain tractography were computed for the motion corrected sequences as well as the motion-free sequences (gold standard generated in Section 2.2, followed by automatic tractography selection for seven major fiber bundles (see Figure 1(d)).

We employed the constrained spherical deconvolution (CSD) technique (**Tournier et al.**, 2007) to reconstruct fiber orientation distributions functions (fODFs) from the DWI data using the diffusion imaging Python (DiPy) library (**Garyfallidis et al.**, 2014). The fiber response function was estimated from the corpus callosum region, defined by the white matter parcellation (see 2.2.2), where it is known to have single fibers. In particular, we used an ROI at the center of the corpus callosum and of a radius that would include all its voxels. The response function was estimated in that region from the voxels with FA higher than 0.7.

Part of our analysis is based on comparing brain connectivity graphs, which are represented as weighted graphs and computed from fiber tractography results. Whole brain tractography was performed using the EuDX deterministic tracking technique (**Garyfallidis**, 2012), which is implemented in the DiPy library (**Garyfallidis et al.**, 2014), using random seeding inside the brain region and a turning-angle threshold of 30^o between two connected voxels (as suggested by **Parizel et al.** (2007) to provide sufficient fiber density while minimizing the number of spurious tracts).

To extract brain connectivity graphs from the fiber tractography results, we used the 176 core and peripheral ROIs defined in the white matter parcellation (see 2.2.2). Let N_{ij} denote the total number of streamlines connecting the $i-$th and $j-$th ROIs, each with length $l_k^{ij}\ \forall k\ \in\ [1, N_{ij}]$, and the edge weights w_{ij} computed as follows (**Hammond et al.**, 2013b): $w_{ij} = \frac{1}{N_{ij}} \sum_{k=1}^{N_{ij}} \frac{1}{l_k^{ij}}$. The normalization by the tracts length gives a higher connection strength to short tracts to compensate for the signal attenuation as a function of tract length. It is worth noting that the concept of using the connection strength or other measures to weight the graph edges was previously discussed in several papers (e.g., **Kaiser** (2011); **Rubinov and Sporns** (2010)).

For tract-based analysis, an automatic tractography selection method was performed to select a subset of detected tracts from the whole brain tractography result corresponding to a specific white matter structure. Starting from the Eve-atlas-based white matter parcellation map defined in the subject space (see 2.2.2), the pass-through and not-pass-through volumes of seven fundamental fiber bundles (left and right hemispheres) were defined. To remove fibers that do not belong to the pathway of interest, we used the geometrical constraints specific for different fiber bundles as defined in **de Luis-García et al.** (2013), where the anatomical characteristics of these fiber bundles are defined in **Jellison et al.** (2004). We report the matching results from seven major fiber bundles: corpus callosum (CC), cingulum of the cingulate gyrus (CG), corticospinal tract (CST), fornix (FX), inferior fronto-occipital tract (IFO), inferior longitudinal fasciculus (ILF) and uncinate fasciculus (UNC).

2.5 MOTION CORRECTION CONSEQUENCES: EVALUATION METRICS

The influence of various motion correction choices on subsequent reconstruction and tractography is evaluated according to voxel-based, global connectivity-based as well as tract-based metrics (see Figure 1(e)), detailed as follows.

2.5.1 Voxel-based Metrics In order to measure similarities between the original motion-free fODFs and the fODFs corresponding to the motion corrected images, we use the Jensen-Shannon divergence (JSD), which has been used to quantify differences between ODFs in various studies, e.g., **Chiang et al.** (2008); **Cohen-Adad et al.** (2011). Given two probability distributions P and Q, the JSD metric is defined as follows:

$$\mathrm{JSD}(\mathrm{P}\|\mathrm{Q}) = \frac{1}{2}\left[D_{\mathrm{KL}}(P\|M) + D_{\mathrm{KL}}(Q\|M)\right], \tag{1}$$

where $M = (P+Q)/2$ and D_{KL} is the Kullback-Leibler divergence. In our case, P and Q are represented as discrete distributions; therefore, the KL divergence takes the following form: $D_{\mathrm{KL}}(P\|Q) = \sum_i P_i \log \frac{P_i}{Q_i}$, where i is the discrete sample index. The JSD is for PDFs, but we compute it for normalized fODFs. We believe it is a good measure since it reveals subtle changes in PDFs so we can also keep track of changes in fiber volumes as well as orientations.

In addition to comparing fODFs, we are interested in quantifying local deviations in fiber orientations due to motion correction. Since brain connectivity maps are inferred by tracking local fiber orientations extracted from fODFs, distortions in those directions may lead to unreliable brain connectivity maps. Therefore, it is important to study the impact of motion correction on fiber orientations by directly comparing the local fiber orientations before and after correction. To that end, we use the mean angular deviation measure θ defined as follows:

$$\theta^k_{i,j} = \frac{180}{\pi}\left|\cos^{-1}(v^k_i . v^k_j)\right|, \quad \theta = \frac{1}{N}\sum_{k=1}^{N}\theta^k_{i,j}, \tag{2}$$

where N is the number of fibers compared, and v^k_i and v^k_j correspond to the orientations of fiber k, with and without motion correction. Before averaging the deviations, we match the fibers, such that fiber j has the closest direction to fiber i. If the number of fibers is different, we compare the fibers that are present in both voxels. For example, if we have three fibers after motion correction, whereas before correction there were only two, we compare the two closest fiber directions. The fiber orientations were computed using the DiPy peak extraction tool (with 0.4 relative peak threshold and 20^o minimum separation angle). We allowed up to five orientations in each voxel ($N = 5$). Since general image transformation does not necessarily preserve the original ordering of the fiber orientations, we first match the fibers based on the angular distance between each pair before computing the mean deviation.

2.5.2 Global Connectivity-based Metric Once the brain connectivity graphs were generated for the different sequences, we compared them by means of the graph diffusion distance (GDD) metric, which

has been proposed in **Hammond et al.** (2013a). The GDD is a novel distance measure for comparing weighted graphs, which takes into account the graph structure in addition to the edge weights, compared to the Frobenius norm, which is sensitive only to the edge weights. For an explanation of the differences between the GDD and the Frobenius norm, see the Barbell graph example in **Hammond et al.** (2013a).

The GDD is based on the diffusion maps framework (**Nadler et al.**, 2005). Let A_1 and A_2 be weighted adjacency matrices for N vertices, that is, A_1 and A_2 are symmetric, nonnegative, $N \times N$ real matrices with zeros along the principle diagonal. The (unnormalized) graph Laplacian operator is defined by $L_n = D_n - A_n$ (for $n = 1, 2$), where D_n is a diagonal degree matrix for the adjacency A_n, i.e., $(D_n)_{i,i} = \sum_{j=1}^{N}(A_n)_{i,j}$.

Given adjacency matrices A_1 and A_2, the columns of the Laplacian exponential kernels, $\exp(-tL_1)$ and $\exp(-tL_2)$, describe the different diffusion patterns centered at each vertex generated by diffusion up to time t under the two sets of weighted edges. Measuring the sum of squared differences between these patterns, summed over all the vertices, yields

$$\begin{aligned}\xi^2_{gdd}(A_1, A_2; t) &= \sum_{i,j}((\exp(-tL_1))_{i,j} - (\exp(-tL_2))_{i,j})^2 \\ &= ||\exp(-tL_1) - \exp(-tL_2)||^2_F \end{aligned} \tag{3}$$

where $||\cdot||_F$ is the matrix Frobenius norm. This defines a family of distance measures ξ, depending on the information propagation time t. The graph diffusion distance is given by ξ at the time of maximal difference, i.e., $d_{gdd}(A_1, A_2) = \max_t \xi_{gdd}(A_1, A_2; t)$. Here, we compute $d_{gdd}(A_1, A_2)$ by first diagonalizing L_1 and L_2 and using the exponential mapping. Then, Eq. (3) allows the computation of $\xi(A_1, A_2; t)$ for any fixed t. Finally, we optimize over t by a line search to give $d_{gdd}(A_1, A_2)$.

2.5.3 Tract-based Metric The spatial matching between motion-free and motion-corrected tracts was examined using Cohen's Kappa statistic (**Landis et al.**, 1977). The streamlines for a specific fiber tract (e.g., CST, IFO ...) are first converted to a binary volume with the same dimension and spacing of the raw DWI, where voxels that were occupied by at least one streamline were assigned a value 1. The two tracking results to be matched were then superimposed to identify: (1) voxels that did not contain streamlines in either result (NN), (2) voxels that contain streamlines in both results (PP) and (3) voxel that contain streamlines in one of the results (PN or NP)[4]. The Kappa statistic measures the level of agreement of the tracking results and corrects for agreement expected by chance. Hence Kappa is computed based on the probability of agreement $P(a)$ and the probability of expected agreement due to chance $P(e)$ as (**Hallgren**, 2012),

$$\kappa = \frac{P(a) - P(e)}{1 - P(e)}, \tag{4}$$

[4] P denotes positive and N denotes negative.

where,

$$P(a) = \frac{NN + PP}{PP + PN + NP + NN},$$
$$P(e) = \frac{(NP + PP)(PN + PP) + (NP + NN)(PN + NN)}{(PP + PN + NP + NN)^2}.$$

3 RESULTS

The fODFs and the whole brain tractography were computed for the $3,150$ motion corrected sequences (175 datasets $\times$ 18 correction schemes), as well as the motion-free sequences, followed by automatic tractography selection for seven major fiber bundles.

Voxel-based Metrics: The average JSD metric was computed using the fODF reconstruction from the "motion-free" dataset, not corrupted by mixing DWI directions from the tilted-brain scan, as a reference (i.e., presenting only subtle motion inherent to a scan). We differentiated between regions where multiple fibers were detected versus single fiber regions. Figure 3 shows the average JSD values for single and multiple fiber regions as a function of motion corrupted percentage for different SNR levels and as a function of SNR levels for different motion corrupted percentages. Figure 4 illustrates sample reconstructions from motion-free versus motion-corrected datasets for different corrupted percentages and different motion correction choices. Table 1 shows the effect of the denoising process prior to applying motion correction on the average JSD values for single and multiple fiber regions as a function of SNR levels for different motion corrupted percentages. Figure 5 shows the average deviation of local fiber orientations (for the first two dominant detected fibers per voxel) as a function of motion corrupted percentage, as well as SNR levels.

Global Connectivity Metric: Figure 6 shows the average graph diffusion distance (GDD) metric as function of both the corrupted directions percentage and the SNR levels. The metric compares the weighted connectivity graphs from the whole brain tractography of the "motion-free" dataset to that of the motion-corrected datasets. It is worth noting that the tractography of the tilted brain dataset is used as a reference for model-based corrections when the corrupted percentage exceeds 50%. Figure 7 visualizes the brain connectivity being represented circularly using the Circos software (**Krzywinski et al.**, 2009) where the parcellated structures (refer to Table 3 for their full names) are displayed on a connectogram representing left and right hemispheres symmetrically positioned along the vertical axis. The weighted connectivity matrix computed as described in 2.4 was normalized to attain a unit maximum. Each entry in the normalized connectivity matrix corresponds to an interregion link with thickness proportional to the entry weight. To avoid dense visualization, all entries with weight < 0.15 were discarded.

Tract-based Metric: Table 2 shows the average Cohen's Kappa statistic computed for corpus callosum (CC), corticospinal tract (CST) and inferior fronto-occipital tract (IFO) (where other pathways showed similar trend) based on automatic tractography selection using whole brain tractography of raw datasets (denoised datasets showed similar trends due to the robust fODF estimation, yet their graphs were omitted due to space limitation). Figures 8-12 show sample tractography selections for the aforementioned tracts from the untilted motion-free dataset as well as selections from motion-corrected datasets with different

corrupted gradient directions. Correction choices shown include outlier-based (i.e., motion scrubbing) and registration-based (using baseline and model-based reference volumes). Also pass-through (in green) and not-pass-through (in red) volumes (i.e., isosurfaces) are shown. Their definitions along with the geometric constraints employed to remove fibers, which do not belong to the pathway of interest, can be found in **de Luis-García et al.** (2013).

4 DISCUSSION

In this section, we discuss the impact of different motion correction choices using local as well as global metrics.

4.1 VOXEL-BASED RESULTS

Heterogeneous regions are more affected by motion correction, showing larger average JSD in general when compared to the single fiber regions, regardless of the correction mode, interpolation scheme or reference volume employed (see Figure 3).

The impact of motion scrubbing (removing gradient directions) becomes more pronounced with more motion-corrupted directions when compared to registration-based correction (see Figure 3(a)). Meanwhile, the JSD values indicate minimal deformations in fODFs reconstructed for baseline-based correction at high SNR levels compared to model-based correction, whereas both choices show comparable average JSD values at low SNR levels. This complies with the conclusions presented in **Sakaie and Lowe** (2010).

Forcing the correction and interpolation of all gradient directions shows comparable performance compared to the correction and interpolation of only the corrupted directions (see Figure 3(a)). This observation discourages the choice of heuristic parameters on motion parameters beyond which directions are claimed to be corrupted and interpolated. Further, interpolation of all directions causes less impact on the reconstructed fODFs at low corrupted percentages ($< 50\%$). We can assume, therefore, that motion is omnipresent and can be corrected for by the alignment and interpolation of all gradient directions.

On the interpolation aspect of correction, the sampling theory suggests the sinc kernel as the ideal interpolation kernel; nonetheless, this gives rise to the Gibbs phenomena (i.e., ringing) due to kernel truncation. This explains the smaller fODF deformation when using trilinear interpolation compared to sinc interpolation. Trilinear interpolation, which is much faster, is probably sufficient for motion correction.

In Figure 3(b), one can observe the comparable impact of different motion correction choices at low motion corruption percentages ($< 30\%$). Whereas with higher motion corruption, a situation that is encountered in studies including infants, for example, motion scrubbing shows a significant impact on the reconstructed fODFs even at high SNR levels. This effect is more pronounced in regions with crossing fibers where the ability to resolve fiber crossings is deteriorated especially as the separation angle of the fibers decreases.

Table 1. The effect of denoising on the average +/- standard deviation of Jensen-Shannon divergence (JSD) values for single fiber regions and multiple fiber regions as a function of SNR levels for different motion corrupted percentages

Baseline-based Motion Correction (Single Fiber Regions)							
Corrupted Directions Percentage	**SNR Levels**						
30%	**4**	**8**	**10**	**12**	**14**	**16**	**20**
Interpolate Corrupted Directions (trilinear): raw	0.360240 +/- 0.045598	0.233751 +/- 0.057661	0.206071 +/- 0.056631	0.185168 +/- 0.053253	0.168574 +/- 0.050516	0.155135 +/- 0.047892	0.135391 +/- 0.043761
Interpolate ALL Directions (trilinear): raw	0.334243 +/- 0.059883	0.215623 +/- 0.062502	0.194581 +/- 0.060456	0.176716 +/- 0.055924	0.162974 +/- 0.052476	0.150333 +/- 0.050609	**0.133550 +/- 0.046793**
Interpolate Corrupted Directions (trilinear): denoised	0.352849 +/- 0.040460	0.231980 +/- 0.055656	0.202870 +/- 0.054822	0.184626 +/- 0.052840	0.167100 +/- 0.049906	0.153834 +/- 0.047516	0.135224 +/- 0.043508
Interpolate ALL Directions (trilinear): denoised	**0.329328 +/- 0.059218**	**0.211255 +/- 0.061444**	**0.190899 +/- 0.058799**	**0.175393 +/- 0.055475**	**0.161398 +/- 0.051984**	**0.150292 +/- 0.050625**	0.133699 +/- 0.046593
70%	**4**	**8**	**10**	**12**	**14**	**16**	**20**
Interpolate Corrupted Directions (trilinear): raw	0.410600 +/- 0.031331	0.318959 +/- 0.050478	0.286541 +/- 0.055425	0.252046 +/- 0.057168	0.230605 +/- 0.055139	0.214138 +/- 0.053860	**0.190215 +/- 0.048155**
Interpolate ALL Directions (trilinear): raw	0.402799 +/- 0.036878	0.314221 +/- 0.054400	0.284747 +/- 0.059198	0.252745 +/- 0.059908	0.233456 +/- 0.057853	0.216581 +/- 0.056865	0.192958 +/- 0.051800
Interpolate Corrupted Directions (trilinear): denoised	0.402564 +/- 0.029625	0.313242 +/- 0.049651	0.284339 +/- 0.052764	**0.250802 +/- 0.056173**	**0.231370 +/- 0.054575**	**0.208578 +/- 0.052739**	0.190779 +/- 0.047920
Interpolate ALL Directions (trilinear): denoised	**0.398054 +/- 0.038274**	**0.310018 +/- 0.054334**	**0.282856 +/- 0.057399**	0.251609 +/- 0.058948	0.234186 +/- 0.057342	0.210697 +/- 0.055260	0.194545 +/- 0.051417
Baseline-based Motion Correction (Multiple Fiber Regions)							
Corrupted Directions Percentage	**SNR Levels**						
30%	**4**	**8**	**10**	**12**	**14**	**16**	**20**
Interpolate Corrupted Directions (trilinear): raw	0.429747 +/- 0.014377	0.374981 +/- 0.028244	0.357056 +/- 0.032396	0.335591 +/- 0.035592	0.319182 +/- 0.036475	0.304666 +/- 0.037237	0.281099 +/- 0.037361
Interpolate ALL Directions (trilinear): raw	0.420579 +/- 0.017062	0.365135 +/- 0.029617	0.349272 +/- 0.032369	0.330066 +/- 0.034023	0.316609 +/- 0.034180	0.300648 +/- 0.035244	**0.278706 +/- 0.034553**
Interpolate Corrupted Directions (trilinear): denoised	**0.408211 +/- 0.013658**	0.361212 +/- 0.027941	0.345386 +/- 0.032180	0.328742 +/- 0.035251	0.314137 +/- 0.036364	0.300747 +/- 0.036693	0.279468 +/- 0.036236
Interpolate ALL Directions (trilinear): denoised	0.415004 +/- 0.016909	**0.357705 +/- 0.029817**	**0.342500 +/- 0.032743**	**0.325475 +/- 0.033581**	**0.312227 +/- 0.034740**	**0.300567 +/- 0.034321**	0.279097 +/- 0.033992
70%	**4**	**8**	**10**	**12**	**14**	**16**	**20**
Interpolate Corrupted Directions (trilinear): raw	0.441668 +/- 0.009944	0.406974 +/- 0.020601	0.394218 +/- 0.025143	0.371914 +/- 0.028719	0.359955 +/- 0.030159	0.349868 +/- 0.030453	0.326731 +/- 0.028539
Interpolate ALL Directions (trilinear): raw	0.438858 +/- 0.010475	0.400544 +/- 0.019045	0.387314 +/- 0.024494	0.369097 +/- 0.027007	0.358079 +/- 0.028551	0.348511 +/- 0.028087	0.327357 +/- 0.026651
Interpolate Corrupted Directions (trilinear): denoised	**0.428357 +/- 0.009559**	0.398647 +/- 0.020752	0.387550 +/- 0.024170	**0.364179 +/- 0.028253**	0.353670 +/- 0.029828	**0.342159 +/- 0.029064**	**0.324294 +/- 0.027985**
Interpolate ALL Directions (trilinear): denoised	0.434734 +/- 0.010609	**0.396470 +/- 0.019646**	**0.385026 +/- 0.023381**	0.364260 +/- 0.026451	**0.353628 +/- 0.027830**	0.343153 +/- 0.026418	0.326491 +/- 0.025983
Model-based Motion Correction (Single Fiber Regions)							
Corrupted Directions Percentage	**SNR Levels**						
30%	**4**	**8**	**10**	**12**	**14**	**16**	**20**
Interpolate Corrupted Directions (trilinear): raw	0.362824 +/- 0.044584	0.234789 +/- 0.057672	0.202436 +/- 0.054992	0.185053 +/- 0.052774	0.168921 +/- 0.051011	0.154412 +/- 0.048760	**0.137016 +/- 0.044368**
Interpolate ALL Directions (trilinear): raw	0.341529 +/- 0.055268	0.216935 +/- 0.061572	0.190353 +/- 0.057184	0.177338 +/- 0.053845	0.164345 +/- 0.051940	**0.151763 +/- 0.049245**	0.137129 +/- 0.044860
Interpolate Corrupted Directions (trilinear): denoised	0.355832 +/- 0.038952	0.233300 +/- 0.055329	0.200942 +/- 0.053886	0.183426 +/- 0.051732	0.168669 +/- 0.050471	0.156186 +/- 0.048501	0.137568 +/- 0.044271
Interpolate ALL Directions (trilinear): denoised	**0.337800 +/- 0.054166**	**0.214965 +/- 0.060579**	**0.188046 +/- 0.056021**	**0.174993 +/- 0.053265**	**0.163170 +/- 0.051443**	0.153666 +/- 0.049561	0.139043 +/- 0.045272
70%	**4**	**8**	**10**	**12**	**14**	**16**	**20**
Interpolate Corrupted Directions (trilinear): raw	0.437995 +/- 0.020398	0.401102 +/- 0.027875	0.395116 +/- 0.029187	0.392917 +/- 0.028751	0.394547 +/- 0.029421	0.394157 +/- 0.029765	0.393072 +/- 0.029627
Interpolate ALL Directions (trilinear): raw	0.424515 +/- 0.023969	0.389511 +/- 0.029832	0.385682 +/- 0.030147	0.385935 +/- 0.029025	0.389897 +/- 0.030270	0.390043 +/- 0.030282	0.389524 +/- 0.030102
Interpolate Corrupted Directions (trilinear): denoised	0.433672 +/- 0.019802	0.392704 +/- 0.026322	0.385278 +/- 0.027166	0.382104 +/- 0.026334	0.382479 +/- 0.027435	0.383047 +/- 0.028099	0.382639 +/- 0.027850
Interpolate ALL Directions (trilinear): denoised	**0.423772 +/- 0.023811**	**0.386144 +/- 0.029342**	**0.380362 +/- 0.029573**	**0.378202 +/- 0.028394**	**0.380366 +/- 0.029042**	**0.381810 +/- 0.029858**	**0.382169 +/- 0.029354**
Model-based Motion Correction (Multiple Fiber Regions)							
Corrupted Directions Percentage	**SNR Levels**						
30%	**4**	**8**	**10**	**12**	**14**	**16**	**20**
Interpolate Corrupted Directions (trilinear): raw	0.431485 +/- 0.013919	0.374890 +/- 0.027681	0.355161 +/- 0.030889	0.340584 +/- 0.034561	0.322469 +/- 0.035754	0.306188 +/- 0.036843	0.282984 +/- 0.037859
Interpolate ALL Directions (trilinear): raw	0.424731 +/- 0.016649	0.366541 +/- 0.029399	0.348776 +/- 0.031170	0.336014 +/- 0.033028	0.319378 +/- 0.034755	0.303846 +/- 0.035464	0.283342 +/- 0.035411
Interpolate Corrupted Directions (trilinear): denoised	**0.409946 +/- 0.012715**	0.361122 +/- 0.027406	0.344173 +/- 0.030754	0.331062 +/- 0.034374	0.316224 +/- 0.035922	0.302532 +/- 0.036504	**0.281484 +/- 0.036670**
Interpolate ALL Directions (trilinear): denoised	0.420245 +/- 0.016382	**0.360538 +/- 0.029426**	**0.342159 +/- 0.030712**	**0.328595 +/- 0.032733**	**0.314057 +/- 0.034207**	**0.300968 +/- 0.035322**	0.283460 +/- 0.034162
70%	**4**	**8**	**10**	**12**	**14**	**16**	**20**
Interpolate Corrupted Directions (trilinear): raw	0.452994 +/- 0.010220	0.417146 +/- 0.015161	0.410814 +/- 0.016068	0.406033 +/- 0.016404	0.408391 +/- 0.017131	0.407360 +/- 0.017038	0.402509 +/- 0.019258
Interpolate ALL Directions (trilinear): raw	0.441239 +/- 0.011757	0.402898 +/- 0.020790	0.398061 +/- 0.020777	0.397453 +/- 0.019135	0.401401 +/- 0.018495	0.401485 +/- 0.017602	0.395939 +/- 0.018965
Interpolate Corrupted Directions (trilinear): denoised	0.448519 +/- 0.009997	0.407138 +/- 0.014850	0.399052 +/- 0.015820	0.393196 +/- 0.015699	0.393496 +/- 0.016721	0.393950 +/- 0.016122	0.389538 +/- 0.017035
Interpolate ALL Directions (trilinear): denoised	**0.440633 +/- 0.011008**	**0.399459 +/- 0.020710**	**0.390507 +/- 0.021431**	**0.385709 +/- 0.020485**	**0.387375 +/- 0.020483**	**0.389474 +/- 0.019294**	**0.384891 +/- 0.019050**

Further, baseline-based motion corrections show minimal JSD values with higher corruption levels ($>$ 50%) when compared to model-based corrections, regardless of the interpolation scheme employed. The difference in performance between baseline-based and model-based becomes more significant as the SNR level increases.

The denoising process yields smaller JSD values for low SNR levels ($<$ 12) (see Table 1), while providing comparable performance for baseline-based and model-based motion correction choices. The slight decrease of JSD values for denoised datasets compared to the raw ones is due to the fODF reconstruction processes where we use the constrained spherical deconvolution (CSD) technique (**Tournier et al.**, 2007). In an iterative manner, the deconvolution process in CSD applies a nonnegativity constraint on the estimated fODFs as negative fiber orientation densities are physically impossible. This process provides fODFs estimates that preserve the angular resolution while being robust to noise. Yet, as a word of caution, the denoising process, when applied to motion-corrupted datasets, should not take into consideration the joint information from diffusion gradients since voxel-wise correspondence between different diffusion volumes is not guaranteed.

In Figure 4, one can observe the significant impact of motion scrubbing (i.e., outlier-based correction) on the reconstructed fODFs for mildly corrupted datasets (e.g., 30% corrupted directions). Further, it can

be noticed that with $> 50\%$ motion corruption, model-based reconstruction infers the spatial position from the gradients corresponding to the tilted brain due to its majority (i.e., gradients of the untilted brain are considered the motion-corrupted directions).

Due to the insufficient number of gradients and unbalanced sampling of the q-space, the impact of motion scrubbing on the estimated fiber orientations becomes evident as SNR decreases and/or corrupted directions increase (see Figure 5).

Although interpolating all directions versus corrupted directions reports comparable orientation deviation with lower impact on fractionally corrupted datasets ($< 50\%$), we still favor forcing such a process to all directions to avoid the ad-hoc process of thresholding motion parameters.

Nonetheless, one can notice the peaked performance of the orientation deviation at 50% corrupted directions for model-based motion correction choices. The explanation of this phenomenon is based on the fact that, with $> 50\%$ of the gradients being corrupted (i.e., corresponding to the tilted brain), the formed reference volumes would instead infer its anatomical structure from the tilted brain. For highly corrupted datasets, the gradients corresponding to the untilted brains become the corrupted directions (i.e., a 70% corruption will have a performance similar to the 30% case).

Model-based corrections display higher impact on the JSD of the reconstructed fODFs at higher levels of motion corruption, but such corrections have a smaller impact on the fiber orientation deviations especially when interpolating all directions (trilinear interpolant). This change of JSD metric implies an increase in the overall fODF volume when compared to the reconstructions from the *motion-free* dataset, yet the fODFs maintain the voxel-wise fiber crossing structure. This observation is more pronounced for fibers with the largest fiber volume fraction.

4.2 GLOBAL CONNECTIVITY-BASED RESULTS

Whereas there is a slight performance difference between GDD values computed based on raw datasets versus those from denoised dataset, thanks to the fODF reconstruction that is robust to noise contamination, one may observe consistent findings when GDD is compared to the JSD metric. In particular, the global brain connectivity is least affected by the motion correction step when forcing the alignment and interpolation of all gradient directions without setting a predefined threshold to claim corrupted volumes. There is a significant difference between GDD values obtained from trilinear interpolation compared to sinc interpolation. This implies that the impact of sinc interpolation on the fODFs, being encoded by the JSD metric, yields global brain connectivity that is different from the ”motion-free”-based brain connectivity.

Whereas the effect of motion correction is evident at higher corrupted percentages (except for motion scrubbing), one can notice the effect of noise where the impact of motion correction becomes more significant at low SNRs (< 12), while different correction choices (except motion scrubbing) render slight performance difference at high SNRs (> 12). Moreover, being consistent with different SNR levels, the baseline-based correction choices yield connectivity graphs with minimal deviations (smaller GDD) compared to their corresponding model-based choices.

On the contrary, motion scrubbing displays a different behavior. The GDD values from the scrubbed datasets, though maximal compared to the other correction choices, are decreasing with higher SNR levels for $< 50\%$ corrupted directions, but this behavior is soon changed to the opposition direction for $\geq 50\%$ corrupted directions, see Figure 6(b). This change of behavior is perceivable in Figure 6(a) where the GDD values are maximal at 50% corruption percentage for high SNR levels (> 12), whereas such a peak occurs even at low corrupted percentages (e.g., 30%) for low SNRs (< 12). This phenomenon can be explained as follows: with high percentage of motion-contaminated gradients, the scrubbing (outlier-based) option tends to produce an inadequate set of gradients for accurate fODF estimation due to the exclusion of too many gradients. This unbalanced sampling of the q-space, henceforth, biases the CSD process to converge to an incorrect solution, producing inaccurate fiber orientation and in turn imprecise brain connectivity. Hence, the increase of the GDD values with higher SNRs beyond 30% corrupted directions is due to having more short tracts connecting nearby region of interests while being assigned to larger weights in the graph construction step (see 2.4).

In Figure 7, one can observe the motion scrubbing behavior where the links become denser with higher corrupted percentages, implying the detection of more short tracts connecting nearby ROIs. On the other hand, the baseline-based choice reveals comparable connectograms to the motion-free ones while model-based counterpart tend to add more shorter tracts.

4.3 TRACT-BASED RESULTS

Being consistent with the results from the other metrics, motion scrubbing shows a significant decrease in the degree of tract agreement when increasing the percentage of motion corruption, which in turn leads to discarding more gradient directions. With $< 50\%$ corrupted directions, the tract agreement degree increases with higher SNR levels, yet such a trend changes with $\geq 50\%$ where shorter or no tracts being detected, which deviates from being anatomically realistic; see for example the top row of Figures 8-12 where tracts can be even missed even at 70% corruption. The CST and IFO tracts are good examples of long tracts that are not recovered by motion scrubbing beyond 10% motion corruption, see Figures 10 and 11. Nonetheless, the maximal agreement is achieved when aligning and interpolating all gradient directions to correct for motion regardless of the reference volume used in the registration process (i.e., baseline versus model-based). It can be observed in Figures 8-12 that model-based motion correction is able to recover longer tracts at high corruption percentages compared to the baseline-based motion correction.

5 CONCLUSIONS: GUIDELINES FOR MOTION CORRECTION IN HARDI ACQUISITIONS

Although there is excellent theoretical work on DWI acquisition parameters and ODF reconstruction schemes, as well as their effects on the quality and crossing fiber resolution, standard users lack clear guidelines and recommendations on the best ways to approach and correct for motion in practical settings. This work investigated motion correction using transformation and interpolation of affected DWI directions versus the exclusion of subsets of DWIs, and its impact on the reconstructed fODFs, local

fiber orientations, brain connectivity and detection of fiber tracts. The various effects were systematically explored and illustrated via living phantom data, leading to the general conclusion that motion, even subtle, exists in every acquired DW scan and special care is needed to correct for motion. In the following, we summarize the findings of our analysis, which might serve as guidelines for users in practice:

- Although least recommended, motion scrubbing (removing corrupted gradient directions) can be used in studies with well-controlled environments and involving not-in-pain adults or sedated subjects, where minimal subject motion is anticipated (i.e., $< 10\%$ motion corruption). Yet, this gradient removal should not result in unbalanced sampling of the q-space since the gradient distribution should be as uniform as possible on the sphere.
- Voxel-wise reconstructions, tractography and global brain connectivity are least affected by the motion correction step when forcing the alignment and interpolation of all gradient directions without setting predefined thresholds to claim corrupted volumes.
- Using voxel-wise reconstructions that are robust to noise, the denoising process can be considered unnecessary prior to applying motion correction. Nonetheless, if applied, the denoising algorithms should not take into account joint information from different diffusion gradients since voxel-wise correspondence is not guaranteed.
- Baseline-based correction choices can be used in studies involving voxel-wise scalars, which depend on the volume of the reconstructed ODFs, especially with highly motion-corrupted datasets.
- Model-based correction choices, on the other hand, are recommended for studies requiring the recovery and analysis of long tracts, e.g., CST and IFO, especially with highly motion-corrupted datasets.
- Trilinear interpolation, although much faster compared to sinc, is probably sufficient for motion correction, where the global brain connectivity is least affected.

One may wonder that using a gold standard which was obtained by motion correction (among other QC steps) using some of the methods under investigation could raise questions on reliability of the conclusions presented. Hence, in order to support the validity of the conclusions drawn from this study, we conducted the same set of experiments using the raw acquired data without performing any quality control. Figure 13 shows a sample result of the average JSD and local fiber orientation deviation metric for reconstructions based on gold standards generated from the QCed phantom datasets as well as the raw phantom datasets. Being consistent with the conclusions drawn from the reconstructions based on the QCed datasets, regions with crossing fibers are more affected by motion correction, showing larger average JSD in general when compared to the single fiber regions. The impact of motion scrubbing becomes more evident with more motion-corrupted directions when compared to the registration-based correction. Moreover, the peaked performance of the orientation deviation at 50% corrupted directions for model-based motion correction is also maintained. Further, forcing the interpolation of all gradients directions would have minimal impact on the reconstructions when compared to the choice of interpolating motion corrupted directions via setting a predefined threshold beyond which a direction is claimed to be corrupted.

6 LIMITATIONS AND FUTURE WORK

The primary message of this paper is that care should be taken in deciding the processing pipeline for any DW-MRI (esp. HARDI) at hand, this involves, for example, the acquisition protocol (i.e. less redundant gradients would discourage the choice of motion scrubbing) and the participating subjects (i.e. elderly in pain, infants, unsedated subjects versus healthy adults where variable motion severity levels are anticipated). Nonetheless, the presented analysis attains some limitations which can be outlined as follow:

- **One-subject analysis:** As a controlled motion experiment, we could use a scan session of subjects with repeated scans where the second shows bulk motion relative to the first one. The existing phantom data contains repeated scans taken in different sessions within 24 hours and hence they have to be seen as independent scans for the same subject. As a pilot study, we therefore asked one healthy volunteer to be scanned twice in a single scan session while tilting the head between the two scans. This enables us to mix gradients between the two scans from the same subject this cannot be done with the existing repeated independent scans. We understand that reporting our results with more than a pair of datasets (tilted and untilted brain) would support our analysis, and we will collect more scans with this experimental design in our future annual phantom scan sessions. Nonetheless, we think that this experiment, even with its limitations, contributes to establish an experimental framework that would guide the scientific community in systematically evaluating the outcomes of different preprocessing steps. In the future, we will prospectively plan to obtain more of such datasets, also including navigator shots for estimation of rotation, to extend this analysis.
- **Anatomical geometric correction:** Echo-planner imaging (EPI) distortion, in contrast to Eddy current that affects only diffusion-weighted images, would affect all images in the acquired sequence regardless of their level of diffusion sensitization. Hence, EPI distortion correction would involve acquiring additional data for either B0 mapping or a dedicated T1 or T2-weighted structural target. That's a primary reason behind ignoring EPI correction in most MRI processing pipelines (**Irfanoglu et al.**, 2012). With the availability of such additional data, EPI correction would involve nonlinear spatial warping that employ interpolation, a decision variable under investigation of the presented work. Hence, we favored to bypass this step in order not to inter-mingle interpolation due to motion correction and that of EPI correction. However, we think that the analysis/correction of inter-gradient spatial distortions, and its effect on ODF reconstruction, is an important issue which we together with the scientific community need to address.
- **Better gold standard generation:** The living phantoms were healthy volunteers who were aware of the whole process and were keen to remain without motion. Nonetheless, the investigation of prospective navigators is an promising idea for future work to provide different types of ground truth data and to get motion estimates directly from the scanner rather than only via postprocessing.

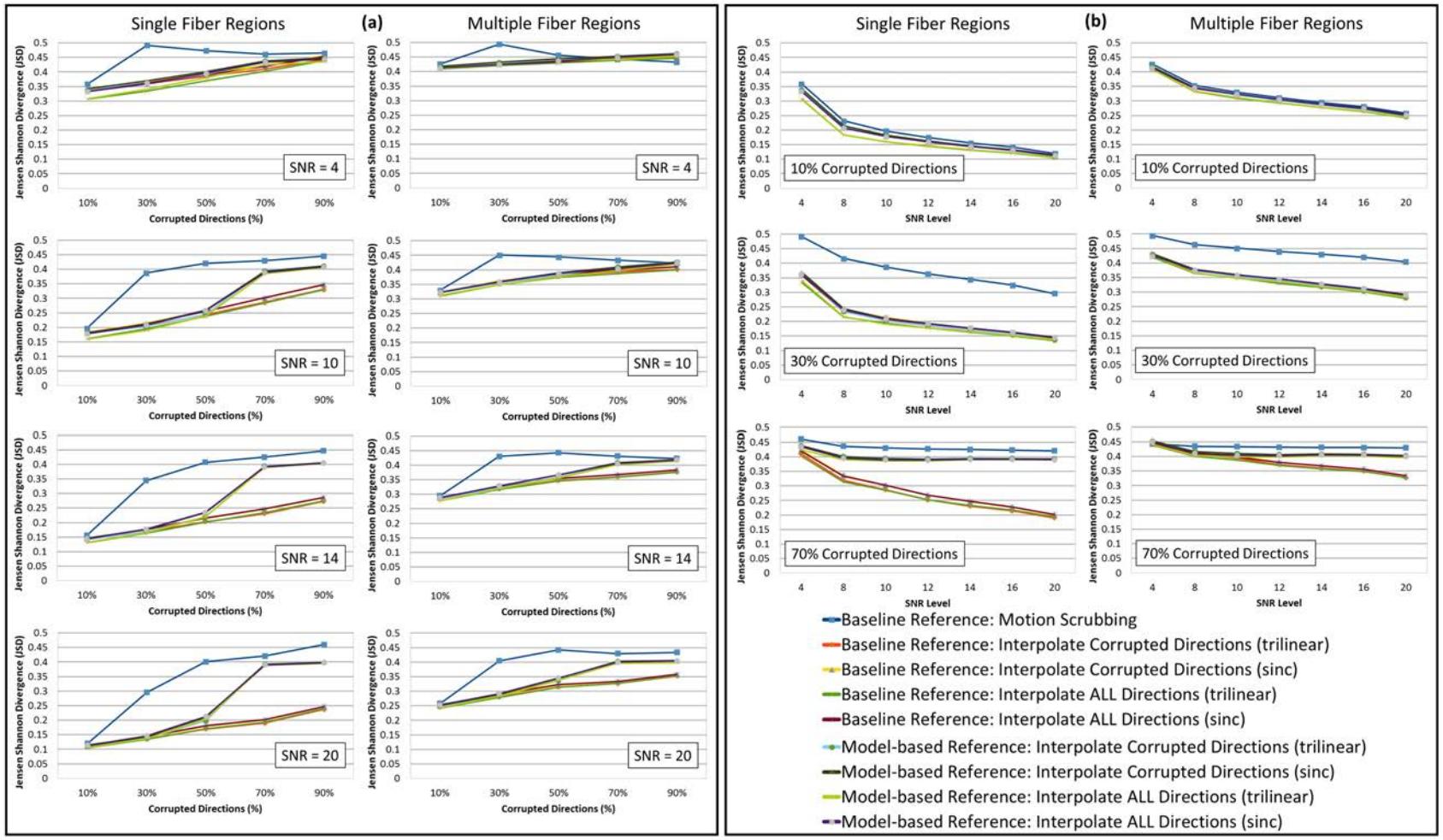


Figure 3. The average Jensen-Shannon divergence (JSD) values (lower is better) for reconstructions based on raw datasets (denoised ones share similar performance) as (a) a function of motion corrupted percentage for different SNR levels and (b) a function of SNR levels for different motion corrupted percentage. The first and third columns show JSDs single fiber regions while the second and fourth columns show such values for reconstructions based on multiple fiber regions. Notice the impact of motion scrubbing (removing gradient directions), which becomes more significant with more motion-corrupted directions when compared to registration-based correction. Further the impact of motion scrubbing is rendered evident for 10% corrupted gradients.

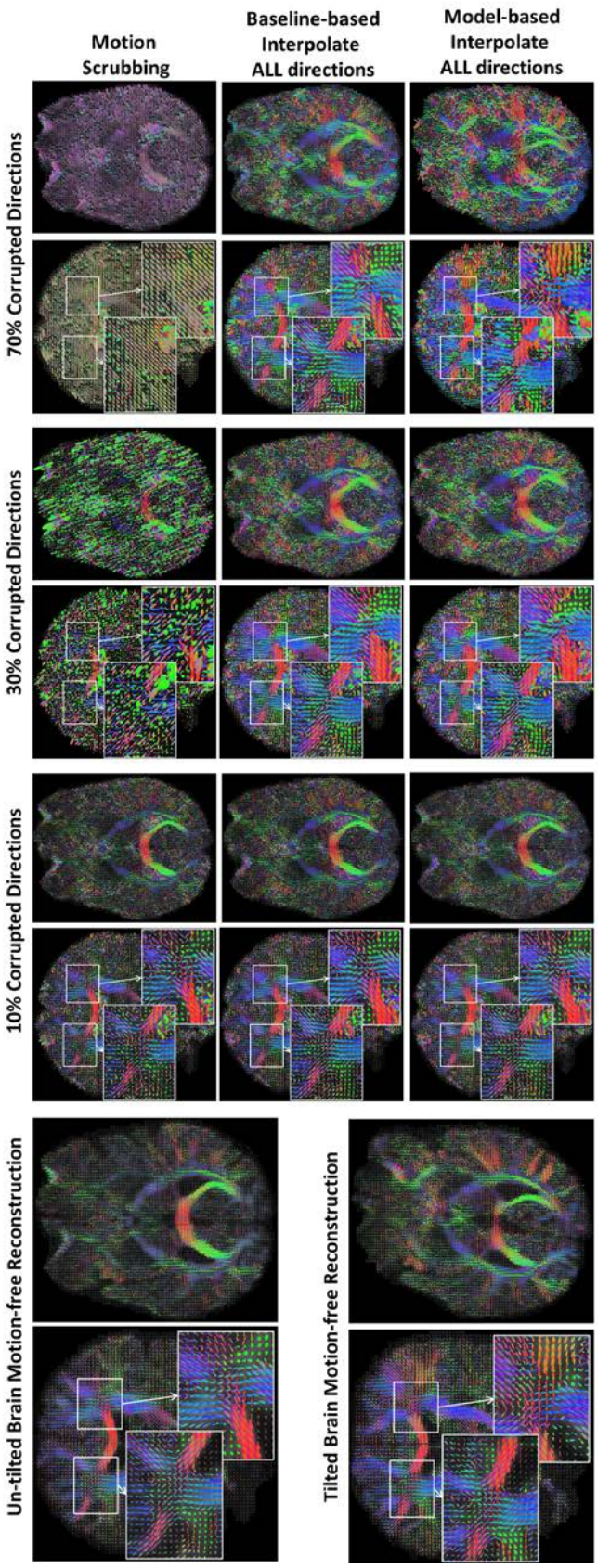


Figure 4. Sample fODFs reconstruction from untilted and tilted motion-free datasets as well as reconstruction from motion-corrected datasets with 10%, 30% and 70% corrupted gradient directions. Correction choices shown include outlier-based (i.e., motion scrubbing) and registration-based (using baseline and model-based reference volumes).

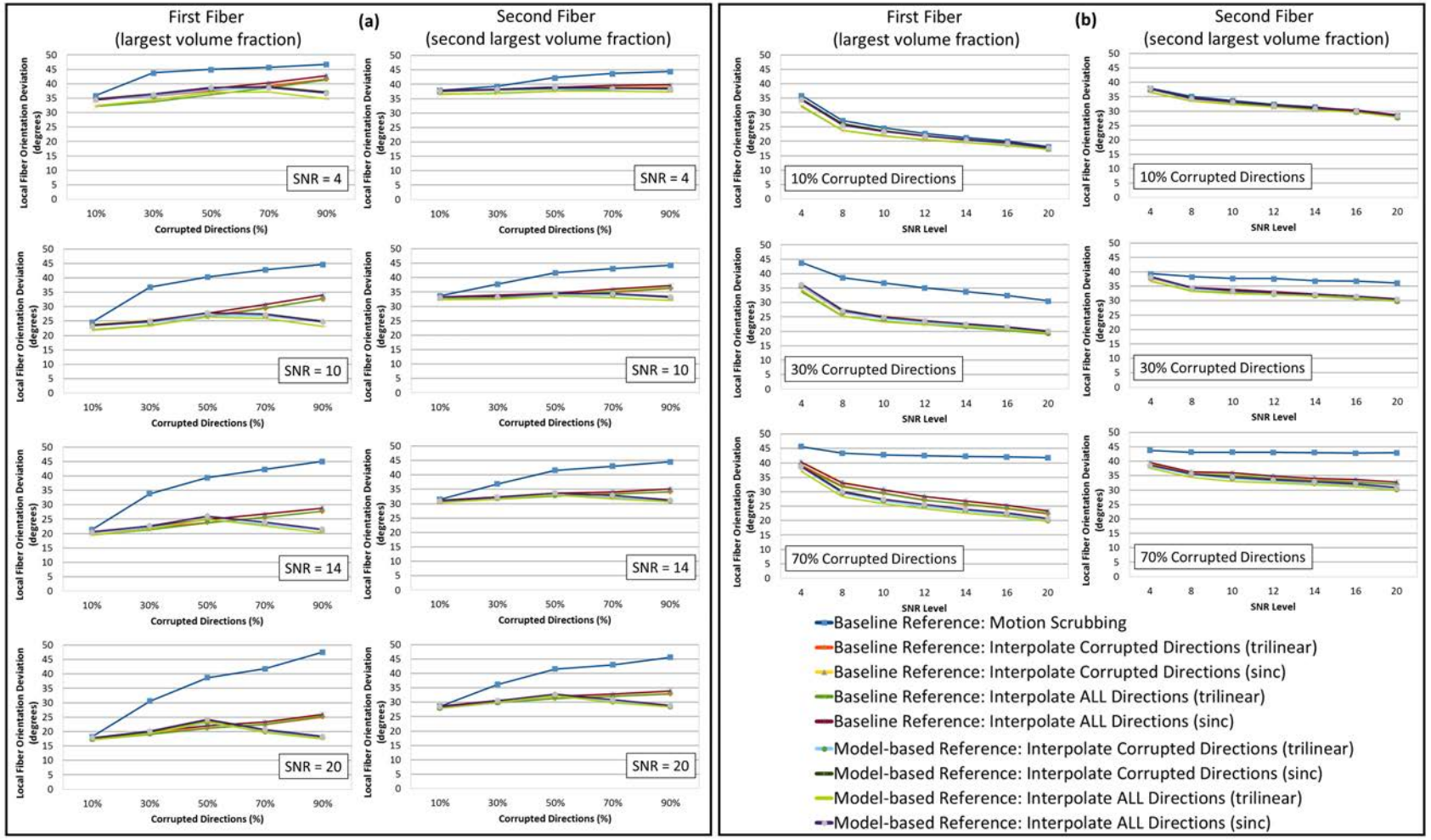


Figure 5. The average fiber orientation deviation (lower is better) for reconstructions based on raw datasets (denoised ones share similar performance) as (a) a function of motion corrupted percentage for different SNR levels and (b) a function of SNR levels for different motion corrupted percentage. The first and third columns show orientation deviation for the first detected fiber having the largest volume fraction while the second and fourth columns show such values for the second detected fiber having the second largest volume fraction. Notice that local fiber orientations are more affected by motion scrubbing as SNR decreases and/or corrupted directions increase.

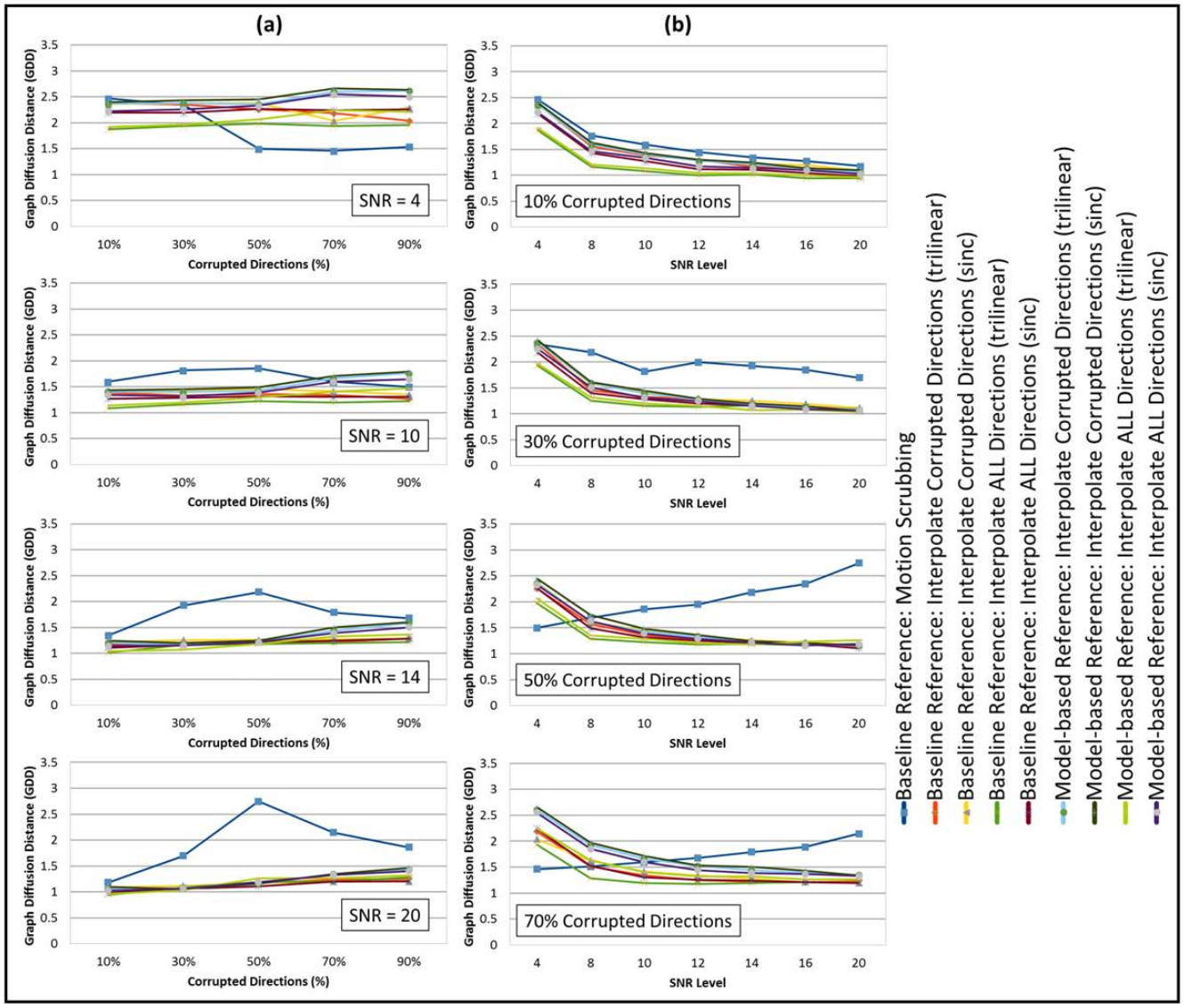


Figure 6. The average graph diffusion distance (GDD) (lower is better) for the whole brain tractography derived from the raw datasets (denoised ones share similar performance) as (a) a function of the corrupted directions percentage for different SNR levels and (b) a function of SNR levels for different motion corrupted percentages. Notice the different behavior displayed by motion scrubbing for $\geq 50\%$ corrupted directions, which due to having more short tracts connecting nearby region of interests while being assigned to larger weights in the graph construction step

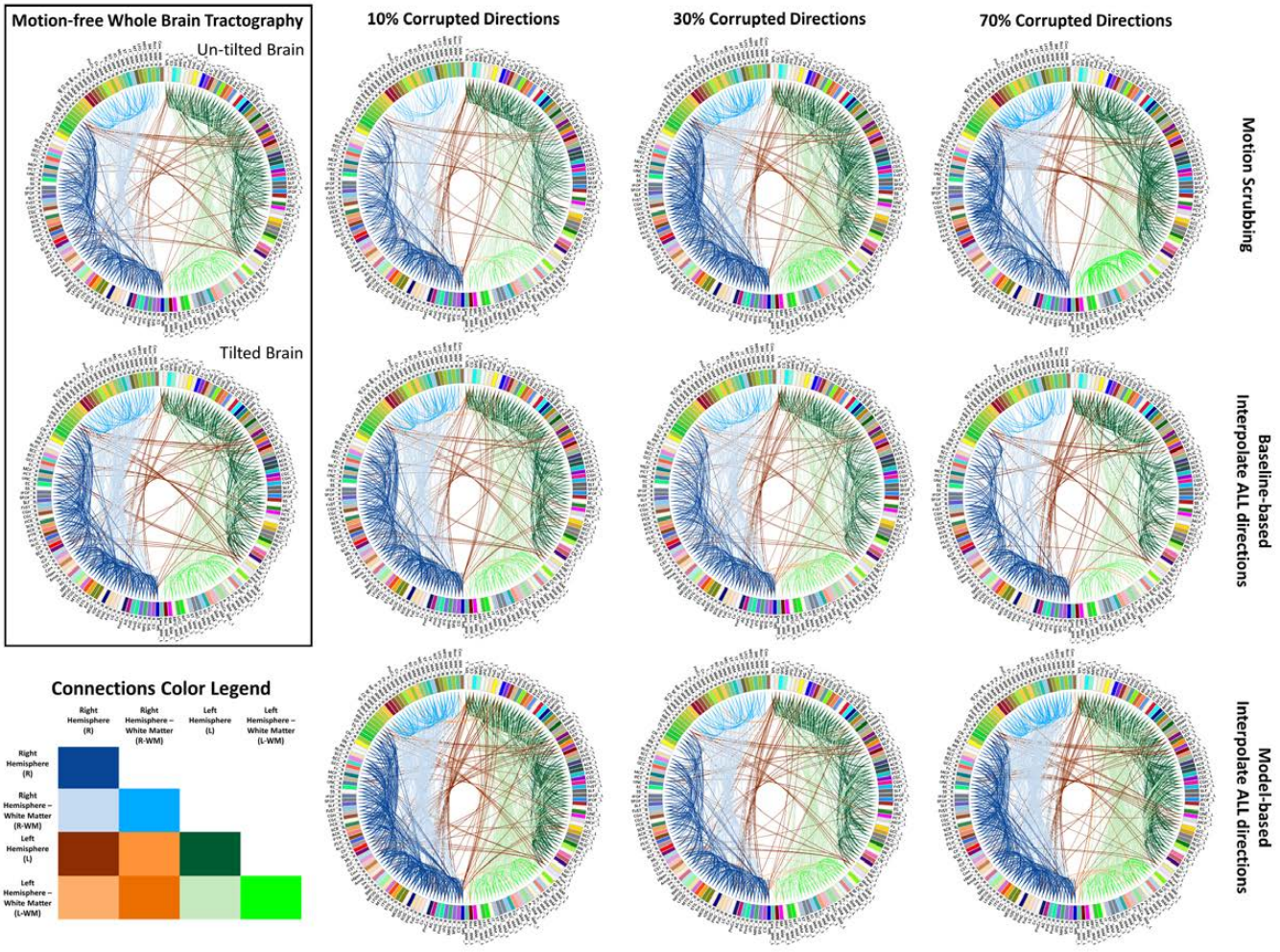


Figure 7. Sample reconstructed connectomic profile (i.e., connectogram) from untilted and tilted motion-free datasets as well as connectograms from motion-corrected datasets with $10\%, 30\%$ and 70% corrupted gradient directions. Correction choices shown include outlier-based (i.e., motion scrubbing) and registration-based (using baseline and model-based reference volumes). Notice the tendency of motion scrubbing to add more links between nearby ROIs at corruption percentages, implying the detection of more short tracts.

Table 2. The average Cohen's Kappa statistic (higher is better) of different anatomically-defined fiber pathways (other pathways show similar trend) based on automatic tractography selection based on whole brain tractography of raw datasets (denoised ones share similar performance) for different corrupted directions percentages.

Corpus callosum (CC)	**SNR Levels**						
10%	**4**	**8**	**10**	**12**	**14**	**16**	**20**
Baseline Reference: Motion Scrubbing	0.337569	0.512323	0.549884	0.583176	0.608084	0.623132	**0.653595**
Baseline Reference: Interpolate Corrupted Directions (trilinear)	0.371873	0.527897	0.560684	0.597392	0.610865	0.625269	0.641443
Baseline Reference: Interpolate ALL Directions (trilinear)	0.430934	**0.565719**	**0.604035**	0.612666	0.623576	**0.645637**	0.650286
Model-based Reference: Interpolate Corrupted Directions (trilinear)	0.372998	0.533756	0.56661	0.597997	0.610078	0.625306	0.645364
Model-based Reference: Interpolate ALL Directions (trilinear)	**0.432495**	0.56421	0.590059	**0.616106**	**0.628666**	0.643482	0.648159
30%	**4**	**8**	**10**	**12**	**14**	**16**	**20**
Baseline Reference: Motion Scrubbing	0.121185	0.240279	0.295196	0.342217	0.367067	0.397206	0.426858
Baseline Reference: Interpolate Corrupted Directions (trilinear)	0.344391	0.480168	0.510159	0.517193	0.519918	0.529172	0.536126
Baseline Reference: Interpolate ALL Directions (trilinear)	**0.397277**	0.508548	0.520689	0.522688	0.528747	0.52865	0.531048
Model-based Reference: Interpolate Corrupted Directions (trilinear)	0.34037	0.483498	0.511051	0.522758	0.53595	0.536208	0.54493
Model-based Reference: Interpolate ALL Directions (trilinear)	0.391228	**0.510303**	**0.522294**	**0.531674**	**0.541776**	**0.538634**	**0.546074**
50%	**4**	**8**	**10**	**12**	**14**	**16**	**20**
Baseline Reference: Motion Scrubbing	0.195245	0.234216	0.24072	0.239568	0.234356	0.228969	0.212593
Baseline Reference: Interpolate Corrupted Directions (trilinear)	0.32114	0.43416	0.463943	0.456179	0.456507	0.455066	0.441334
Baseline Reference: Interpolate ALL Directions (trilinear)	**0.354402**	**0.447591**	**0.476502**	**0.460336**	**0.464282**	0.455952	0.435115
Model-based Reference: Interpolate Corrupted Directions (trilinear)	0.308208	0.424219	0.454309	0.455936	0.456871	0.465054	**0.477503**
Model-based Reference: Interpolate ALL Directions (trilinear)	0.344133	0.443797	0.459322	0.459972	0.462699	**0.46546**	0.47007
70%	**4**	**8**	**10**	**12**	**14**	**16**	**20**
Baseline Reference: Motion Scrubbing	0.178267	0.178831	0.178042	0.163248	0.164247	0.158891	0.152246
Baseline Reference: Interpolate Corrupted Directions (trilinear)	0.314508	0.391142	0.408141	0.417553	0.420105	0.412033	0.405891
Baseline Reference: Interpolate ALL Directions (trilinear)	0.327764	0.395117	0.405833	0.415301	0.421643	0.402026	0.401629
Model-based Reference: Interpolate Corrupted Directions (trilinear)	0.290382	0.405799	0.440177	0.452235	0.479685	**0.479166**	**0.504993**
Model-based Reference: Interpolate ALL Directions (trilinear)	**0.330574**	**0.428458**	**0.455009**	**0.46612**	**0.482178**	0.478169	0.496215

Corticospinal tract (CST)	**SNR Levels**						
10%	**4**	**8**	**10**	**12**	**14**	**16**	**20**
Baseline Reference: Motion Scrubbing	0.255609	0.511582	0.59193	0.647802	0.681451	0.708302	0.733556
Baseline Reference: Interpolate Corrupted Directions (trilinear)	0.288567	0.568014	0.636962	0.674537	0.700027	0.714401	0.741004
Baseline Reference: Interpolate ALL Directions (trilinear)	**0.383782**	**0.673544**	**0.709478**	**0.732921**	**0.735943**	0.74033	**0.753895**
Model-based Reference: Interpolate Corrupted Directions (trilinear)	0.290589	0.561405	0.636853	0.673316	0.699448	0.713441	0.739494
Model-based Reference: Interpolate ALL Directions (trilinear)	0.377735	0.663838	0.703892	0.723948	0.732407	**0.741558**	0.751347
30%	**4**	**8**	**10**	**12**	**14**	**16**	**20**
Baseline Reference: Motion Scrubbing	0.08445	0.181213	0.223927	0.250415	0.278556	0.304034	0.338546
Baseline Reference: Interpolate Corrupted Directions (trilinear)	0.282041	0.52623	0.598869	0.626852	0.636781	0.643479	0.651485
Baseline Reference: Interpolate ALL Directions (trilinear)	**0.366403**	**0.61723**	**0.665015**	0.67397	0.663454	0.668005	0.671065
Model-based Reference: Interpolate Corrupted Directions (trilinear)	0.273331	0.521025	0.603001	0.639991	0.658976	0.671995	0.67978
Model-based Reference: Interpolate ALL Directions (trilinear)	0.347568	0.612903	0.655906	**0.678086**	**0.6801**	**0.692068**	**0.686448**
50%	**4**	**8**	**10**	**12**	**14**	**16**	**20**
Baseline Reference: Motion Scrubbing	0.167928	0.215559	0.227706	0.231256	0.238757	0.233776	0.237475
Baseline Reference: Interpolate Corrupted Directions (trilinear)	0.274988	0.47285	0.528849	0.56997	0.578886	0.581467	0.58602
Baseline Reference: Interpolate ALL Directions (trilinear)	**0.329036**	**0.540045**	**0.567838**	**0.60573**	**0.596126**	0.594135	0.599463
Model-based Reference: Interpolate Corrupted Directions (trilinear)	0.25215	0.466852	0.518179	0.553814	0.56339	0.584337	0.596757
Model-based Reference: Interpolate ALL Directions (trilinear)	0.300971	0.519456	0.551479	0.574403	0.578095	**0.603384**	**0.609796**
70%	**4**	**8**	**10**	**12**	**14**	**16**	**20**
Baseline Reference: Motion Scrubbing	0.209264	0.213673	0.214178	0.214527	0.219481	0.206522	0.210003
Baseline Reference: Interpolate Corrupted Directions (trilinear)	0.268415	0.449839	0.495255	0.54836	0.560483	0.565741	0.553228
Baseline Reference: Interpolate ALL Directions (trilinear)	**0.30485**	0.486243	0.531024	0.563493	0.569974	0.578043	0.561712
Model-based Reference: Interpolate Corrupted Directions (trilinear)	0.237249	0.43595	0.511829	0.537773	0.579497	0.591882	0.617357
Model-based Reference: Interpolate ALL Directions (trilinear)	0.304681	**0.493711**	**0.554253**	**0.575569**	**0.605093**	**0.610975**	**0.62865**

Inferior fronto-occipital tract (IFO)	**SNR Levels**						
10%	**4**	**8**	**10**	**12**	**14**	**16**	**20**
Baseline Reference: Motion Scrubbing	0.021174	0.164388	0.253179	0.360941	0.432971	0.522092	0.538292
Baseline Reference: Interpolate Corrupted Directions (trilinear)	0.036125	0.248734	0.350268	0.41586	0.467664	0.496367	0.525266
Baseline Reference: Interpolate ALL Directions (trilinear)	**0.066164**	**0.404309**	0.453207	**0.49633**	**0.497915**	**0.542272**	**0.55883**
Model-based Reference: Interpolate Corrupted Directions (trilinear)	0.036877	0.241989	0.355478	0.41203	0.450946	0.497335	0.530589
Model-based Reference: Interpolate ALL Directions (trilinear)	0.061719	0.397744	**0.476713**	0.481677	0.484301	0.532199	0.553205
30%	**4**	**8**	**10**	**12**	**14**	**16**	**20**
Baseline Reference: Motion Scrubbing	0.017605	0.032846	0.036941	0.054356	0.069675	0.096068	0.120592
Baseline Reference: Interpolate Corrupted Directions (trilinear)	0.021015	0.193547	0.30352	0.358774	0.391792	0.417395	0.448851
Baseline Reference: Interpolate ALL Directions (trilinear)	**0.040243**	**0.286059**	**0.375406**	0.407854	0.415109	**0.449185**	0.45364
Model-based Reference: Interpolate Corrupted Directions (trilinear)	0.019676	0.190149	0.298156	0.374981	0.394936	0.440116	0.450691
Model-based Reference: Interpolate ALL Directions (trilinear)	0.036533	0.269425	0.356155	**0.41894**	**0.417197**	0.448939	**0.457867**
50%	**4**	**8**	**10**	**12**	**14**	**16**	**20**
Baseline Reference: Motion Scrubbing	0.079802	0.096081	0.088847	0.08342	0.084004	0.078977	0.074762
Baseline Reference: Interpolate Corrupted Directions (trilinear)	0.023615	0.173538	0.255646	0.303376	0.355478	0.370352	**0.389862**
Baseline Reference: Interpolate ALL Directions (trilinear)	**0.037031**	**0.21137**	**0.287878**	**0.332891**	**0.361154**	**0.383244**	0.388926
Model-based Reference: Interpolate Corrupted Directions (trilinear)	0.017528	0.155743	0.226579	0.260404	0.306267	0.343361	0.363517
Model-based Reference: Interpolate ALL Directions (trilinear)	0.031326	0.187443	0.250407	0.282972	0.310855	0.353314	0.350486
70%	**4**	**8**	**10**	**12**	**14**	**16**	**20**
Baseline Reference: Motion Scrubbing	0.103989	0.111274	0.105098	0.103238	0.107333	0.11481	0.110321
Baseline Reference: Interpolate Corrupted Directions (trilinear)	0.02605	0.137873	0.203676	0.292856	0.312459	0.329589	0.36963
Baseline Reference: Interpolate ALL Directions (trilinear)	0.034815	0.169664	0.252561	0.300767	0.314673	0.347346	0.371886
Model-based Reference: Interpolate Corrupted Directions (trilinear)	0.021983	0.185256	0.299656	0.364072	0.433785	0.458342	0.477853
Model-based Reference: Interpolate ALL Directions (trilinear)	**0.043441**	**0.249726**	**0.354267**	**0.409827**	**0.471097**	**0.479884**	**0.485465**

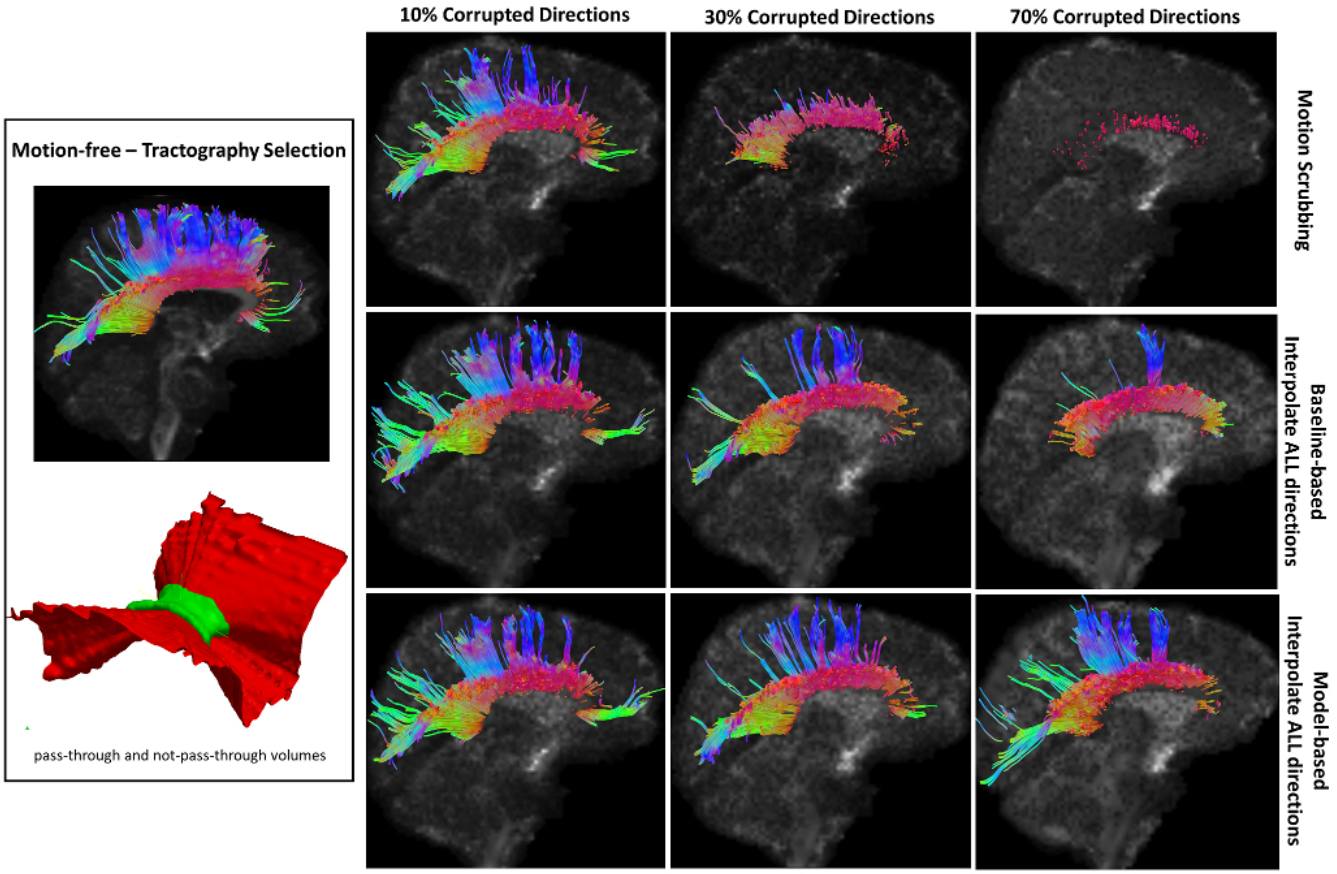


Figure 8. Sample tractography selection for the corpus callosum (CC) from the untilted motion-free dataset as well as selections from motion-corrected datasets with $10\%, 30\%$ and 70% corrupted gradient directions. Correction choices shown include outlier-based (i.e., motion scrubbing) and registration-based (using baseline and model-based reference volumes). One can observe the short tracts being detected by motion scrubbing at high corruption percentages due to the exclusion of too many gradient directions.

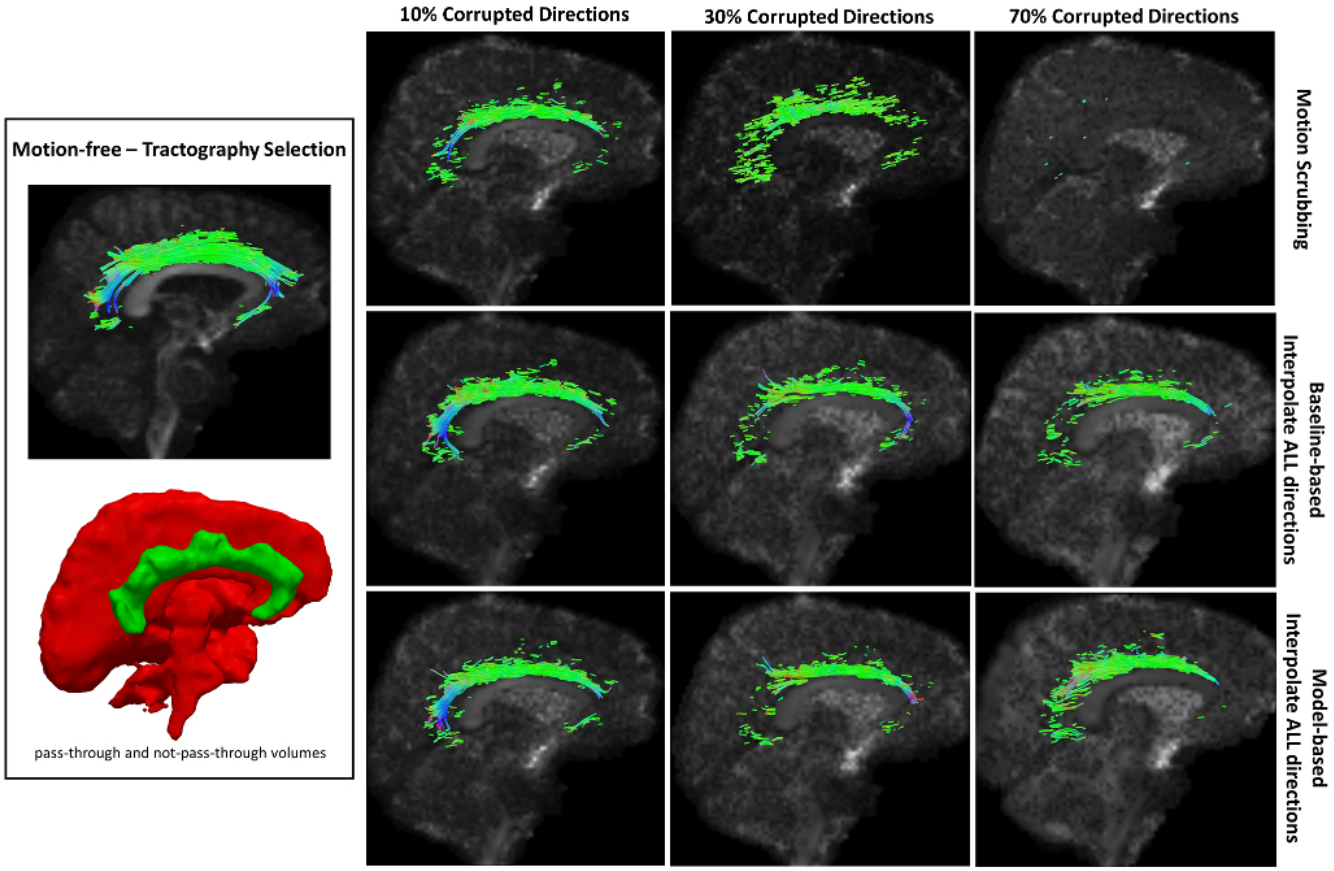


Figure 9. Sample tractography selection for the cingulum of the cingulate gyrus (CG) from the untilted motion-free dataset as well as selections from motion-corrected datasets with $10\%, 30\%$ and 70% corrupted gradient directions. Correction choices shown include outlier-based (i.e., motion scrubbing) and registration-based (using baseline and model-based reference volumes). Notice the inability of motion scrubbing to detect an anatomically realized CG at high corrupted percentages.

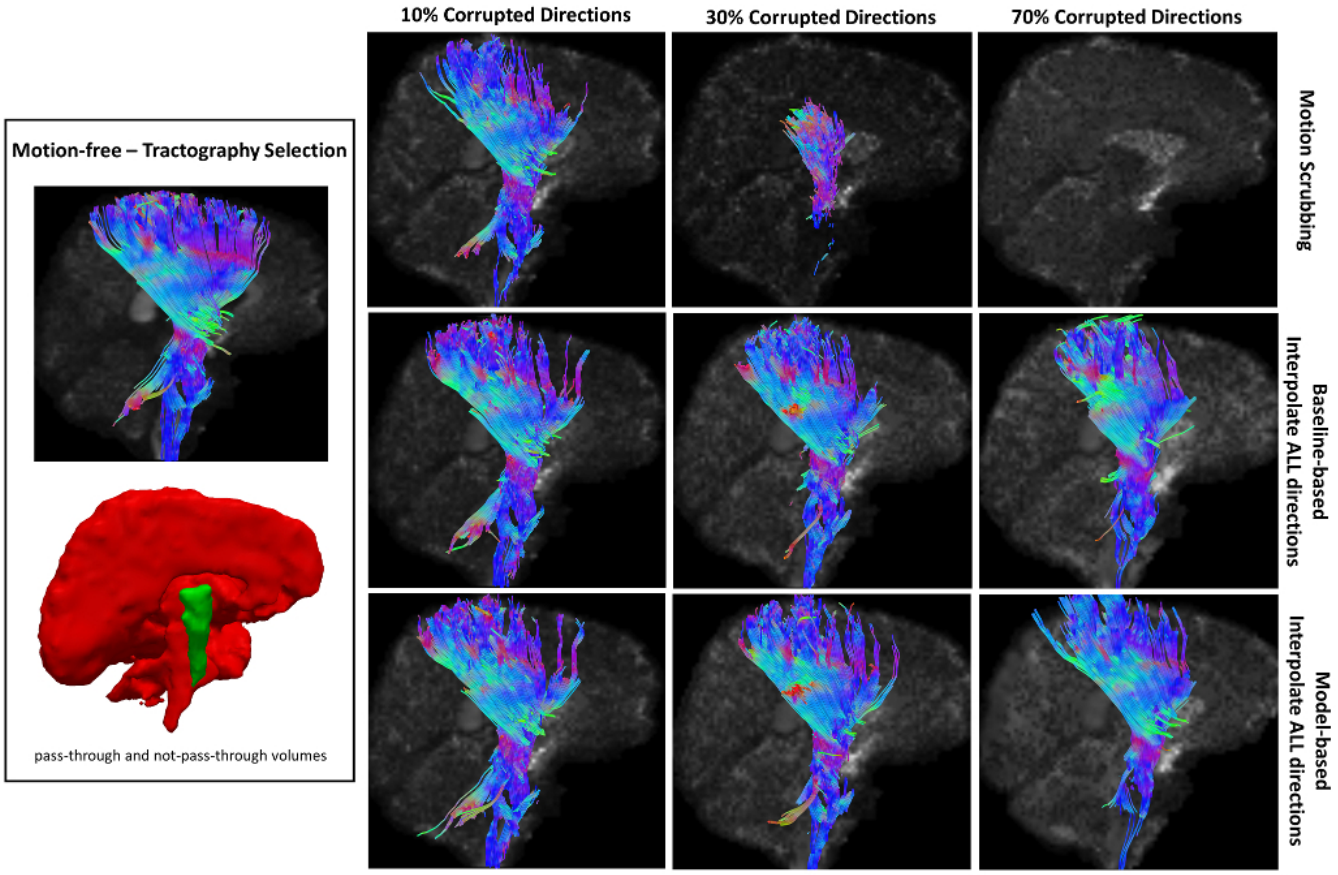


Figure 10. Sample tractography selection for the corticospinal tract (CST) from the untilted motion-free dataset as well as selections from motion-corrected datasets with 10%, 30% and 70% corrupted gradient directions. Correction choices shown include outlier-based (i.e., motion scrubbing) and registration-based (using baseline and model-based reference volumes). Note that motion scrubbing cannot recover long tracts such as CST beyond 10% motion corruption.

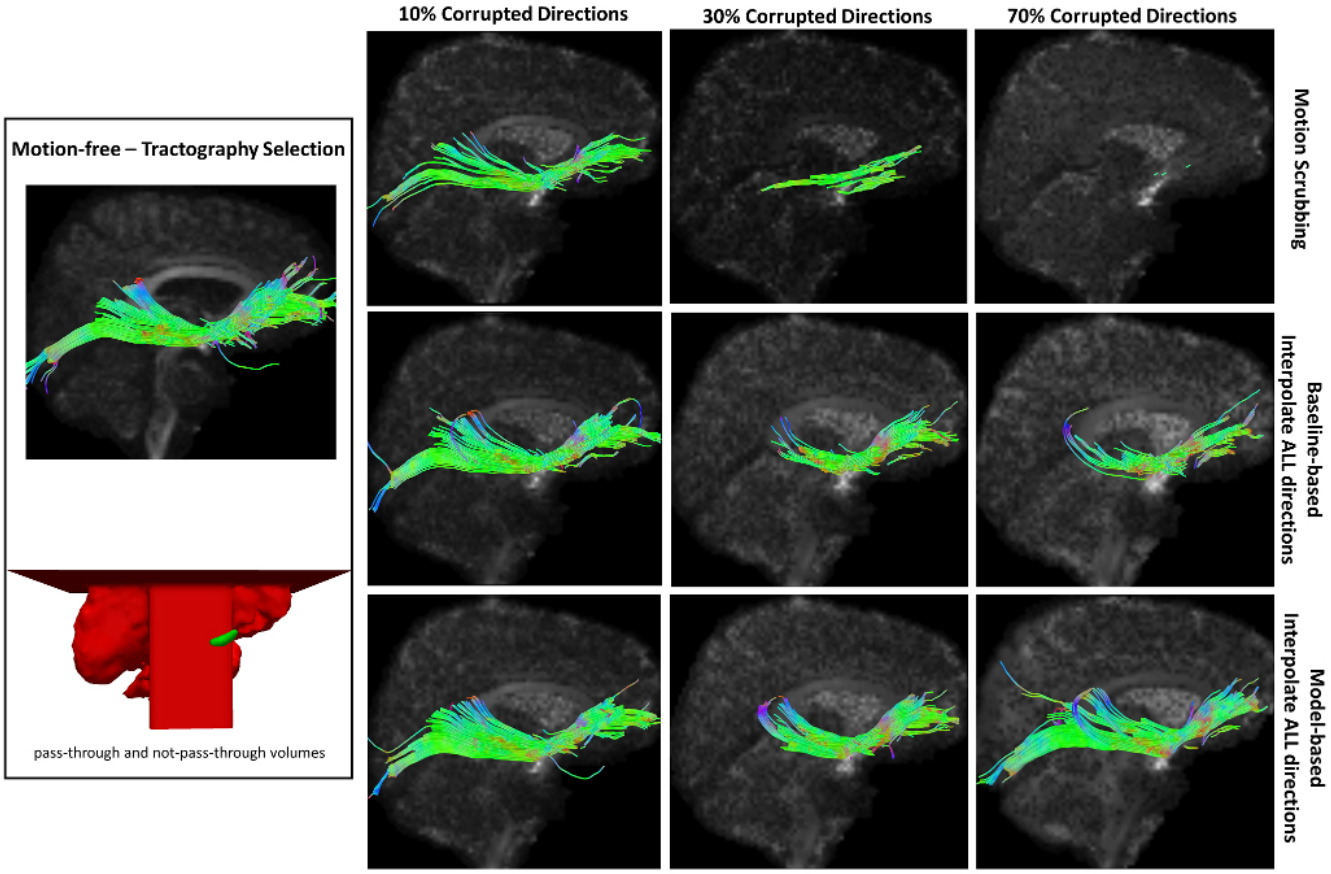


Figure 11. Sample tractography selection for the inferior fronto-occipital tract (IFO) from the untilted motion-free dataset as well as selections from motion-corrected datasets with $10\%, 30\%$ and 70% corrupted gradient directions. Correction choices shown include outlier-based (i.e., motion scrubbing) and registration-based (using baseline and model-based reference volumes). Note that motion scrubbing cannot recover long tracts such as IFO beyond 10% motion corruption. Further, motion-based motion correction tends to recover longer tracts at high motion corruption compared to baseline-based correction.

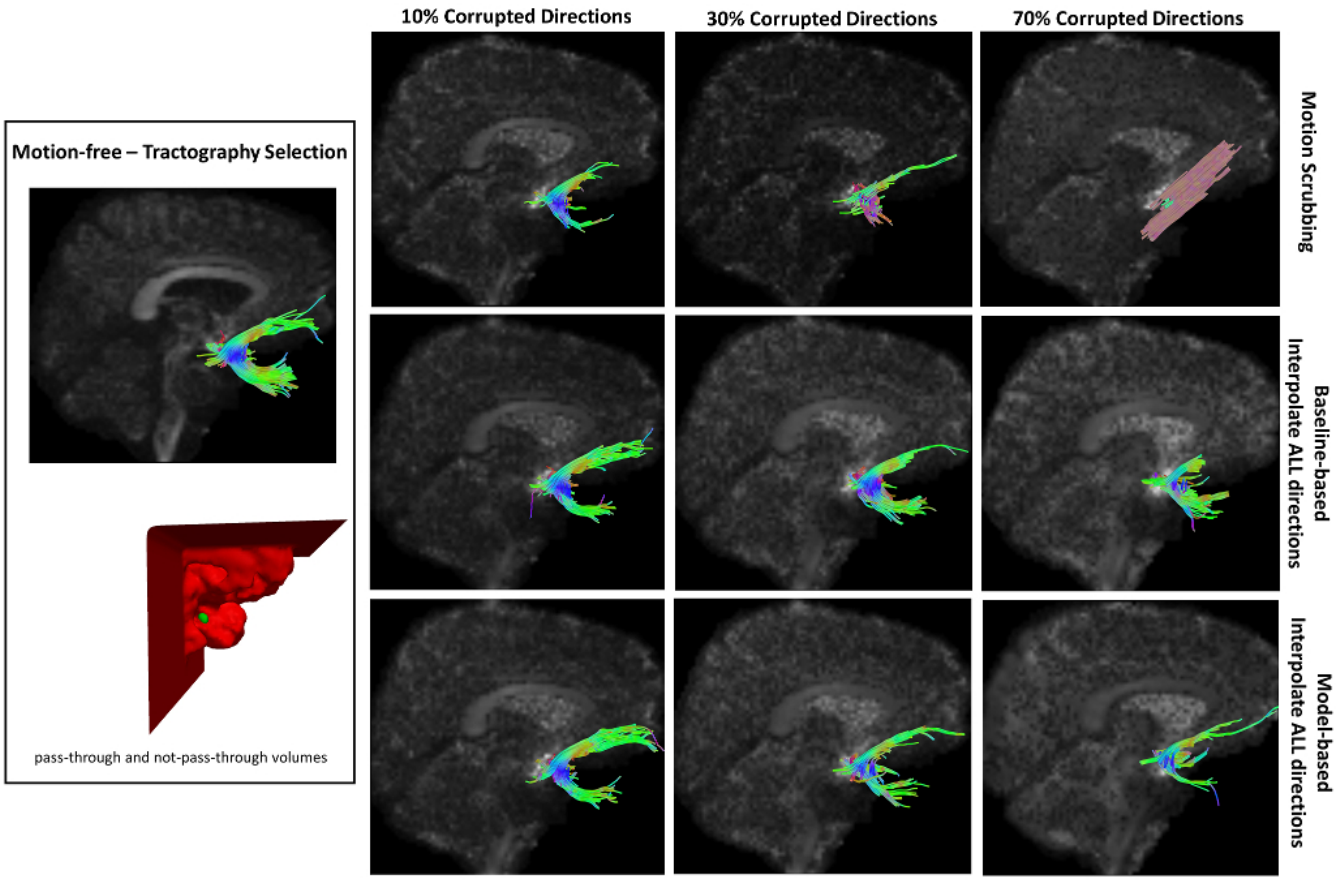


Figure 12. Sample tractography selection for the uncinate fasciculus (UNC) from the untilted motion-free dataset as well as selections from motion-corrected datasets with 10%, 30% and 70% corrupted gradient directions. Correction choices shown include outlier-based (i.e., motion scrubbing) and registration-based (using baseline and model-based reference volumes). Notice the inaccurate UNC tract being detected from the motion scrubbing choice at high percentages of motion corruption.

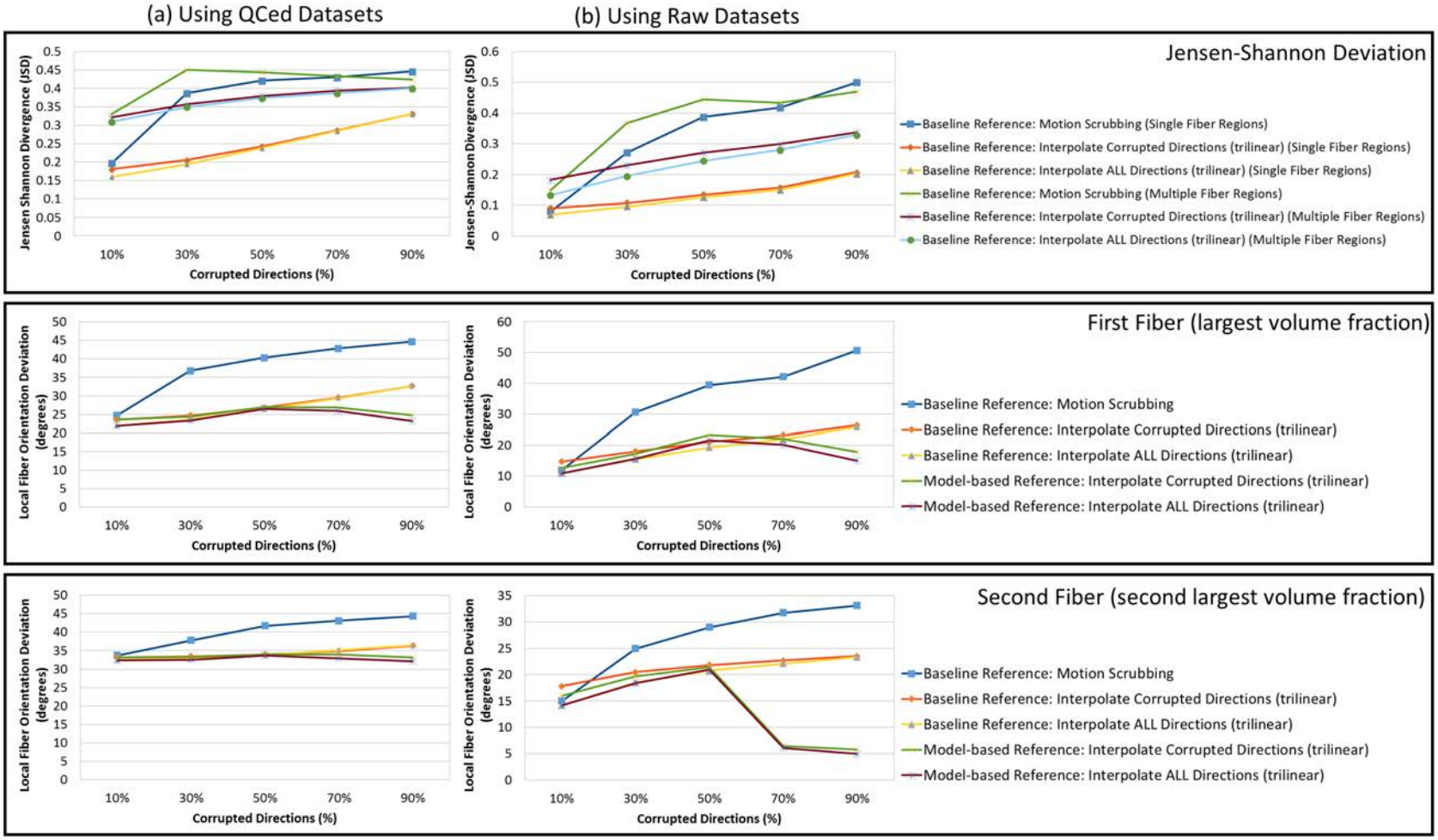


Figure 13. The average Jensen-Shannon divergence (JSD) values (first row) and the average fiber orientation deviation (second and third row) a function of motion corrupted percentage for reconstructions based on gold standards generated from (a) the QCed phantom dataset and (b) the raw phantom dataset. Notice the agreement between (a) and (b) where the impact of motion scrubbing becomes more significant with more motion-corrupted directions when compared to registration-based correction. This effect is also rendered evident for local fiber orientations.

ACKNOWLEDGMENT

Funding: This work is supported by NIH grants ACE RO1 HD 055741 and NA-MIC Roadmap U54 EB005149 and the cocaine infant project (CAMID NIDA DA022446-01). *The NIH funded Autism Centers of Excellence Infant Brain Imaging Study (ACE-IBIS) Network: Clinical Sites: University of North Carolina: J. Piven (IBIS Network PI), H.C. Hazlett, C. Chappell; University of Washington: S. Dager, A. Estes, D. Shaw; Washington University: K. Botteron, R. McKinstry, J. Constantino, J. Pruett; Children's Hospital of Philadelphia: R. Schultz, S. Paterson; University of Alberta: L. Zwaigenbaum; Data Coordinating Center: Montreal Neurological Institute: A.C. Evans, D.L. Collins, G.B. Pike, V. Fonov, P. Kostopoulos; Samir Das; Image Processing Core: University of Utah: G. Gerig; University of North Carolina: M. Styner; Statistical Analysis Core: University of North Carolina: H. Gu. This work was also funded in part by the NIH/NCRR Center for Integrative Biomedical Computing, P41RR12553.

REFERENCES

Aja-Fernández, S., Niethammer, M., Kubicki, M., Shenton, M. E., and Westin, C.-F. (2008), Restoration of dwi data using a rician lmmse estimator, *Medical Imaging, IEEE Transactions on*, 27, 10, 1389–1403

Aksoy, M., Forman, C., Straka, M., Skare, S., Holdsworth, S., Hornegger, J., et al. (2011), Real-time optical motion correction for diffusion tensor imaging, *Magnetic Resonance in Medicine*, 66, 2, 366–378

Alexander, A. L., Lee, J. E., Lazar, M., Boudos, R., DuBray, M. B., Oakes, T. R., et al. (2007), Diffusion tensor imaging of the corpus callosum in autism, *Neuroimage*, 34, 1, 61–73

Alexander, D. C., Pierpaoli, C., Basser, P. J., and Gee, J. C. (2001), Spatial transformations of diffusion tensor magnetic resonance images, *Medical Imaging, IEEE Transactions on*, 20, 11, 1131–1139

Alhamud, A., Tisdall, M. D., Hess, A. T., Hasan, K. M., Meintjes, E. M., and van der Kouwe, A. J. (2012), Volumetric navigators for real-time motion correction in diffusion tensor imaging, *Magnetic Resonance in Medicine*, 68, 4, 1097–1108

Avants, B. B., Epstein, C. L., Grossman, M., and Gee, J. C. (2008), Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain, *Medical image analysis*, 12, 1, 26–41

Bai, Y. and Alexander, D. C. (2008), Model-based registration to correct for motion between acquisitions in diffusion mr imaging, in Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on (IEEE), 947–950

Ben-Amitay, S., Jones, D. K., and Assaf, Y. (2012), Motion correction and registration of high b-value diffusion weighted images, *Magnetic Resonance in Medicine*, 67, 6, 1694–1702

Benner, T., van der Kouwe, A. J., and Sorensen, A. G. (2011), Diffusion imaging with prospective motion correction and reacquisition, *Magnetic Resonance in Medicine*, 66, 1, 154–167

Burdette, J. H., Durden, D. D., Elster, A. D., and Yen, Y.-F. (2001), High b-value diffusion-weighted MRI of normal brain, *Journal of computer assisted tomography*, 25, 4, 515–519

Caruyer, E., Aganj, I., Lenglet, C., Sapiro, G., and Deriche, R. (2013), Motion detection in diffusion MRI via online odf estimation, *International journal of biomedical imaging*, 2013

Table 3. Parcellated Structures

ROI#	Label	Region	ROI#	Label	Region
1	SPL_L	Superior parietal lobule left	89	SPL_R	Superior parietal lobule right
2	CG_L	Cingulate gyrus left	90	CG_R	Cingulate gyrus right
3	SFG_L	Superior frontal gyrus left	91	SFG_R	Superior frontal gyrus right
4	MFG_L	Middle frontal gyrus left	92	MFG_R	Middle frontal gyrus right
5	IFG_L	Inferior frontal gyrus left	93	IFG_R	Inferior frontal gyrus right
6	PreG_L	Precentral gyrus left	94	PreG_R	Precentral gyrus right
7	PoG_L	Postcentral gyrus left	95	PoG_R	Postcentral gyrus right
8	AG_L	Angular gyrus left	96	AG_R	Angular gyrus right
9	PreCu_L	Pre-cuneus left	97	PreCu_R	Pre-cuneus right
10	Cu_L	Cuneus left	98	Cu_R	Cuneus right
11	LG_L	Lingual gyrus left	99	LG_R	Lingual gyrus right
12	FuG_L	Fusiform gyrus left	100	FuG_R	Fusiform gyrus right
13	PHG_L	Parahippocampal gyrus left	101	PHG_R	Parahippocampal gyrus right
14	SOG_L	Superior occipital gyrus left	102	SOG_R	Superior occipital gyrus right
15	IOG_L	Inferior occipital gyrus left	103	IOG_R	Inferior occipital gyrus right
16	MOG_L	Middle occipital gyrus left	104	MOG_R	Middle occipital gyrus right
17	Ent_L	Entorhinal area left	105	Ent_R	Entorhinal area right
18	STG_L	Superior temporal gyrus left	106	STG_R	Superior temporal gyrus right
19	ITG_L	Inferior temporal gyrus left	107	ITG_R	Inferior temporal gyrus right
20	MTG_L	Middle temporal gyrus left	108	MTG_R	Middle temporal gyrus right
21	LFOG_L	Lateral fronto-orbital gyrus left	109	LFOG_R	Lateral fronto-orbital gyrus right
22	MFOG_L	Middle fronto-orbital gyrus left	110	MFOG_R	Middle fronto-orbital gyrus right
23	SMG_L	Supramarginal gyrus left	111	SMG_R	Supramarginal gyrus right
24	RG_L	Gyrus rectus left	112	RG_R	Gyrus rectus right
25	Ins_L	Insular left	113	Ins_R	Insular right
26	Amyg_L	Amygdala left	114	Amyg_R	Amygdala right
27	Hippo_L	Hippocampus left	115	Hippo_R	Hippocampus right
28	Cere_L	Cerebellum left	116	Cere_R	Cerebellum right
29	CST_L	Corticospinal tract left	117	CST_R	Corticospinal tract right
30	ICP_L	Inferior cerebellar peduncle left	118	ICP_R	Inferior cerebellar peduncle right
31	ML_L	Medial lemniscus left	119	ML_R	Medial lemniscus right
32	SCP_L	Superior cerebellar peduncle left	120	SCP_R	Superior cerebellar peduncle right
33	CP_L	Cerebral peduncle left	121	CP_R	Cerebral peduncle right
34	ALIC_L	Anterior limb of internal capsule left	122	ALIC_R	Anterior limb of internal capsule right
35	PLIC_L	Posterior limb of internal capsule left	123	PLIC_R	Posterior limb of internal capsule right
36	PTR_L	Posterior thalamic radiation left	124	PTR_R	Posterior thalamic radiation right
37	ACR_L	Anterior corona radiata left	125	ACR_R	Anterior corona radiata right
38	SCR_L	Superior corona radiata left	126	SCR_R	Superior corona radiata right
39	PCR_L	Posterior corona radiata left	127	PCR_R	Posterior corona radiata right
40	CGC_L	Cingulum (cingulate gyrus) left	128	CGC_R	Cingulum (cingulate gyrus) right
41	CGH_L	Cingulum (hippocampus) left	129	CGH_R	Cingulum (hippocampus) right
42	Fx/ST_L	Fornix(cres) stria terminalis left	130	Fx/ST_R	Fornix(cres) stria terminalis right
43	SLF_L	Superior longitudinal fasciculus left	131	SLF_R	Superior longitudinal fasciculus right
44	SFOF_L	Superior fronto-occipital fasciculus left	132	SFOF_R	Superior fronto-occipital fasciculus right
45	IFOF_L	Inferior fronto-occipital fasciculus left	133	IFOF_R	Inferior fronto-occipital fasciculus right
46	SS_L	Sagittal stratum left	134	SS_R	Sagittal stratum right
47	EC_L	External capsule left	135	EC_R	External capsule right
48	UNC_L	Uncinate fasciculus left	136	UNC_R	Uncinate fasciculus right
49	PCT_L	Pontine crossing tract left	137	PCT_R	Pontine crossing tract right
50	MCP_L	Middle cerebellar peduncle left	138	MCP_R	Middle cerebellar peduncle right
51	Fx_L	Fornix (column and body) left	139	Fx_R	Fornix right
52	GCC_L	Genu of corpus callosum left	140	GCC_R	Genu of corpus callosum right
53	BCC_L	Body of corpus callosum left	141	BCC_R	Body of corpus callosum right
54	SCC_L	Splenium of corpus callosum left	142	SCC_R	Splenium of corpus callosum right
55	RLIC_L	Retrolenticular part of internal capsule left	143	RLIC_R	Retrolenticular part of internal capsule right
56	RN_L	Red nucleus left	144	RN_R	Red nucleus right
57	SN_L	Substancia nigra left	145	SN_R	Substancia nigra right
58	Tp_L	Tapatum left	146	Tp_R	Tapatum right
59	CN_L	Caudate nucleus left	147	CN_R	Caudate nucleus right
60	P_L	Putamen left	148	P_R	Putamen right
61	Th_L	Thalamus left	149	Th_R	Thalamus right
62	GP_L	Globus pallidus left	150	GP_R	Globus pallidus right
63	MB_L	Midbrain left	151	MB_R	Midbrain right
64	Pons_L	Pons left	152	Pons_R	Pons right
65	Med_L	Medulla left	153	Med_R	Medulla right
66	SP_WM_L	Superior parietal wm left	154	SP_WM_R	Superior parietal wm right
67	CG_WM_L	Cingulum wm left	155	CG_WM_R	Cingulum wm right
68	SF_WM_L	Superior frontal wm left	156	SF_WM_R	Superior frontal wm right
69	MF_WM_L	Middle frontal wm left	157	MF_WM_R	Middle frontal wm right
70	IF_WM_L	Inferior frontal wm left	158	IF_WM_R	Inferior frontal wm right
71	Pr_WM_L	Precentral wm left	159	Pr_WM_R	Precentral wm right
72	Po_WM_L	Postcentral wm left	160	Po_WM_R	Postcentral wm right
73	A_WM_L	Angular wm left	161	A_WM_R	Angular wm right
74	PreCu_WM_L	Pre-cuneus wm left	162	PreCu_WM_R	Pre-cuneus wm right
75	Cu_WM_L	Cuneus wm left	163	Cu_WM_R	Cuneus wm right
76	L_WM_L	Lingual wm left	164	L_WM_R	Lingual wm right
77	Fu_WM_L	Fusiform wm left	165	Fu_WM_R	Fusiform wm right
78	SO_WM_L	Superior occipital wm left	166	SO_WM_R	Superior occipital wm right
79	IO_WM_L	Inferior occipital wm left	167	IO_WM_R	Inferior occipital wm right
80	MO_WM_L	Middle occipital wm left	168	MO_WM_R	Middle occipital wm right
81	ST_WM_L	Superior temporal wm left	169	ST_WM_R	Superior temporal wm right
82	IT_WM_L	Inferior temporal wm left	170	IT_WM_R	Inferior temporal wm right
83	MT_WM_L	Middle temporal wm left	171	MT_WM_R	Middle temporal wm right
84	LFO_WM_L	Lateral fronto-orbital wm left	172	LFO_WM_R	Lateral fronto-orbital wm right
85	MFO_WM_L	Middle fronto-orbital wm left	173	MFO_WM_R	Middle fronto-orbital wm right
86	SM_WM_L	Supramarginal wm left	174	SM_WM_R	Supramarginal wm right
87	Rect_WM_L	Rectus wm left	175	Rect_WM_R	Rectus wm right
88	Cere_WM_L	Cerebellum wm left	176	Cere_WM_R	Cerebellum wm right

Table 4. The average and standard deviation of the affine transformation parameters

Dataset ID	Translation Vector Magnitude (mm)	Rotation Angle (degrees)	Scale in x direction	Scale in y direction	Scale in z direction	Skew A	Skew B	Skew C
phan1_time1_chop	0.263283 +/- 0.129780	0.4258 +/- 0.1765	0.974387 +/- 0.122766	0.991650 +/- 0.125002	0.978987 +/- 0.123368	-0.001053 +/- 0.001162	-0.000104 +/- 0.005986	-0.000753 +/- 0.001325
phan1_time1_unc_hos	0.470658 +/- 0.232117	0.1534 +/- 0.0918	0.967640 +/- 0.120958	0.987012 +/- 0.123419	0.969454 +/- 0.121189	-0.002346 +/- 0.001614	0.001971 +/- 0.000873	-0.002854 +/- 0.002759
phan1_time1_unc_res	0.189921 +/- 0.063812	0.2667 +/- 0.0947	0.969042 +/- 0.121136	0.981819 +/- 0.122801	0.968367 +/- 0.121047	0.001694 +/- 0.001812	-0.000951 +/- 0.001204	-0.003887 +/- 0.002544
phan1_time1_washu_res	0.592816 +/- 0.167149	0.0998 +/- 0.108	0.972356 +/- 0.122509	0.985866 +/- 0.124311	0.980857 +/- 0.123619	-0.000613 +/- 0.001537	0.001895 +/- 0.001171	0.000042 +/- 0.002984
phan1_time2_chop	0.476715 +/- 0.158225	0.2922 +/- 0.1342	0.970214 +/- 0.122241	0.981341 +/- 0.123652	0.974277 +/- 0.122824	-0.004799 +/- 0.002259	0.001672 +/- 0.001119	-0.004175 +/- 0.002521
phan1_time2_unc_hos	0.580458 +/- 0.205031	0.1544 +/- 0.0845	0.968968 +/- 0.121126	0.985392 +/- 0.123211	0.969308 +/- 0.121165	-0.000668 +/- 0.001327	0.003067 +/- 0.001256	-0.003470 +/- 0.001561
phan1_time2_unc_res	0.282364 +/- 0.103415	0.3059 +/- 0.186	0.966004 +/- 0.120753	0.974620 +/- 0.121831	0.967014 +/- 0.120879	-0.000146 +/- 0.000979	0.001095 +/- 0.000984	0.001062 +/- 0.001672
phan1_time2_washu_res	0.519812 +/- 0.201197	0.1394 +/- 0.1097	0.971916 +/- 0.124449	0.988555 +/- 0.126727	0.974220 +/- 0.124811	-0.001642 +/- 0.001679	-0.001064 +/- 0.001624	-0.000475 +/- 0.002170
phan2_time1_chop	0.476104 +/- 0.150039	0.2744 +/- 0.2404	0.973545 +/- 0.124653	0.987570 +/- 0.126611	0.977957 +/- 0.125338	-0.000196 +/- 0.001119	0.002787 +/- 0.000843	0.000057 +/- 0.002648
phan2_time1_unc_hos	0.310917 +/- 0.145060	0.1143 +/- 0.1271	0.973924 +/- 0.124700	0.989896 +/- 0.126798	0.972212 +/- 0.124521	0.001524 +/- 0.001987	0.001644 +/- 0.003226	-0.001149 +/- 0.002316
phan2_time1_unc_res	0.573942 +/- 0.159912	0.338 +/- 0.2473	0.972004 +/- 0.125488	0.992603 +/- 0.128153	0.978598 +/- 0.126395	-0.001017 +/- 0.000745	-0.001052 +/- 0.001142	0.000181 +/- 0.001112
phan2_time1_washu_res	0.399943 +/- 0.159193	0.2744 +/- 0.1721	0.973515 +/- 0.122656	0.990590 +/- 0.124862	0.978706 +/- 0.123327	0.001399 +/- 0.001533	-0.000429 +/- 0.002762	-0.002402 +/- 0.002564
phan2_time2_chop	0.334249 +/- 0.148992	0.5177 +/- 0.2077	0.975330 +/- 0.124881	0.986590 +/- 0.126375	0.977786 +/- 0.125299	0.000482 +/- 0.001227	0.000831 +/- 0.001404	0.001103 +/- 0.002349
phan2_time2_unc_hos	0.657453 +/- 0.159169	0.551 +/- 0.1521	0.972614 +/- 0.122543	0.986810 +/- 0.124331	0.979568 +/- 0.123450	0.000448 +/- 0.001303	0.001030 +/- 0.000899	0.001600 +/- 0.002339
phan2_time2_unc_res	0.166599 +/- 0.063907	0.4218 +/- 0.231	0.971802 +/- 0.122443	0.982748 +/- 0.123822	0.969650 +/- 0.122169	0.007760 +/- 0.001445	0.001021 +/- 0.001453	-0.000299 +/- 0.001389
phan2_time2_washu_res	0.201228 +/- 0.058342	0.4382 +/- 0.1562	0.973766 +/- 0.123671	0.983821 +/- 0.124961	0.972836 +/- 0.123570	-0.002218 +/- 0.001863	-0.001634 +/- 0.002000	-0.001353 +/- 0.002329
phan3_time1_chop	0.259133 +/- 0.175041	0.6439 +/- 0.126	0.971986 +/- 0.123448	0.986653 +/- 0.125336	0.976689 +/- 0.124069	0.001603 +/- 0.001702	-0.004551 +/- 0.002228	0.003174 +/- 0.004675
phan3_time1_sea	0.452670 +/- 0.297523	0.3769 +/- 0.2054	0.988928 +/- 0.127704	0.991992 +/- 0.128127	0.989708 +/- 0.127798	0.003240 +/- 0.002314	0.000590 +/- 0.001049	0.000753 +/- 0.001993
phan3_time1_unc_hos	0.635246 +/- 0.158109	0.1311 +/- 0.127	0.975263 +/- 0.122875	0.989361 +/- 0.124665	0.982201 +/- 0.123788	0.001640 +/- 0.001893	0.001942 +/- 0.000933	0.000695 +/- 0.003627
phan3_time1_unc_res	0.539050 +/- 0.152979	0.4504 +/- 0.1164	0.973626 +/- 0.122668	0.987238 +/- 0.124383	0.979584 +/- 0.123445	0.001921 +/- 0.001540	0.004630 +/- 0.002217	0.000202 +/- 0.002536
phan3_time2_chop	0.280374 +/- 0.086683	0.1584 +/- 0.1157	0.973434 +/- 0.123631	0.984695 +/- 0.125073	0.972197 +/- 0.123512	0.002663 +/- 0.002568	0.003467 +/- 0.001624	0.000037 +/- 0.002711
phan3_time2_sea	0.556044 +/- 0.238928	0.2208 +/- 0.133	0.988590 +/- 0.127661	0.992080 +/- 0.128118	0.984150 +/- 0.127070	0.002273 +/- 0.001335	0.000950 +/- 0.000859	0.001297 +/- 0.002274
phan3_time2_unc_hos	0.526151 +/- 0.216508	0.1569 +/- 0.1836	0.974415 +/- 0.122768	0.990469 +/- 0.124820	0.980314 +/- 0.123656	-0.001155 +/- 0.002091	0.002274 +/- 0.000998	0.003932 +/- 0.002886
phan3_time2_unc_res	0.743160 +/- 0.172370	0.1276 +/- 0.0877	0.972536 +/- 0.124526	0.986215 +/- 0.126279	0.981740 +/- 0.125742	-0.000120 +/- 0.001058	0.002701 +/- 0.001266	0.002270 +/- 0.004077

Chang, L.-C., Jones, D. K., and Pierpaoli, C. (2005), Restore: robust estimation of tensors by outlier rejection, *Magnetic Resonance in Medicine*, 53, 5, 1088–1095

Chang, L.-C., Walker, L., and Pierpaoli, C. (2012), Informed restore: A method for robust estimation of diffusion tensor from low redundancy datasets in the presence of physiological noise artifacts, *Magnetic Resonance in Medicine*, 68, 5, 1654–1663

Chiang, M.-C., Barysheva, M., Lee, A. D., Madsen, S., Klunder, A. D., Toga, A. W., et al. (2008), Brain fiber architecture, genetics, and intelligence: a high angular resolution diffusion imaging (hardi) study, in Medical Image Computing and Computer-Assisted Intervention–MICCAI 2008 (Springer), 1060–1067

Cohen-Adad, J., Descoteaux, M., and Wald, L. L. (2011), Quality assessment of high angular resolution diffusion imaging data using bootstrap on q-ball reconstruction, *Journal of Magnetic Resonance Imaging*, 33, 5, 1194–1208

Collins, D. L., Zijdenbos, A. P., Kollokian, V., Sled, J. G., Kabani, N. J., Holmes, C. J., et al. (1998), Design and construction of a realistic digital brain phantom, *Medical Imaging, IEEE Transactions on*, 17, 3, 463–468

de Luis-García, R., Westin, C.-F., and Alberola-López, C. (2013), Geometrical constraints for robust tractography selection, *NeuroImage*, 81, 26–48

Dumas, E. M., van den Bogaard, S. J., Ruber, M. E., Reilmann, R., Stout, J. C., Craufurd, D., et al. (2012), Early changes in white matter pathways of the sensorimotor cortex in premanifest huntington's disease, *Human brain mapping*, 33, 1, 203–212

Feinberg, D. A. and Setsompop, K. (2013), Ultra-fast MRI of the human brain with simultaneous multi-slice imaging, *Journal of Magnetic Resonance*, 229, 90–100

Fletcher, P. T. and Joshi, S. (2007), Riemannian geometry for the statistical analysis of diffusion tensor data, *Signal Processing*, 87, 2, 250–262

Garyfallidis, E. (2012), Towards an accurate brain tractography, Ph.D. thesis, PhD thesis, University of Cambridge

Garyfallidis, E., Brett, M., Amirbekian, B., Rokem, A., Van Der Walt, S., Descoteaux, M., et al. (2014), Dipy, a library for the analysis of diffusion MRI data, *Frontiers in Neuroinformatics*, 8, 8

Gilmore, J. H., Kang, C., Evans, D. D., Wolfe, H. M., Smith, J. K., Lieberman, J. A., et al. (2010), Prenatal and neonatal brain structure and white matter maturation in children at high risk for schizophrenia, *American Journal of Psychiatry*, 167, 9, 1083–1091

Goodlett, C. B., Fletcher, P. T., Gilmore, J. H., and G., G. (2009), Group analysis of dti fiber tract statistics with application to neurodevelopment, *Neuroimage*, 45, 1, 133–142

Gumus, K., Keating, B., Poser, B. A., Armstrong, B., Chang, L., Maclaren, J., et al. (2013), Prevention of motion-induced signal loss in diffusion-weighted echo-planar imaging by dynamic restoration of gradient moments, *Magnetic Resonance in Medicine*

Hallgren, K. A. (2012), Computing inter-rater reliability for observational data: An overview and tutorial, *Tutorials in quantitative methods for psychology*, 8, 1, 23

Hammond, D. K., Gur, Y., and Johnson, C. R. (2013a), Graph diffusion distance: A difference measure for weighted graphs based on the graph laplacian exponential kernel, in Global Conference on Signal and Information Processing (GlobalSIP), 2013 IEEE, 419–422

Hammond, D. K., Scherrer, B., and Warfield, S. K. (2013b), Cortical graph smoothing: a novel method for exploiting dwi-derived anatomical brain connectivity to improve eeg source estimation, 32, 10, 1952–1963

Herbst, M., Maclaren, J., Weigel, M., Korvink, J., Hennig, J., and Zaitsev, M. (2012), Prospective motion correction with continuous gradient updates in diffusion weighted imaging, *Magnetic Resonance in Medicine*, 67, 2, 326–338

Irfanoglu, M. O., Walker, L., Sarlls, J., Marenco, S., and Pierpaoli, C. (2012), Effects of image distortions originating from susceptibility variations and concomitant fields on diffusion mri tractography results, *NeuroImage*, 61, 1, 275–288

Jellison, B. J., Field, A. S., Medow, J., Lazar, M., Salamat, M. S., and Alexander, A. L. (2004), Diffusion tensor imaging of cerebral white matter: a pictorial review of physics, fiber tract anatomy, and tumor imaging patterns, *American Journal of Neuroradiology*, 25, 3, 356–369

Jenkinson, M., Bannister, P., Brady, M., and Smith, S. (2002), Improved optimization for the robust and accurate linear registration and motion correction of brain images, *Neuroimage*, 17, 2, 825–841

Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W., and Smith, S. M. (2012), Fsl, *Neuroimage*, 62, 2, 782–790

Johnson, H., Harris, G., and Williams, K. (2007), Brainsfit: mutual information rigid registrations of whole-brain 3d images, using the insight toolkit, *Insight J*

Jones, D. K. and Basser, P. J. (2004), Squashing peanuts and smashing pumpkins: How noise distorts diffusion-weighted mr data, *Magnetic Resonance in Medicine*, 52, 5, 979–993

Jones, D. K., Knösche, T. R., and Turner, R. (2013), White matter integrity, fiber count, and other fallacies: the do's and don'ts of diffusion MRI, *Neuroimage*, 73, 239–254

Joshi, S., Davis, B., Jomier, M., and Gerig, G. (2004), Unbiased diffeomorphic atlas construction for computational anatomy, *NeuroImage*, 23, S151–S160

Kaiser, M. (2011), A tutorial in connectome analysis: topological and spatial features of brain networks, *Neuroimage*, 57, 3, 892–907

Kober, T., Gruetter, R., and Krueger, G. (2012), Prospective and retrospective motion correction in diffusion magnetic resonance imaging of the human brain, *Neuroimage*, 59, 1, 389–398

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., et al. (2009), Circos: an information aesthetic for comparative genomics, *Genome research*, 19, 9, 1639–1645

Landis, J. R., Koch, G. G., et al. (1977), The measurement of observer agreement for categorical data., *biometrics*, 33, 1, 159–174

Le Bihan, D., Poupon, C., Amadon, A., and Lethimonnier, F. (2006), Artifacts and pitfalls in diffusion MRI, *Journal of magnetic resonance imaging*, 24, 3, 478–488

Leemans, A. and Jones, D. K. (2009), The b-matrix must be rotated when correcting for subject motion in dti data, *Magnetic Resonance in Medicine*, 61, 6, 1336–1349

Liu, Z., Wang, Y., Gerig, G., Gouttard, S., Tao, R., Fletcher, T., et al. (2010), Quality control of diffusion weighted images, in SPIE Medical Imaging (International Society for Optics and Photonics), 76280J–76280J

Mohammadi, S., Möller, H. E., Kugel, H., Müller, D. K., and Deppe, M. (2010), Correcting eddy current and motion effects by affine whole-brain registrations: Evaluation of three-dimensional distortions and comparison with slicewise correction, *Magnetic Resonance in Medicine*, 64, 4, 1047–1056

Nadler, B., Lafon, S., Coifman, R., and Kevrekidis, I. (2005), Diffusion maps, spectral clustering and eigenfunctions of fokker-planck operators, in NIPS

Netsch, T. and van Muiswinkel, A. (2004), Quantitative evaluation of image-based distortion correction in diffusion tensor imaging, *Medical Imaging, IEEE Transactions on*, 23, 7, 789–798

Nunes, R. G., Jezzard, P., and Clare, S. (2005), Investigations on the efficiency of cardiac-gated methods for the acquisition of diffusion-weighted images, *Journal of Magnetic Resonance*, 177, 1, 102–110

Oakes, T., Johnstone, T., Ores Walsh, K., Greischar, L., Alexander, A., Fox, A., et al. (2005), Comparison of fMRI motion correction software tools, *Neuroimage*, 28, 3, 529–543

Oguz, I., Farzinfar, M., Matsui, J., Budin, F., Liu, Z., Gerig, G., et al. (2014), DTIPrep: Quality control of diffusion-weighted images, *Frontiers in Neuroinformatics*, 8, 4

Oishi, K., Faria, A., Jiang, H., Li, X., Akhter, K., Zhang, J., et al. (2009), Atlas-based whole brain white matter analysis using large deformation diffeomorphic metric mapping: application to normal elderly and alzheimer's disease participants, *Neuroimage*, 46, 2, 486–499

Pannek, K., Raffelt, D., Bell, C., Mathias, J. L., and Rose, S. E. (2012), Homor: Higher order model outlier rejection for high b-value mr diffusion data, *Neuroimage*, 63, 2, 835–842

Parizel, P., Van Rompaey, V., Van Loock, R., Van Hecke, W., Van Goethem, J., Leemans, A., et al. (2007), Influence of user-defined parameters on diffusion tensor tractography of the corticospinal tract, *Neuroradiology Journal*, 20, 139–147

Pierpaoli, C. (2010), Artifacts in diffusion MRI, *Diffusion MRI: Theory, Methods and Applications, Oxford University Press, New York*, 303–318

Pierpaoli, C., Walker, L., Irfanoglu, M., Barnett, A., Basser, P., Chang, L., et al. (2010), Tortoise: an integrated software package for processing of diffusion MRI data, *Book TORTOISE: an integrated software package for processing of diffusion MRI data (Editor ed. eds.)*, 18, 1597

Porter, D. A. and Heidemann, R. M. (2009), High resolution diffusion-weighted imaging using readout-segmented echo-planar imaging, parallel imaging and a two-dimensional navigator-based reacquisition, *Magnetic Resonance in Medicine*, 62, 2, 468–475

Ramani, S., Thevenaz, P., and Unser, M. (2010), Regularized interpolation for noisy images, *Medical Imaging, IEEE Transactions on*, 29, 2, 543–558

Rohde, G., Barnett, A., Basser, P., Marenco, S., and Pierpaoli, C. (2004), Comprehensive approach for correction of motion and distortion in diffusion-weighted MRI, *Magnetic Resonance in Medicine*, 51, 1, 103–114

Rubinov, M. and Sporns, O. (2010), Complex network measures of brain connectivity: uses and interpretations, *Neuroimage*, 52, 3, 1059–1069

Sakaie, K. E. and Lowe, M. J. (2010), Quantitative assessment of motion correction for high angular resolution diffusion imaging, *Magnetic resonance imaging*, 28, 2, 290–296

SCIInstitute (2014), Atlaswerks: An open-source (bsd license) software package for medical image atlas generation. scientific computing and imaging institute (sci), university of utah.

Shi, X., Kholmovski, E. G., Kim, S.-E., Parker, D. L., and Jeong, E.-K. (2009), Improvement of accuracy of diffusion MRI using real-time self-gated data acquisition, *NMR in Biomedicine*, 22, 5, 545–550

Smith, S. M. (2002), Fast robust automated brain extraction, *Human brain mapping*, 17, 3, 143–155

Soares, J. M., Marques, P., Alves, V., and Sousa, N. (2013), A hitchhiker's guide to diffusion tensor imaging, *Frontiers in neuroscience*, 7

Tournier, J., Calamante, F., Connelly, A., et al. (2007), Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution, *NeuroImage*, 35, 4, 1459–1472

Tournier, J.-D., Mori, S., and Leemans, A. (2011), Diffusion tensor imaging and beyond, *Magnetic Resonance in Medicine*, 65, 6, 1532–1556

Tristán-Vega, A. and Aja-Fernández, S. (2009), Design and construction of a realistic dwi phantom for filtering performance assessment, in Medical Image Computing and Computer-Assisted Intervention–MICCAI 2009 (Springer), 951–958

Tristán-Vega, A. and Aja-Fernández, S. (2010), Dwi filtering using joint information for dti and hardi, *Medical Image Analysis*, 14, 2, 205–218

Tuch, D. S., Reese, T. G., Wiegell, M. R., Makris, N., Belliveau, J. W., and Wedeen, V. J. (2002), High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity, *Magnetic Resonance in Medicine*, 48, 4, 577–582

Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., et al. (2010), N4itk: improved n3 bias correction, *Medical Imaging, IEEE Transactions on*, 29, 6, 1310–1320

Veraart, J., Rajan, J., Peeters, R. R., Leemans, A., Sunaert, S., and Sijbers, J. (2013), Comprehensive framework for accurate diffusion MRI parameter estimation, *Magnetic Resonance in Medicine*, 70, 4, 972–984

Verde, A. R., Budin, F., Berger, J.-B., Gupta, A., Farzinfar, M., Kaiser, A., et al. (2013), Unc-utah na-mic framework for dti fiber tract analysis, *Frontiers in neuroinformatics*, 7

Wolff, J. J., Gu, H., Gerig, G., Elison, J. T., Styner, M., Gouttard, S., et al. (2012), Differences in white matter fiber tract development present from 6 to 24 months in infants with autism, *American Journal of Psychiatry*, 169, 6, 589–600

Yendiki, A., Koldewyn, K., Kakunoori, S., Kanwisher, N., and Fischl, B. (2014), Spurious group differences due to head motion in a diffusion MRI study, *NeuroImage*, 88, 79–90

Figure 1.JPEG

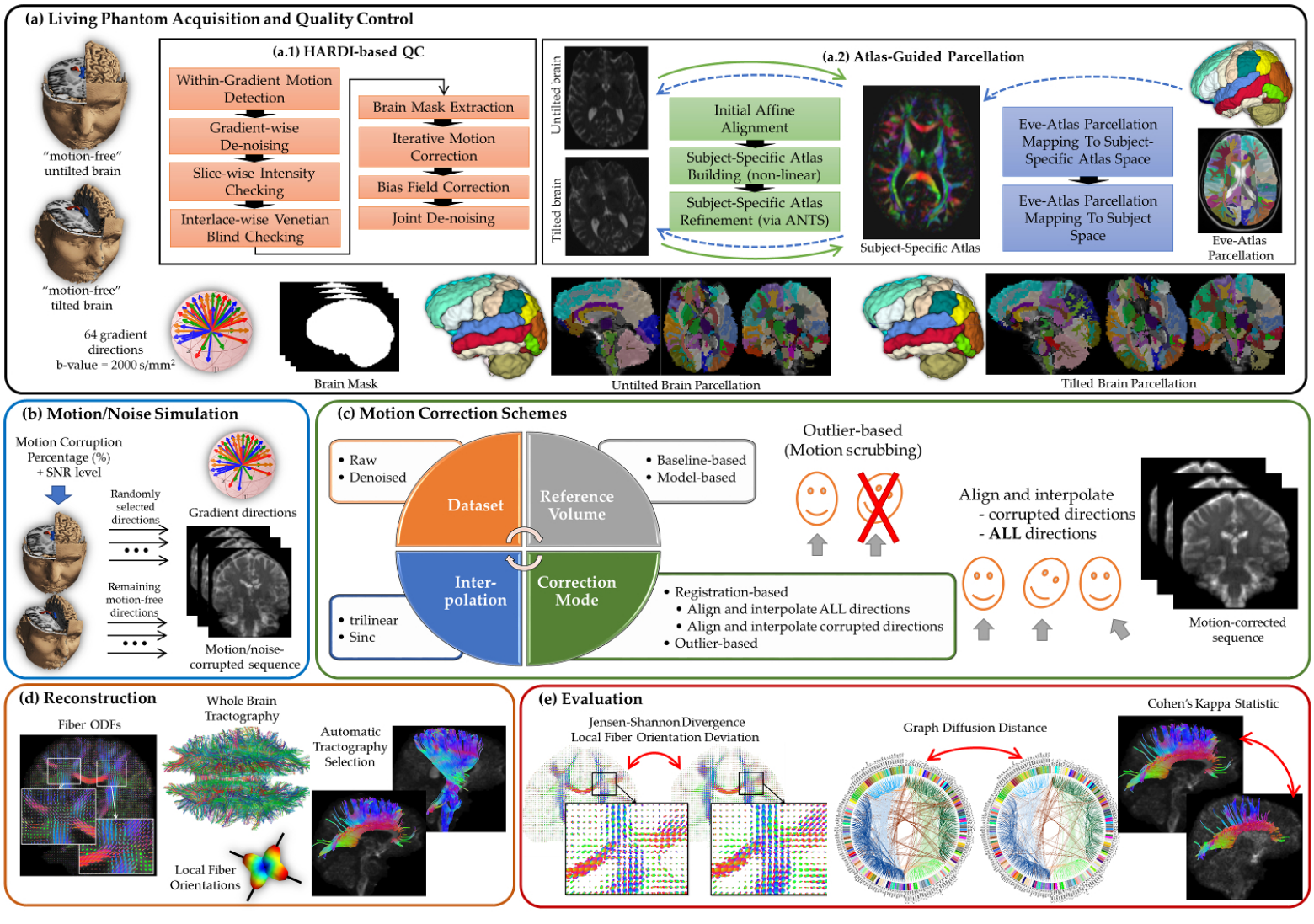
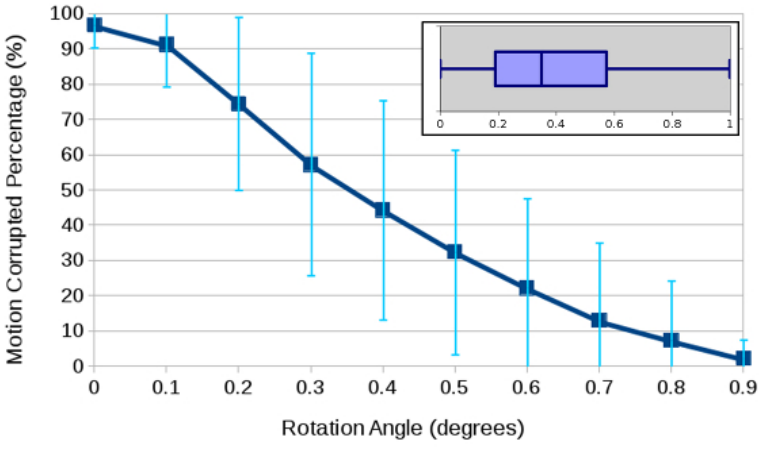


Figure 2.JPEG



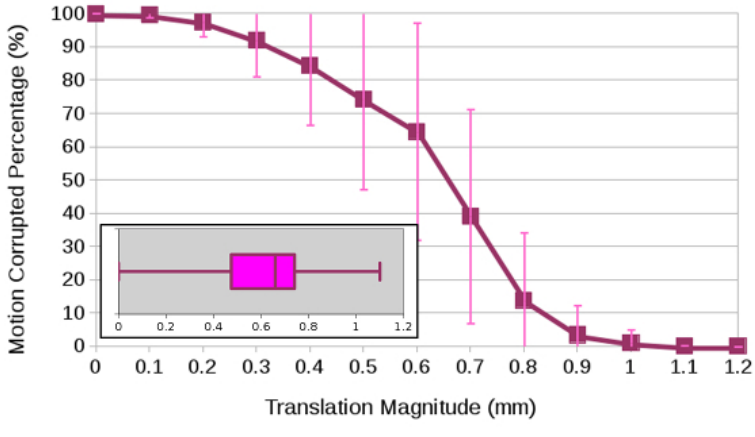


Figure 3.JPEG

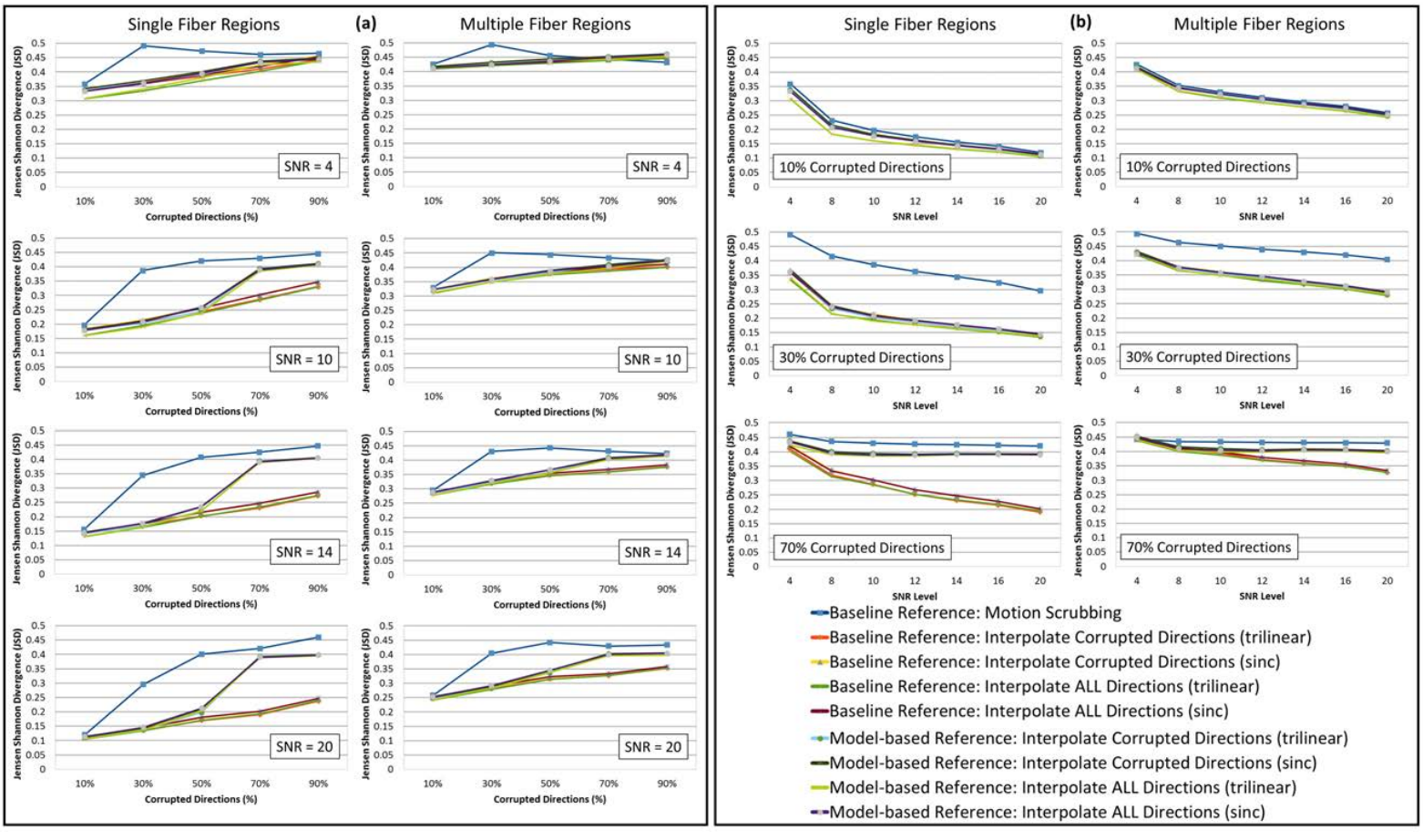


Figure 4.JPEG

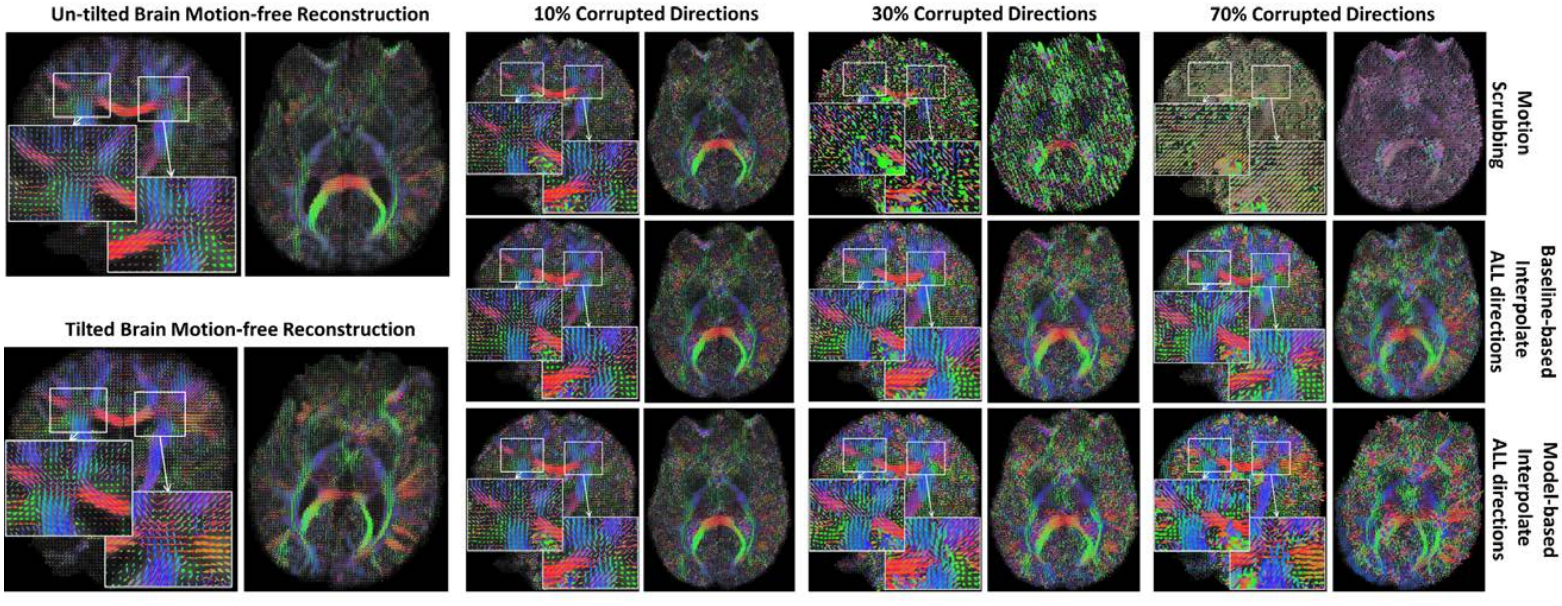


Figure 5.JPEG

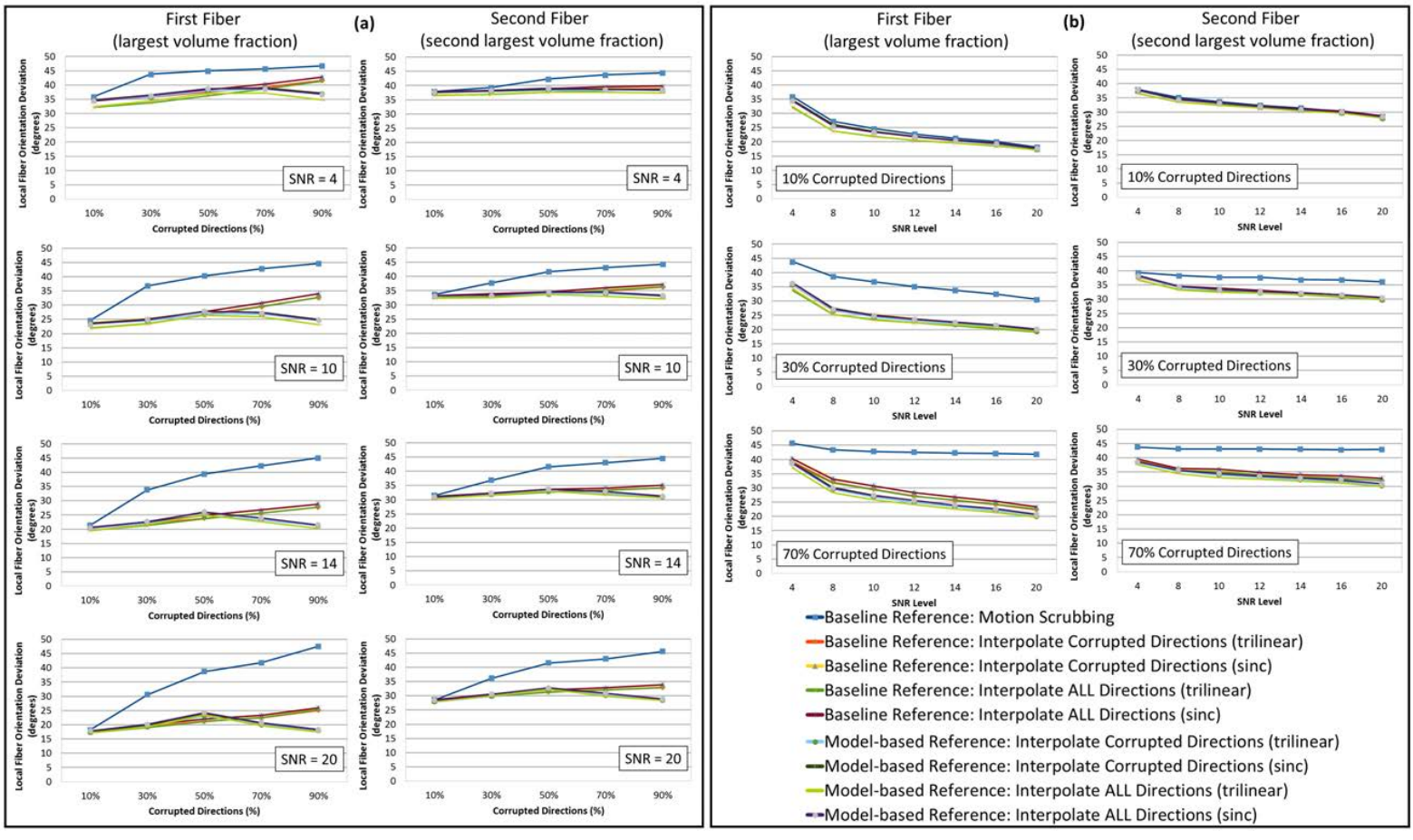


Figure 6.JPEG

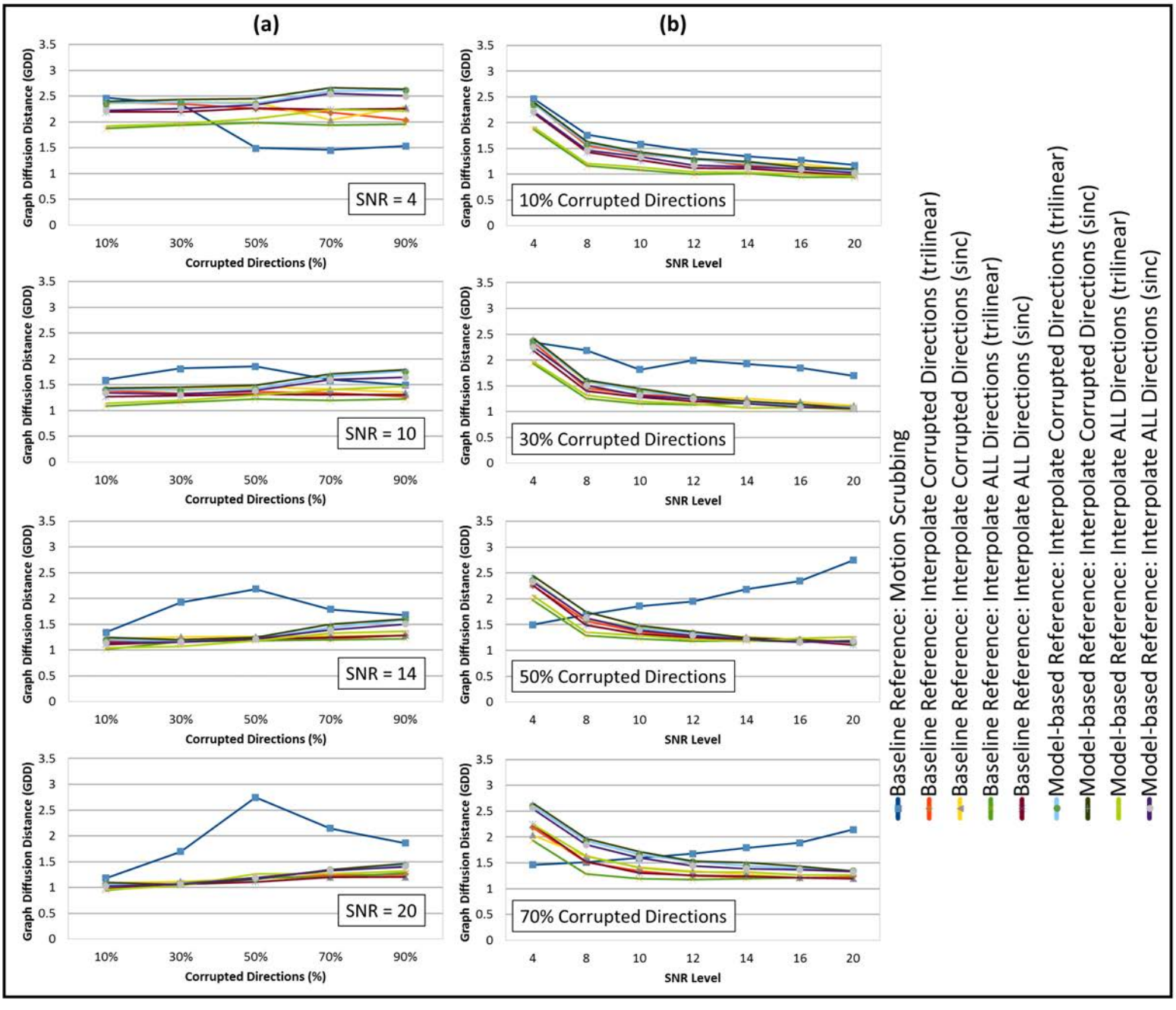


Figure 7.JPEG

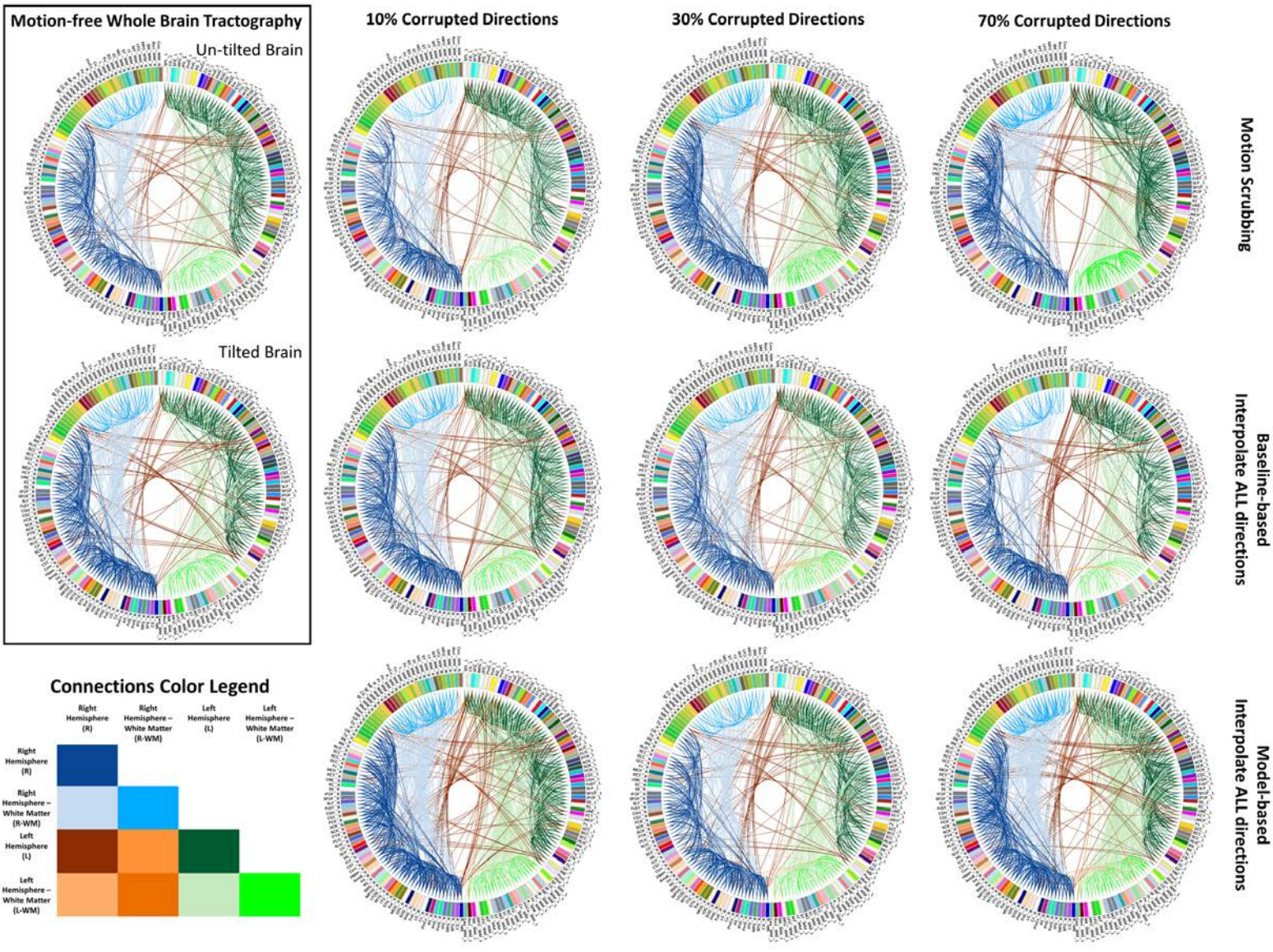


Figure 8.JPEG

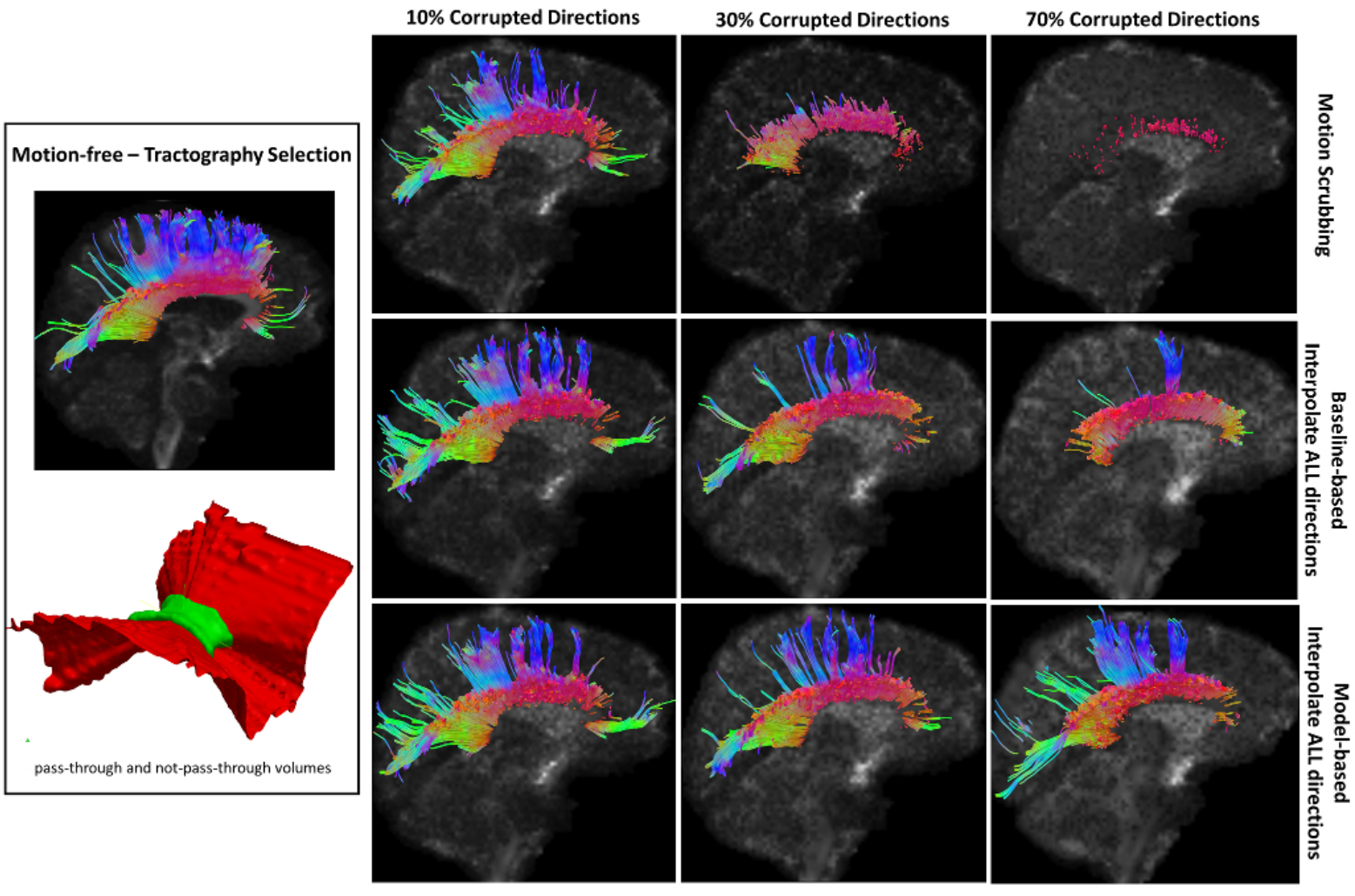


Figure 9.JPEG

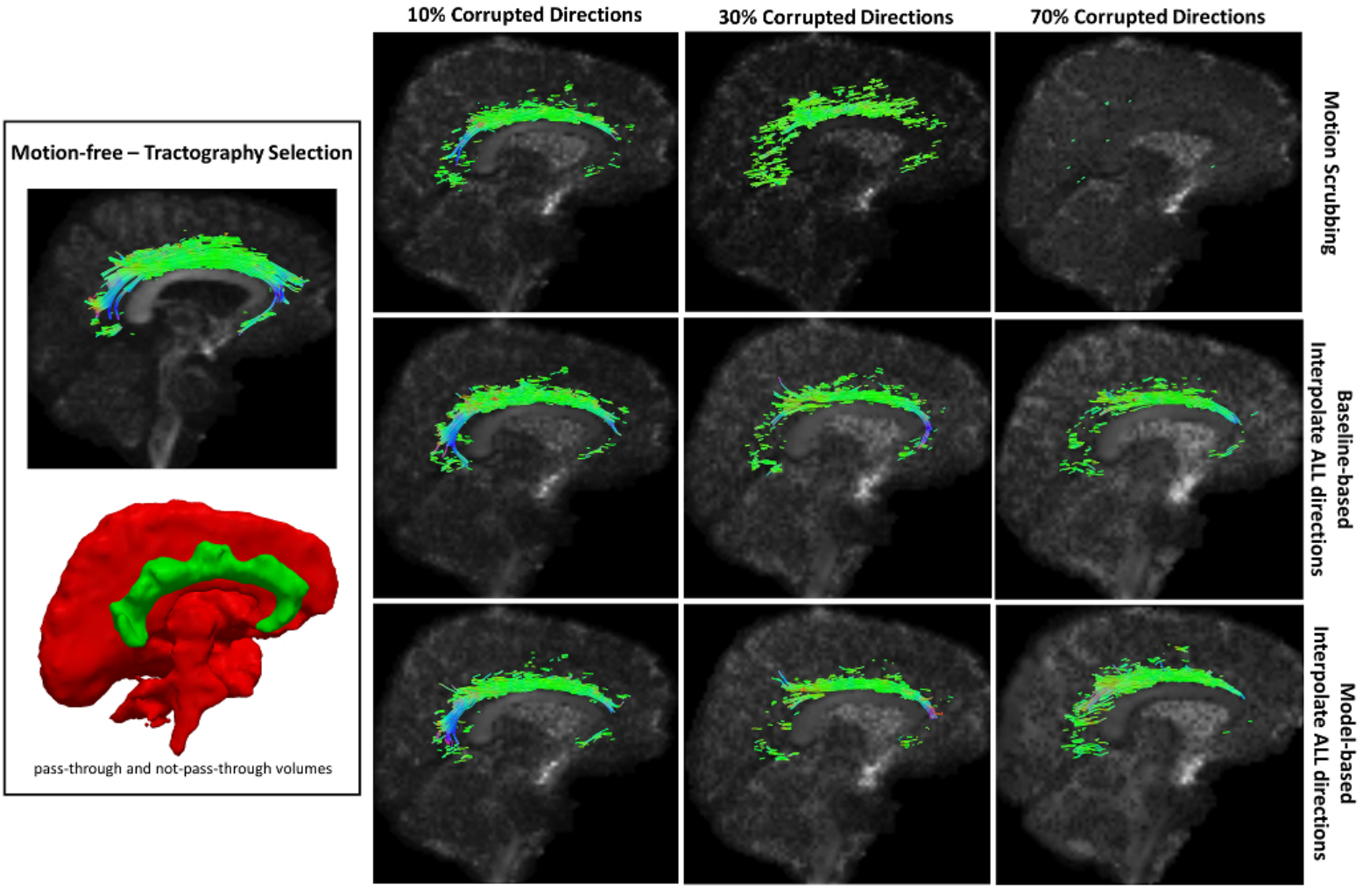


Figure 10.JPEG

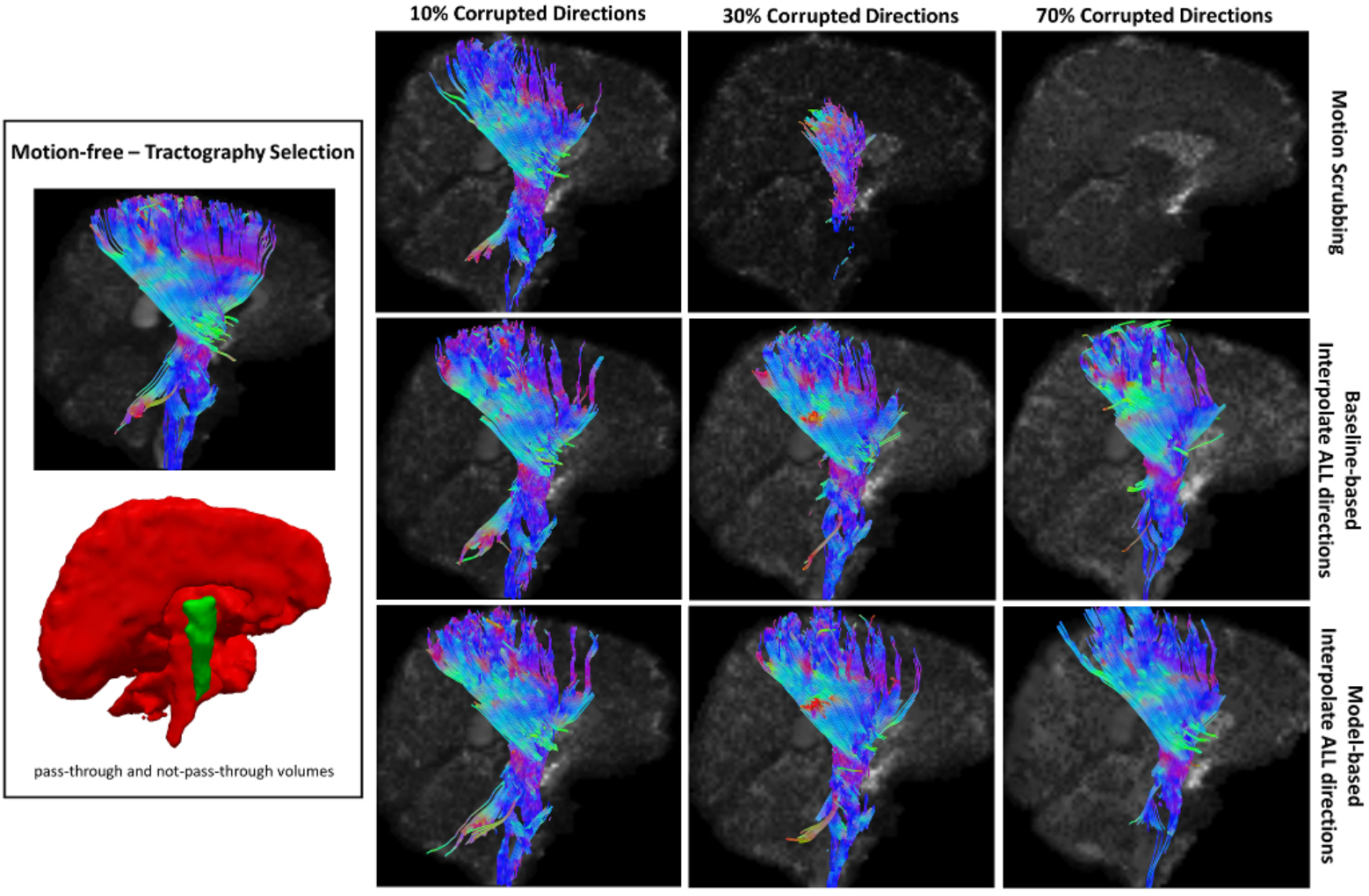


Figure 11.JPEG

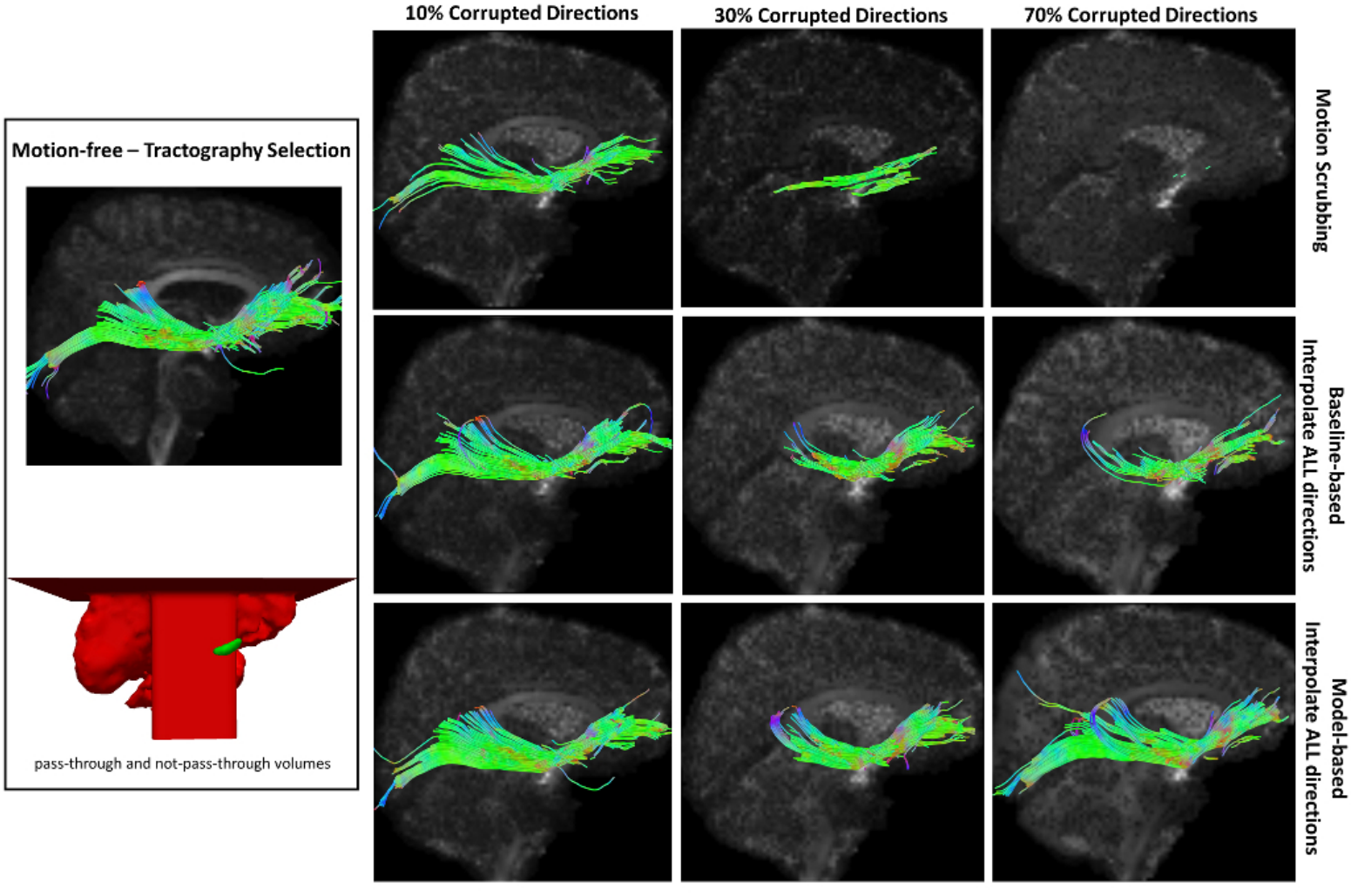


Figure 12.JPEG

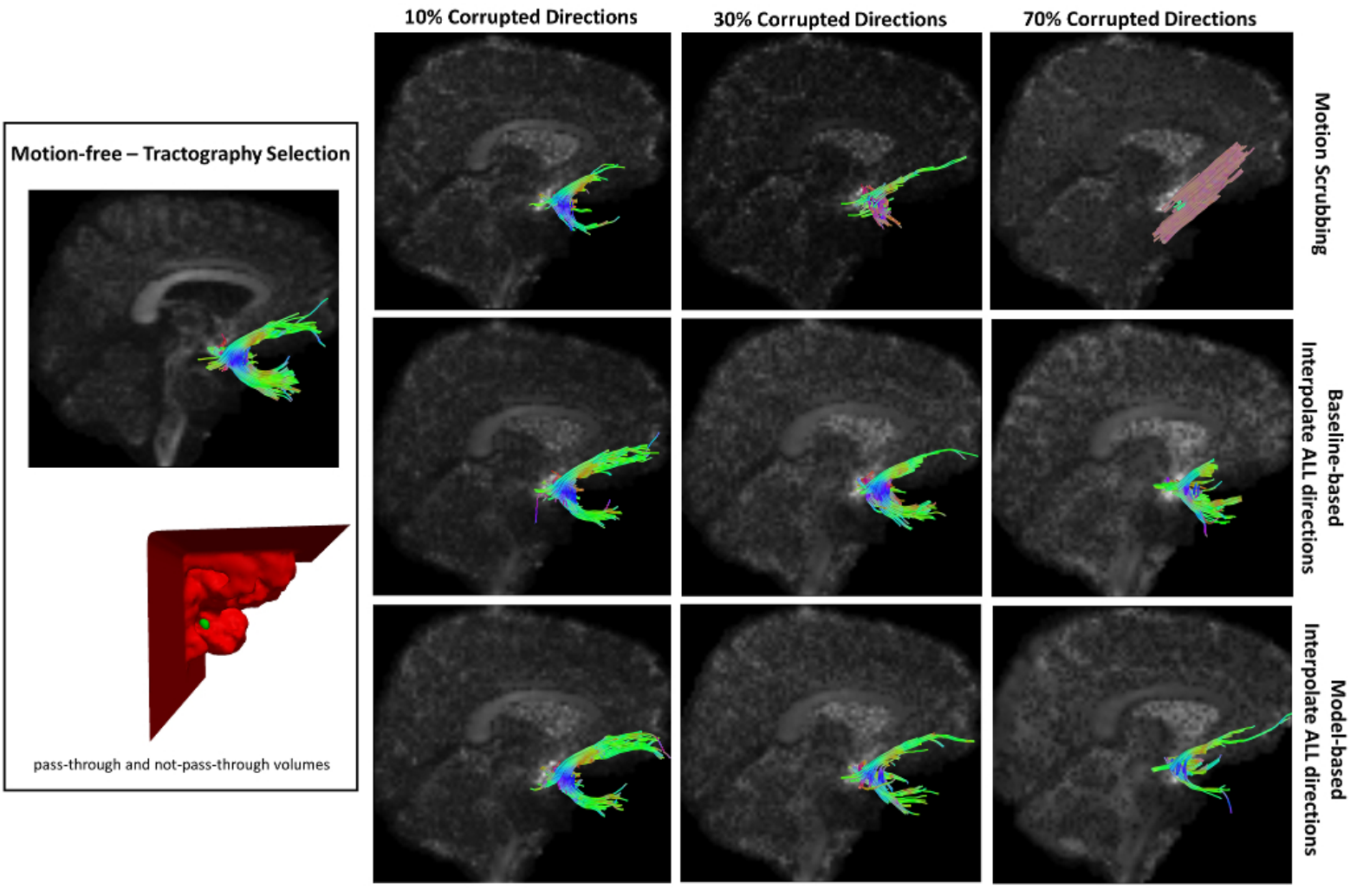


Figure 13.JPEG