PreQual: An automated pipeline for integrated preprocessing and quality assurance of diffusion weighted MRI images

Leon Y. Cai[1], Qi Yang[2], Colin B. Hansen[2], Vishwesh Nath[2], Karthik Ramadass[2], Graham W. Johnson[1], Benjamin N. Conrad[3,7], Brian D. Boyd[4,5], John P. Begnoche[4,5], Lori L. Beason-Held[6], Andrea T. Shafer[6], Susan M. Resnick[6], Warren D. Taylor[4,5], Gavin R. Price[7], Victoria L. Morgan[1,8,9,10], Baxter P. Rogers[1,4,8,10], Kurt G. Schilling[8,10], and Bennett A. Landman[1,2,4,8,10]

[1] Department of Biomedical Engineering, Vanderbilt University, Nashville, TN, USA
[2] Department of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, TN, USA
[3] Neuroscience Graduate Program, Vanderbilt Brain Institute, Vanderbilt University Medical Center, Nashville, TN, USA
[4] Department of Psychiatry and Behavioral Sciences, Vanderbilt University Medical Center, Nashville, TN, USA
[5] Center for Cognitive Medicine, Vanderbilt University Medical Center, Nashville, TN, USA
[6] Laboratory of Behavioral Neuroscience, National Institute on Aging, National Institutes of Health, Baltimore, MD, USA
[7] Department of Psychology and Human Development, Peabody College, Vanderbilt University, Nashville, TN, USA
[8] Department of Radiology and Radiological Sciences, Vanderbilt University Medical Center, Nashville, TN, USA
[9] Department of Neurological Surgery, Vanderbilt University Medical Center, Nashville, TN, USA
[10] Vanderbilt University Institute of Imaging Science, Vanderbilt University, Nashville, TN, USA

Correspondence

Leon Y. Cai
Medical-image Analysis and Statistical Interpretation (MASI) Lab
Vanderbilt University
leon.y.cai@vanderbilt.edu

Word Count

4290 words

Abstract

Purpose

Diffusion weighted MRI imaging (DWI) is often subject to low signal-to-noise ratios (SNRs) and artifacts. Recent work has produced software tools that can each correct an individual problem, but these tools have not been combined with each other or with quality assurance (QA). Thus, a single integrated pipeline is proposed to perform DWI preprocessing and tensor fitting with a spectrum of tools and produce an intuitive QA document.

Methods

The proposed pipeline is built around the FSL, MRTrix3, and ANTs software packages to perform DWI denoising; inter-scan intensity normalization; susceptibility-, eddy current-, and motion-induced artifact correction; slice-wise signal drop-out imputation; and tensor fitting. For QA, each operation is visualized alongside qualitative analyses, gradient and fractional anisotropy verifications, and a tensor goodness-of-fit analysis.

Results

Raw DWI data were preprocessed and quality checked with the proposed pipeline and demonstrated improved SNRs; physiologic intensity ratios; corrected susceptibility-, eddy current-, and motion-induced artifacts; imputed signal-lost slices; and improved tensor fits. The pipeline identified incorrect gradient configurations and file-type conversion errors and was shown to be effective on externally available datasets.

Conclusion

The proposed pipeline is a single integrated pipeline that combines established diffusion preprocessing tools from major MRI-focused software packages with intuitive QA.

Keywords

diffusion weighted imaging, DTI, preprocessing, denoising, distortion correction, quality assurance

Introduction

Diffusion weighted MRI imaging (DWI) is a powerful and noninvasive way of ascertaining the microstructural makeup of the brain (1). It forms the basis for many neurological studies including those investigating autism (2,3), aging (4,5), multiple sclerosis (6,7), and schizophrenia (8,9). It also forms the basis for studies of the structural human connectome (10) and has been increasingly used in context of neurosurgical planning and outcomes (11–14).

However, DWI images are often subject to low signal-to-noise ratios (SNRs) and a multitude of artifacts (15). The echo planar imaging acquisitions normally used to acquire DWI images can result in susceptibility-induced and eddy current-induced distortions as well as slice-wise signal drop-out (16). Imaging the brain repeatedly to acquire volumes along different gradient directions can result in long

acquisition times and increased subject movement. This can cause significant inter-volume motion artifacts (16). Acquisitions with "prescan" turned on can result in different gain settings within a session. This can cause the intensity distributions of DWI scans to violate their expected physical relationships. Intensity distributions of the same shell should be roughly equivalent and those of larger shells should be lower in intensity than those of smaller shells (15).

DWI image file formats present an additional opportunity for data corruption because they must store gradient configurations (b-values and b-vectors) in addition to voxel-wise intensities and coordinate space transformations. The scanner-based DICOM file format stores gradient information in the file header, but the processing-based NIFTI file format cannot. Thus, NIFTI files require separate "bval" and "bvec" text files to store this information. Converting DICOMs to NIFTIs can corrupt the data (17). For instance, some converters will reorder the volumes by shell or b-value while others do not. This can result in a mismatch between the NIFTI volumes and the gradients in the "bval" and "bvec" text files if the latter are not also reordered.

Fortunately, previous work in the field has resulted in many powerful tools across many software packages that can help correct these issues. MRTrix3's *dwidenoise* performs Marchenko-Pastur PCA denoising on DWI images which has been shown to significantly improve the quality of subsequent analyses (18–20). FSL's *topup* corrects susceptibility-induced distortions (21,22). The Synb0-DisCo deep learning framework creates a synthetic b = 0 s/mm^2 volume of infinite bandwidth from T1 images. This volume can be used with *topup* when the typically necessary forward and reverse, or "blip-up and blip-down", phase encoded DWI images are not otherwise available (23). FSL's *eddy* corrects eddy current-induced distortions and motion artifacts and imputes slices with signal drop-out (24,25). *Eddy* also rotates gradients to better align them with DWI signals. Last, MRTrix3's *dwigradcheck* determines the optimal DWI gradient configuration, defined as the gradient axis order and sign permutation that yields the highest average whole-brain streamline length (26).

As currently implemented, these tools exist in separate software packages and to date have not been integrated into one single DWI preprocessing pipeline. Additionally, there are only a few dedicated tools designed to perform intuitive quality assurance (QA) on the preprocessing operations and integrity of the data as a whole. One such QA pipeline, built by Lauzon et al. in 2013, analyzes the quality of DWI data and details the results in an intuitive document for user review. However, it was designed prior to the release of many of these preprocessing tools and thus does not perform extensive distortion correction (27). Another pipeline, *DTIPrep*, does not integrate susceptibility-induced distortion correction into its QA process and requires users learn to use its custom QA software in order to understand its results (28). Last, FSL's *eddyqc* performs QA on *eddy* (29) and details the results in an output document, but does not provide integration with or QA of other tools. Thus, herein we propose PreQual, a single pipeline designed to fill this gap. In summary, we leverage these powerful software tools from a variety of packages and combine them to produce preprocessed DWI and diffusion tensor imaging (DTI) outputs ready for further analysis. In addition, we produce a QA report that details the efficacy of each preprocessing step and supplies high-level QA on the final preprocessed data output by the pipeline.

Methods

We built PreQual around the MRTrix3 (30), FSL (31), and ANTs (32) software packages, MRI-focused tools popular in the field, in order to perform integrated preprocessing, artifact correction, and intuitive QA (Figure 1).

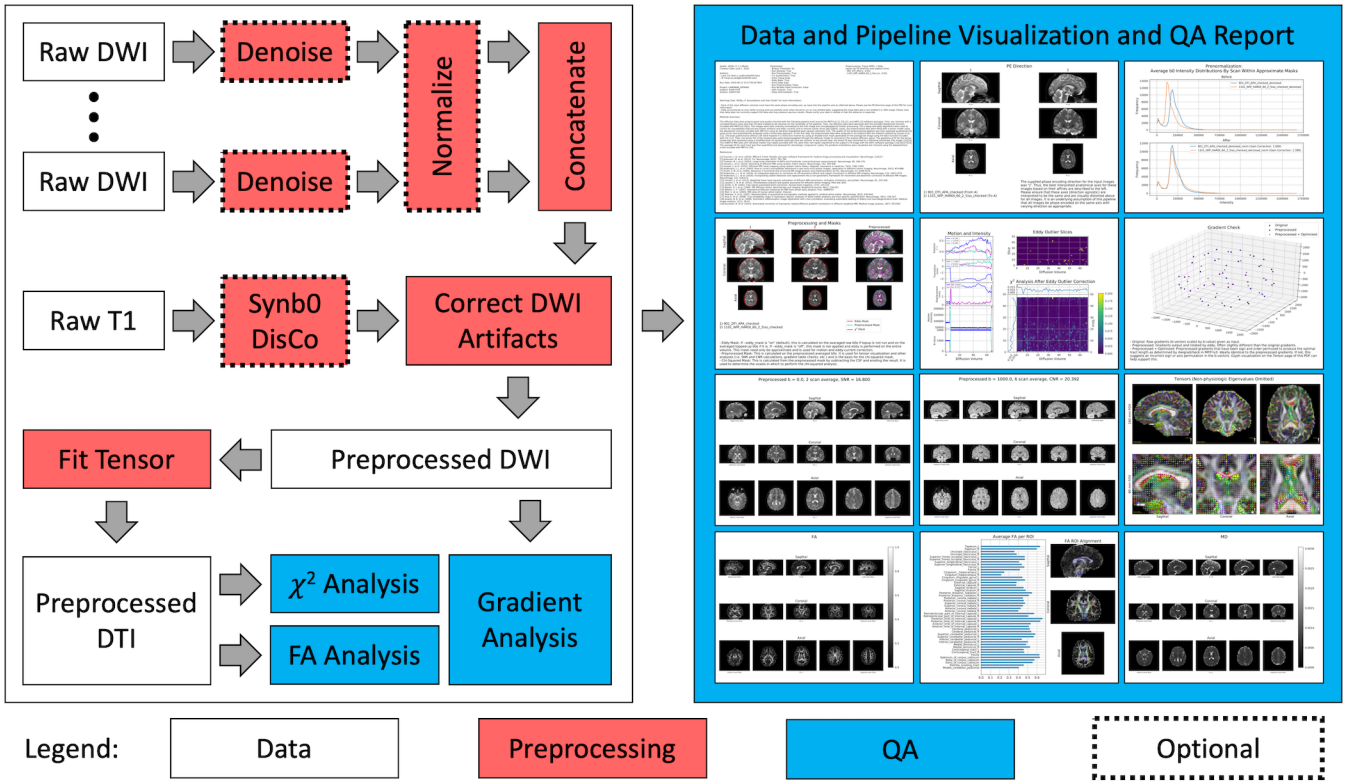

Figure 1. Overview of the PreQual pipeline and QA document. The pipeline takes as input raw DWI data, denoises them, normalizes their intensities, concatenates them, and corrects for artifacts. If no complementary (i.e., forward and reverse) phase encoded images are supplied, PreQual uses Synb0-DisCo to perform susceptibility-induced distortion correction. The pipeline fits the preprocessed DWI data to a tensor to produce a DTI output. QA checks on each preprocessing step and high-level QA checks on the final outputs of the pipeline are then performed. These results are presented in a QA report. All preprocessing steps except for distortion correction and tensor fitting are optional and can be turned off per user preference.

Raw DWI Inputs

We take a series of raw DWI images as input into the pipeline along with a phase encoding axis and a configuration file. We envision the input images to be from the same acquisition session, phase encoded along the same axis. We define the axis as the image dimension along which phase encoding occurred. In the configuration file, we require the user to input the directions along the axis in which the images were encoded and their corresponding readout times. The phase encoding schemes of each image (i.e., axis, direction, and readout time) are required for susceptibility-induced distortion correction with FSL's *topup* (21,22).

We note that the supplied phase encoding axis of the inputs may not always correctly correspond to the expected anatomical axis. For instance, DWI images phase encoded along the anterior to posterior axis and oriented radiologically (right-anterior-superior) or neurologically (left-anterior-superior) would be encoded along the second dimension. Yet, if the image is oriented differently and anterior to posterior phase encoding is expected, the user may need to indicate a different dimension for the phase encoding axis. Thus, for QA, we check that the supplied axis and direction for each image translates to the expected anatomical axis and direction using each image's coordinate transformation. We visualize triplanar slices of the raw data alongside these interpretations for users to verify the presence of susceptibility-induced distortions along the anatomical axes (Supplementary Figure 1).

Denoising

We individually denoise each input image using the Marchenko-Pastur PCA technique as implemented in MRTrix3 (18–20). For QA, we calculate the median intra-brain voxel-wise SNR for the non-diffusion weighted b = 0 s/mm^2 volumes (Supplementary Figure 7). We define SNR in alignment with FSL's *eddy*, as $\mu_{voxel}/\sigma_{voxel}$ or the mean voxel-wise intensity divided by its standard deviation. When only one b = 0 s/mm^2 volume is available, we report the SNR as "Not a Number".

Normalization

We correct the denoised input images for inter-scan gain differences. To achieve this, for the kth input image, we take the average b = 0 s/mm^2 volume, V_k, and mask the result with FSL's brain extraction tool (*bet*) (33). We then determine a multiplicative scale factor, s_k, such that the histogram intersection, I, between the intra-mask intensity histogram, h, of V_k scaled by s_k and that of the first image is maximized. When $k = 1$, $s_1 = 1$. When $k > 1$, we calculate s_k by performing optimization on the negated problem (Eq. 1) with the downhill simplex algorithm with 10 initial guesses evenly spaced between 0.5 and 1.5 (34).

$$s_k = \underset{s_k}{\operatorname{argmin}} -I\big(h(s_k V_k), h(V_1)\big) \tag{1}$$

During each iteration of the algorithm, we calculate the histograms $h(s_k V_k)$ and $h(V_1)$ with 100 bins from a to b where $a = \min(V_1, s_k V_k)$, the minimum intra-mask intensity between the two average b = 0 s/mm^2 volumes, and b is the corresponding maximum. To complete normalization, we intensity-scale all volumes, diffusion-weighted or otherwise, of the kth image by s_k for all k. For QA, we report the calculated scale factors and plot of the average b = 0 s/mm^2 intensity histograms of each image before and after normalization (Supplementary Figure 2).

Artifact Correction

Following normalization, we concatenate the images and perform artifact correction. First, we correct susceptibility-induced distortions using FSL's *topup* tool (21,22). We generate the *topup* acquisition parameters file using the input phase encoding schemes provided and then run *topup* on all the b = 0 s/mm^2 volumes from all input scans. Typically, *topup* requires complementary (i.e., both forward and reverse or "blip-up" and "blip-down") phase encoded images. In the event that the user does not supply complementary images, we use Synb0-DisCo to create a susceptibility-corrected synthetic b = 0 s/mm^2 volume of infinite bandwidth for *topup* instead (23). When we run Synb0-DisCo, we require that the user input a T1 image of the same subject into the pipeline, and we visualize triplanar views of the T1 and synthetic b = 0 s/mm^2 volume in the output QA document (Supplementary Figure 3).

Following *topup*, we use FSL's *eddy* with a brain mask estimated by *bet* on the averaged *topup* output to correct for eddy current-induced distortions, inter-volume motion, and slice-wise signal drop-out (24,25). During this process, *eddy* also performs rotation of the gradients to best suit the DWI signals and is configured to estimate the voxel-wise SNR across b = 0 s/mm^2 volumes and contrast-to-noise ratio (CNR) for each diffusion-weighted shell. *Eddy* defines SNR as $\mu_{voxel}/\sigma_{voxel}$, or the mean voxel-wise value divided by its standard deviation. *Eddy* defines CNR as $\sigma_{GP}/\sigma_{residual}$, or the voxel-wise standard

deviation of the Gaussian process predicted by *eddy* divided by that of the residuals. A final brain mask is calculated with *bet* (33). This results in the pipeline's final preprocessed DWI data.

For QA of the artifact correction steps, we take a combined quantitative and qualitative approach. First, we use the final mask to calculate the median intra-mask SNR for b = 0 s/mm^2 volumes and CNR for each diffusion weighted shell. We display these ratios alongside five central triplanar slices of each corresponding shell-wise averaged preprocessed volume (Supplementary Figure 7). This provides a quantitative metric for SNR for the denoising step as well as qualitative verification of artifact correction. In addition, we plot the *eddy*-calculated inter-volume rotation, translation, and displacement alongside a plot indicating the slices *eddy* imputed due to signal drop-out (Supplementary Figure 5).

Tensor Fitting

We fit the preprocessed DWI volumes to a tensor model using an iterative reweighted least-squares estimator as implemented in MRTrix3 (35). This produces the final preprocessed DTI output of the pipeline (36). We convert the DTI output to fractional anisotropy (FA) and mean diffusivity (MD) scalar maps (37). For QA, we visualize the tensor glyphs (Supplementary Figure 8) as rendered in MRTrix3 (38) and five central triplanar slices of the FA and MD maps in the output document (Supplementary Figure 9).

Gradient Analysis

To determine whether the DWI gradient directions are properly oriented, we calculate the optimal permutation of the preprocessed DWI gradient directions output by *eddy.* We define optimality as it is defined in MRTrix3: the permutation of gradient sign and axis order that produces the highest average whole-brain streamline length (26). We plot the optimal preprocessed gradient permutation, the preprocessed gradient directions as output by *eddy*, and the original input gradient directions (Supplementary Figure 6) in the output QA document. We also visualize triplanar slices of the tensor glyphs (Supplementary Figure 8). We expect properly oriented gradients to be identifiable with this visualization, as the preprocessed and optimal gradients would overlap and produce physiologically oriented tensors.

Fractional Anisotropy Analysis

As a quality check on the preprocessed DTI data, we calculate the average FA value for all 48 white matter regions of interest (ROIs) defined by the Johns Hopkins ICBM DTI 81 white matter atlas (39–41). We deformably register the atlas to the subject's FA image space using the ANTs software package (42) to isolate each region and its average FA value. For QA, we plot these values along with an overlay of the atlas on the subject's FA map. The former allows users to verify FA congruence with expected physiologic values, and the latter allows users to check the registration process (Supplementary Figure 10).

Chi-Squared Analysis

To measure the goodness-of-fit of the tensor model, we perform a modified pixel-wise chi-squared analysis as described by Lauzon et al. on the preprocessed DTI data (27). We constrain this analysis to the brain parenchyma in order to avoid analyzing tensors fit to background or CSF that would produce invalid chi-squared values. To mask the parenchyma, we erode the final mask calculated on the preprocessed

DWI data with a radius of one pixel and remove the CSF. To identify the CSF, we calculate a CSF mask for each shell-wise averaged volume of the preprocessed DWI data with FSL's *fast* tool (43). We then generate an overall CSF probability mask by performing a voxel-wise average across shells and setting a threshold at 15%. For each slice in each volume, we divide the intra-parenchyma sum squared error of tensor fit by the sum squared intensity of the slice to calculate the goodness-of-fit (27). For QA, we visualize the chi-squared values in the QA document (Supplementary Figure 5).

Mask Quality Assurance

In order to quality check the mask used for *eddy*, the final preprocessed mask used for median SNR and CNR calculation, and the parenchyma mask used for the chi-squared analysis, we plot their contours. Specifically, we overlay them on triplanar views of the first b = 0 s/mm^2 volume of each raw input image and that of the preprocessed DWI output (Supplementary Figure 4).

Study Overview

We evaluate the efficacy of each pipeline stage by running the pipeline on one of two DWI datasets acquired on a Phillips 3T scanner at Vanderbilt University. The first dataset, Dataset A, was taken from one subject scanned repeatedly on the same scanner over the course of three sessions, one session each consecutive day. Each session consisted of 6 sets of scans. Each set of scans consisted of one 3-direction b = 1000 s/mm^2 image phase encoded in the anterior to posterior direction (APA), one 96-direction b = 1000 s/mm^2 image phase encoded in the posterior to anterior direction (APP), one 96-direction b = 1500 s/mm^2 APP image, one 96-direction b = 2000 s/mm^2 APP image, one 96-direction b = 2500 s/mm^2 APP image, and one 96-direction b = 3000 s/mm^2 APP image. The second dataset, Dataset B, was taken from another subject scanned once. It consists of a 6-direction b = 1000 s/mm^2 APA image and one 60-direction b = 2000 s/mm^2 APP image. All images were deidentified and all scans were acquired only after informed consent under supervision of the project Institutional Review Board.

To study the denoising step, we calculate changes in SNR of the preprocessed DWI data with and without denoising. To study the normalization step, we introduce artificial inter-scan gain differences and observe how the pipeline corrects them. To study the artifact correction step, we qualitatively visualize representative volumes that display the DWI data before and after artifact correction. To study the gradient analysis, we improperly orient the input gradients by negating them in the second dimension and visualize how the pipeline identifies this error. To study one use case of the FA analysis, we corrupt the input DWI data by shuffling the order of the image volumes without shuffling their corresponding gradients to simulate failed file-type conversion and observe how the pipeline identifies this error. To study the chi-squared analysis, we perform the same tensor fit on the raw input data as is performed on the DWI outputs and measure the differences in goodness-of-fit.

In addition, we study the generalizability of PreQual by running it on one imaging session from each of three externally available datasets: the Human Connectome Project (HCP) Lifespan cohort (44), the Autism Brain Imaging Data Exchange (ABIDE) II (2), and the Baltimore Longitudinal Study of Aging (BLSA) (45,46). Last, we quantify the runtime of PreQual on three different configurations of Dataset B.

Results

Denoising

To study the impact of the denoising step, we ran PreQual on each session of Dataset A and combined the results for analysis. We calculated the median intra-mask pixel-wise SNR for each shell and display them in Figure 2. Qualitative visualization showed improved SNR for each diffusion weighted shell which were each shown to be statistically significant at $p < 0.001$ (Wilcoxon signed rank test).

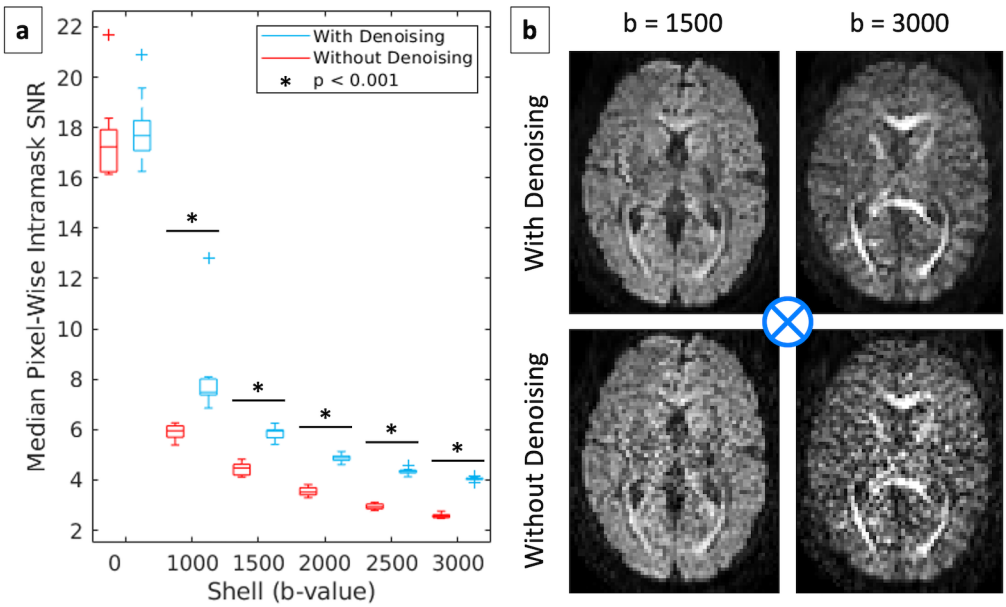

Figure 2. PreQual improves the signal-to-noise ratio (SNR) of diffusion images. (a) The shell-wise SNR for all diffusion weighted shells increased in the preprocessed DWI outputs of PreQual following Marchenko-Pastur PCA denoising of the input data as implemented in MRTrix3 compared to preprocessing without denoising (Wilcoxon signed rank test, $p < 0.001$). (b) Visualization of representative volumes at shells b = 1500 s/mm^2 and b = 3000 s/mm^2 with gradient direction x = 0.002, y = 0.005, and z = 0.999 (best modeled as a unit vector going into the page as shown with the blue fletching) illustrates qualitatively improved SNR.

Normalization

To study the efficacy of the normalization step, we ran PreQual on Dataset B while simulating inter-scan gain changes. When different images have the same gain settings, we expect intensity distributions of larger shells to be lower than those of smaller shells and those of the same shell to be roughly equivalent (15). We use each volume's median intra-mask intensity as a representative measure for its intensity distribution, and we intentionally violate both of these expectations by multiplying the b = 2000 s/mm^2 image of Dataset B by four. This caused the median intensities of the b = 2000 s/mm^2 volumes to be higher than those of the b = 1000 s/mm^2 volumes (Figure 3b). It also caused the median intensities of the b = 0 s/mm^2 volumes to be different from each other (Figure 3a). We show that the expected relationships return after normalization with PreQual (Figures 3c and 3d).

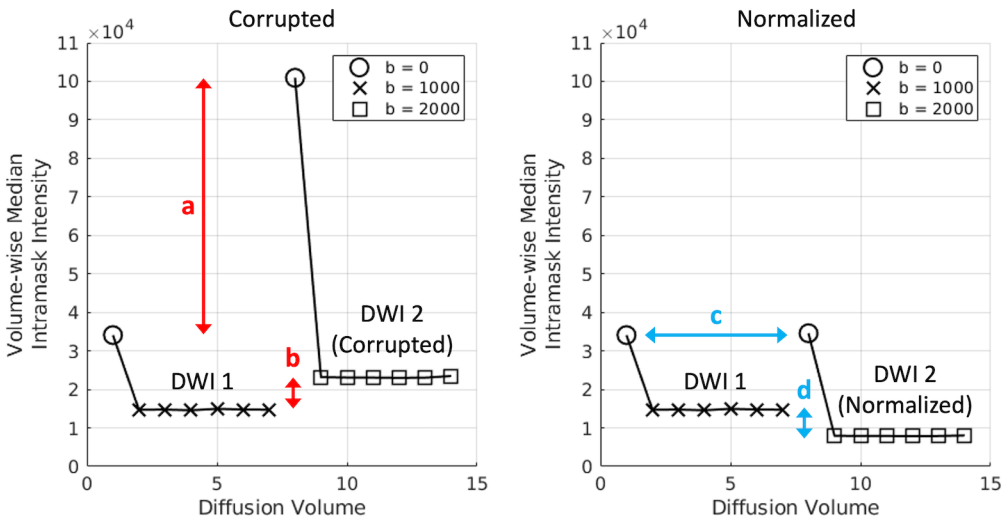

Figure 3. PreQual corrects physically impractical shell-wise intensity values in simulation. Two DWI images of the same session were used as input for this simulation. DWI 1 was a 6-direction b = 1000 s/mm^2 image with an associated b = 0 s/mm^2 volume, and DWI 2 was a 60-direction b = 2000 s/mm^2 image with an associated b = 0 s/mm^2 volume (only the first 6 volumes are shown for simplicity). DWI 2 was corrupted via global intensity-scaling by a factor of 4, emulating gain changes between scans of the same session. This resulted in physically impractical inter-shell intensity relationships. (a) Scans of the same shell (or b-value) are expected to have roughly the same intensity distributions, a phenomenon which is violated here in simulation. The b = 0 s/mm^2 volumes are expected to have roughly the same median intensity. (b) Shell size and intensity are inversely related which is violated here in simulation. The b = 2000 s/mm^2 volumes are expected to have lower median intensity than the b = 1000 s/mm^2 volumes. (c and d) The expected phenomena are observed after normalization with PreQual.

Artifact Correction

To study the effectiveness of susceptibility-, eddy current-, and motion-induced distortion correction and slice-wise signal drop-out imputation by PreQual, we ran the pipeline on both Datasets A and B and visualized representative inputs and outputs of each operation (Figures 4a, 4c, 4d, and 4e). We also ran PreQual on Dataset B with the b = 1000 s/mm^2 APA image omitted, thus destroying the necessary complementary phase encoded pair for *topup* and triggering Synb0-DisCo to run. We visualized the inputs and outputs of this operation as well (Figure 4b). Qualitative visualization demonstrated cases where artifact correction was needed as well as the improvement of the data with PreQual.

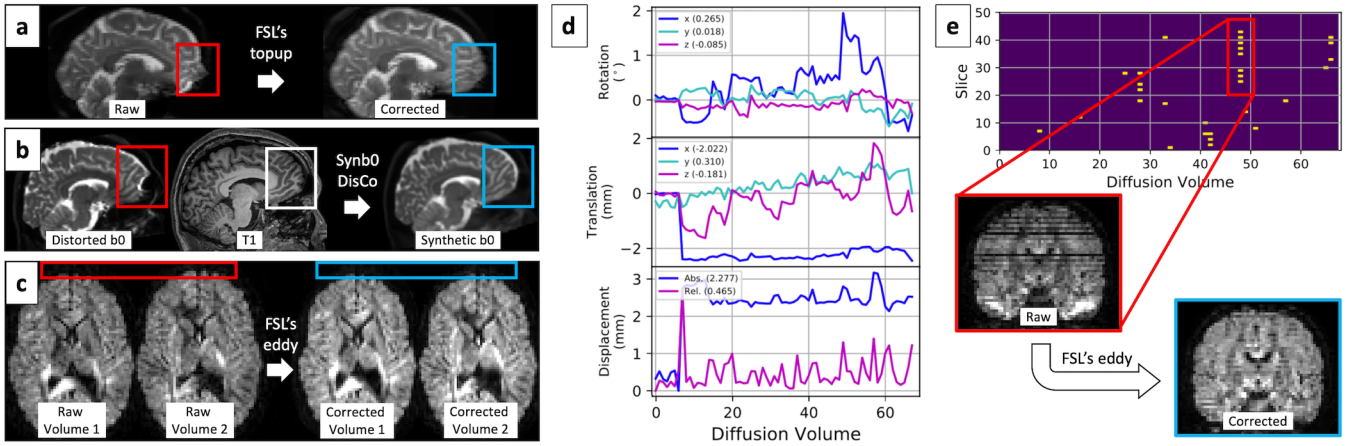

Figure 4. PreQual performs artifact correction. (a) Susceptibility-induced distortions are corrected using FSL's *topup* in conjunction with input complementary (i.e., forward and reverse) phase encoded images. An example of this distortion before and after correction with *topup* is shown in the red and blue bounding boxes, respectively. (b) If no complementary phase encoded images are available, PreQual uses Synb0-DisCo to generate a synthetic distortion-free b = 0 s/mm^2 volume from a distorted one and a T1 image. The synthetic volume is used with *topup* to correct the image as a whole. The bounding boxes highlight an example of how Synb0-DisCo generates distortion-free regions (blue) by using distorted b = 0 s/mm^2 information (red) and distortion-free T1 information (white). (c) Eddy current-induced inter-volume geometric distortions as highlighted between raw volumes 1 and 2 (red) are corrected with FSL's *eddy* as shown between corrected volumes 1 and 2 (blue). (d) FSL's *eddy* estimates and corrects for inter-volume motion. An example output in the PreQual QA report detailing this movement is shown. The overall average movement across volumes is shown in the legend parentheses. For displacement, legend entry "Abs." is short for absolute, and "Rel." is short for relative. (e) *Eddy* identifies slice-wise signal drop-out and imputes lost slices accordingly. An example output plot in the PreQual report details the affected slices. The corresponding raw volume with drop-out slices is shown highlighted in red and the imputed output of *eddy* is shown highlighted in blue. Note that the imputation improves data quality, but that the overall process is still quite motion-sensitive, explaining some of the remaining artifacts.

Gradient Analysis

We ran PreQual on Dataset B in two configurations. In the first configuration, the data were not corrupted. In the second, all gradients were negated in the second dimension. We visualize the resultant tensor glyphs in the splenium of the corpus callosum for both cases and show them to be oriented in physiologically probable and improbable orientations, respectively. We also visualize the optimal gradient tables as determined by PreQual for both cases and show them to be overlapping and non-overlapping with those output in the preprocessed DWI data, respectively. These two visualizations together allow us to correctly identify the case with improperly oriented gradients (Figure 5).

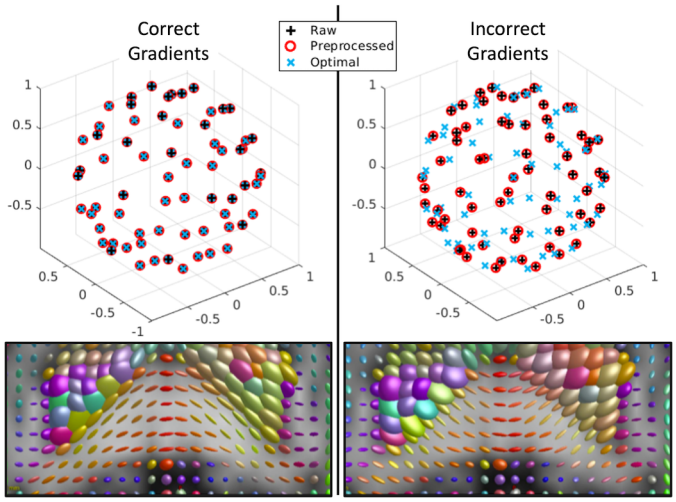

Figure 5. PreQual identifies improperly oriented gradient tables in simulation. The left side of the figure displays expected phenomenon. The optimal gradient configurations are identical to those in the preprocessed DWI output of PreQual. The unit gradient vectors are overlapping, and the tensors calculated from the preprocessed gradients are oriented in a physiologically probable manner, as shown in an axial slice through the splenium of the corpus callosum. The right side of the figure displays the identification of improperly oriented gradients. The preprocessed gradients do not overlap with the optimal ones and the tensor glyphs are oriented along the corpus callosum in a physiologically improbable manner. The raw input gradients are shown here to match what is reported in the QA report such that users can visualize rotational optimization made to input gradients during preprocessing.

Fractional Anisotropy Analysis

To demonstrate one use case for the FA analysis, we simulated a failed file-type conversion that led to the mismatch between image volumes and gradient tables. We took the b = 0 s/mm^2 volume acquired at the beginning of the b = 2000 s/mm^2 APP scan of Dataset B out of its original position and placed in the middle of the sequence without reordering the corresponding gradients. We demonstrate that the average ROI-based FA exhibit an entirely different cross-region profile with this volume-to-gradient mismatch than without (Figure 6a). This is supported by qualitative visualization of the FA maps (Figure 6b) and together result in the identification of faulty file-type conversions.

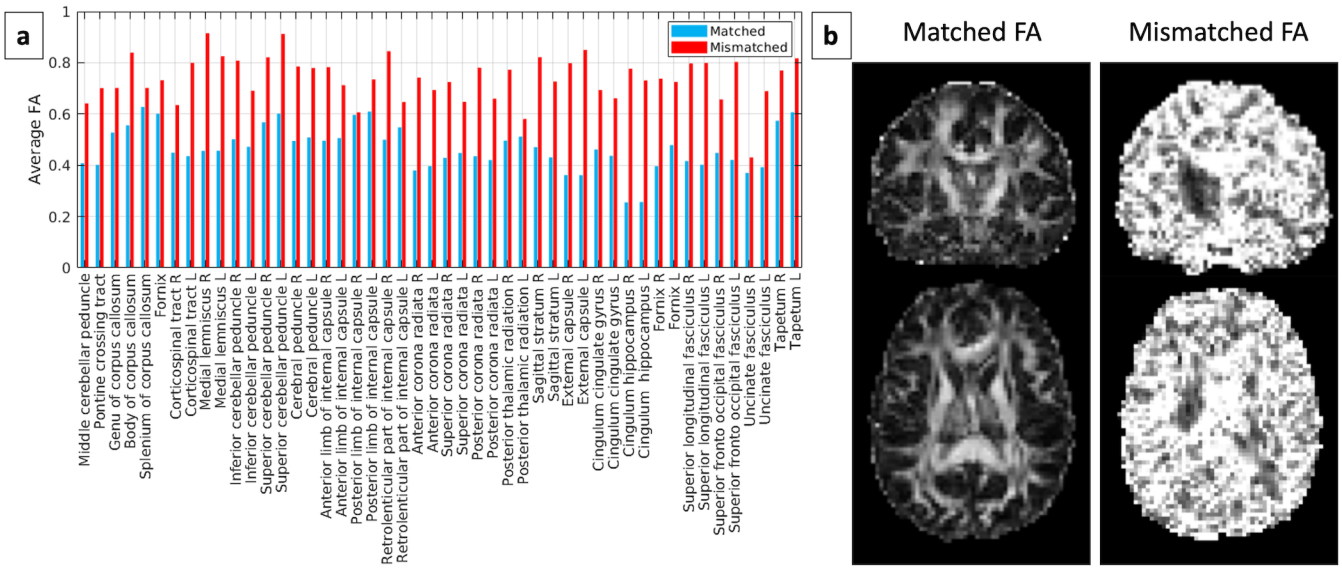

Figure 6. PreQual identifies simulated faulty file-type conversions that lead to volume-to-gradient mismatch. Volume-to-gradient mismatch was achieved by shuffling b = 0 s/mm^2 volumes out of their original position in the DWI sequence without corresponding shuffling of gradient information. (a) The average FAs by white matter region are higher across all regions and exhibit an identifiably different cross-region profile with volume-to-gradient mismatch. (b) Qualitative examination of FA maps with and without mismatch provides additional verification.

Chi-Squared Analysis

We show a chi-squared goodness-of-fit analysis in Figure 7. We performed the analysis on Dataset B and measured the chi-squared values calculated on the raw images and on those processed with PreQual. Since chi-squared is a measurement of error, the lower values in the data processed with PreQual suggest improved tensor fitting of the data. This improvement was shown to be statistically significant at $p < 0.001$ (Wilcoxon signed rank test).

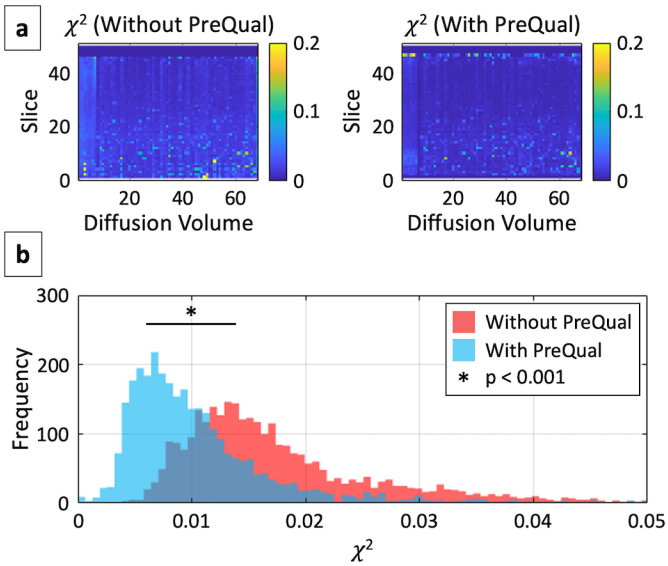

Figure 7. PreQual improves tensor fitting of diffusion weighted images. (a) The modified pixel chi-squared for a multi-shell acquisition with and without PreQual applied are presented per slice per volume. (b) The distributions of the chi-squared values in (a) are presented and demonstrate improved goodness-of-fit with statistical significance (Wilcoxon signed rank test, $p < 0.001$).

External Verification

We ran the pipeline on one imaging session from each of three externally available datasets, as shown in Figure 8, to demonstrate the generalizability of PreQual. We show visually apparent susceptibility-induced distortions corrected in the HCP Lifespan (44) image, slice-wise signal drop-out imputed in the ABIDE II (2) image, and improperly oriented gradients identified in the BLSA (45,46) image, as have been previously found to exist in some uncorrected BLSA scans (47). The full QA documents generated by PreQual corresponding to these three datasets are available in the supplement.

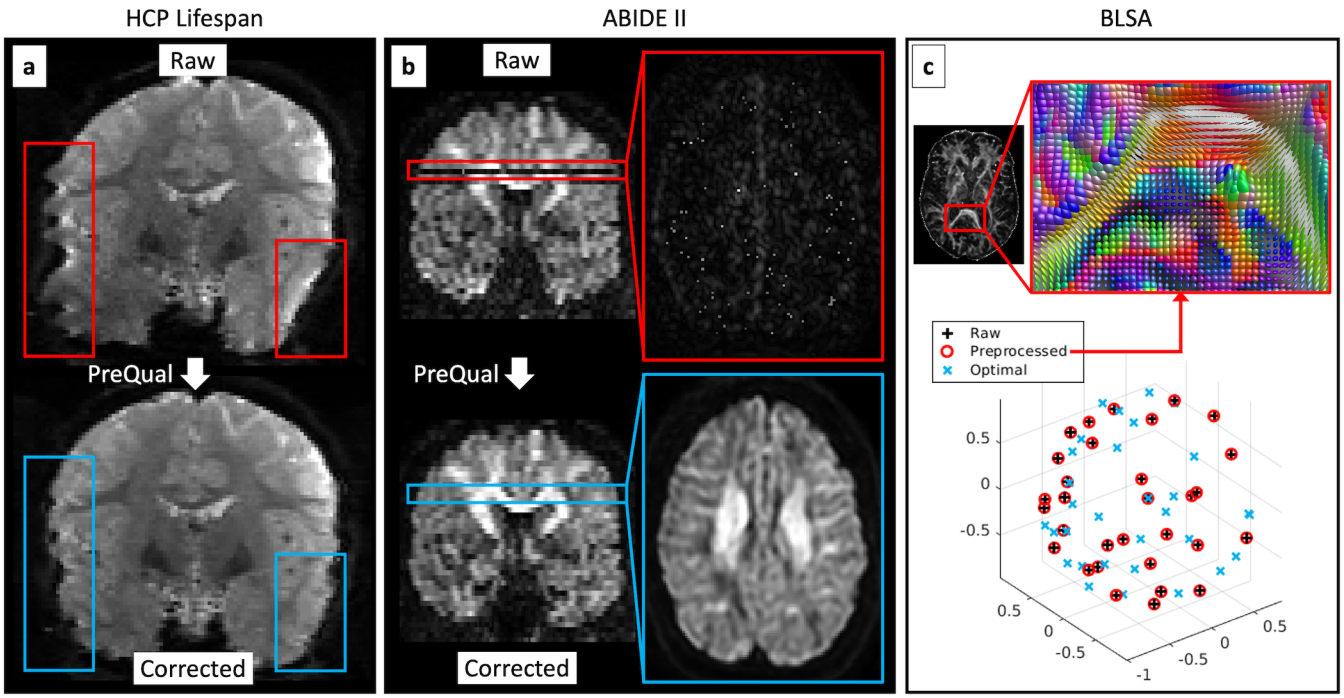

Figure 8. PreQual performs artifact correction and error identification on three externally available datasets. (a) For an HCP Lifespan image, PreQual corrected visually apparent susceptibility-induced distortions. The red and blue bounding boxes highlight corresponding regions before and after preprocessing by PreQual, respectively. (b) For an ABIDE II image, PreQual identified and corrected visually apparent slice-wise signal drop out. The red and blue bounding boxes highlight corresponding slices before and after preprocessing by PreQual, respectively. (c) For an uncorrected BLSA image, PreQual correctly identified improperly oriented gradients as have been previously documented and repaired (47) via visualization of preprocessed gradients against optimal ones and visualization of tensor glyphs.

Typical Runtimes

We quantified the runtime of PreQual on three different configurations of Dataset B (Table 1).

Table 1. Typical runtimes of PreQual on three different configurations of Dataset B.

Input Images	Susceptibility-induced Distortion Correction	Runtime
• 60-direction b = 2000 s/mm^2 APP	None	18m 12s
• 60-direction b = 2000 s/mm^2 APP • T1 weighted	With a synthetic b = 0 s/mm^2 volume	58m 47s
• 60-direction b = 2000 s/mm^2 APP • 6-direction b = 1000 s/mm^2 APA	With a reverse phase encoded b = 0 s/mm^2 volume	48m 30s

First, we ran the pipeline without susceptibility-induced distortion correction on only the 60-direction b = 2000 s/mm^2 APP image and its associated b = 0 s/mm^2 volume. Second, we ran it with susceptibility-induced distortion correction informed by Synb0-DisCo on the same image and a T1 weighted image. Last, we ran it with traditional susceptibility-induced distortion correction on the 6-direction b = 1000 s/mm^2 APA image, the 60-direction b = 2000 s/mm^2 APP image, and their associated b = 0 s/mm^2 volumes. We measured the runtimes to be 18 minutes and 12 seconds, 58 minutes and 47 seconds, and 48 minutes and 30 seconds, respectively. Processing was performed on a desktop computer running 64-bit Ubuntu 18.04 with 24 GB RAM and a 4-core 3.5 GHz Intel(R) Xeon(R) E5 processor with multithreading. We find that under these conditions, the susceptibility-induced distortion correction step with FSL's *topup* was the most time-demanding step.

Discussion and Conclusions

Here, we present PreQual, a single integrated pipeline for DWI preprocessing and QA. We show that by leveraging tools across different software packages, we are able to apply a multitude of different preprocessing techniques to DWI data. We demonstrate that each individual step of the pipeline contributes to the quality of the output data and that the pipeline as a whole produces data that can be better fit to a tensor model for subsequent analysis. We show that QA documentation, both quantitative and qualitative, can be used to demonstrate how raw DWI data is improved with these preprocessing steps. Last, we show that an analysis of gradient orientations and ROI-based average FA can identify problems with the storage and conversion of DWI data.

We envision the use of PreQual as a replacement for traditional DWI preprocessing and QA methods. Regarding preprocessing, a single pipeline reduces the need for data and software manipulation and thus reduces the opportunity for errors and data corruption. In addition, it simplifies the process as a whole. Regarding QA, current methods can be tedious. These include needing to open individual DWI images with specialized software and needing to loop through the individual volumes manually, looking for artifacts and other errors. Other approaches include needing to hand-draw multiple ROIs for investigation and needing to calculate and visualize tensors glyphs and scalar maps manually (15). Instead, we envision that all scans in a session can be run through PreQual and visualization of the output QA report with a common PDF reader can provide an intuitive way for users to understand their data. This approach is centered around the QA document detailing each preprocessing step and the data as a whole. For instance, we display calculated SNR and CNR, intensity distributions, summary motion and slice imputation statistics, gradient and ROI-based FA visualizations, and chi-squared measures of tensor fits, providing thorough quantitative documentation. In addition, our approach displays a multitude of raw and preprocessed triplanar slices, FA and MD maps, tensor glyphs, and mask contours to allow users to qualitatively understand their data. Our intention is that users can preprocess their data with one command and perform QA by visualization of the output document as opposed to relying on traditional, often tedious, methods.

A related QA tool, *eddyqc* (29), exists to interrogate FSL's *eddy*. It provides a document that details the motion, slice-wise imputation, and other distortion-correction operations performed by FSL. With PreQual, we build on this by providing preprocessing tools from other software packages and a way to quality check other aspects of diffusion preprocessing in addition to distortion correction. We note that *eddyqc* can perform study-wise QA, something that PreQual was not explicitly designed to do. However, we save all information needed to do so in the pipeline's outputs should the user require it.

All information presented in this manuscript was extracted directly from either pipeline outputs or from the QA documents. The pipeline was created with this design in mind in order to facilitate interrogation of the preprocessing steps for users to understand what is happening to their data. The only data reported here that are explicitly not reported in the PreQual document are the shell-wise SNRs for non-zero shells, as shown in Figure 2. Instead, in line with the *eddyqc* tool, we report the median intra-mask shell-wise CNR calculated by *eddy* in order to represent the ratio between the diffusion-induced signal variability and non-related signal variability as a measure of *eddy*'s success (29).

Last, PreQual was designed to be modular: each preprocessing step aside from artifact correction and tensor fitting can be turned off with command line options. For instance, should users choose to not perform denoising or normalization, they can do so. In these scenarios, the QA document still reports the corresponding QA metrics, like SNR and the b = 0 s/mm^2 intensity distributions and calculated scale factors that would have been necessary for normalization. Thus, users can still use PreQual to identify potential noise or intensity issues with their data, even without performing the corresponding preprocessing steps. PreQual is also designed to take additional parameters for FSL's *topup* and *eddy*, should users want to use more advanced features, like slice-wise intra-volume motion correction (48) or dynamic susceptibility estimation (49).

The PreQual source code, accompanying documentation, and a Singularity definition file for containerization have been made available to enable evaluation of the proposed pipeline at github.com/MASILab/PreQual.

Acknowledgements

This work was conducted in part using the resources of the Advanced Computing Center for Research and Education at Vanderbilt University, Nashville, TN. This work was supported by the National Institutes of Health (NIH) under award numbers 5R01EB017230, 5T32EB001628, 5T32GM007347, 1UL1RR024975, 5R01NS110130, and 5R01NS108445. This work was also supported by the National Science Foundation under award number 1452485. This research was conducted with the support from the Intramural Research Program of the National Institute on Aging of the NIH. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

References

1. O'Donnell LJ, Westin CF. An introduction to diffusion tensor image analysis. Neurosurg. Clin. N. Am. 2011;22:185–196 doi: 10.1016/j.nec.2010.12.004.
2. Di Martino A, O'Connor D, Chen B, et al. Enhancing studies of the connectome in autism using the autism brain imaging data exchange II. Sci. Data 2017;4:170010 doi: 10.1038/sdata.2017.10.
3. Travers BG, Adluru N, Ennis C, et al. Diffusion Tensor Imaging in Autism Spectrum Disorder: A Review. Autism Res. 2012;5:289–313 doi: 10.1002/aur.1243.
4. Westlye LT, Walhovd KB, Dale AM, et al. Life-span changes of the human brain white matter: Diffusion tensor imaging (DTI) and volumetry. Cereb. Cortex 2010;20:2055–2068 doi: 10.1093/cercor/bhp280.
5. Zavaliangos-Petropulu A, Nir TM, Thomopoulos SI, et al. Diffusion MRI indices and their relation to cognitive impairment in brain aging: The updated multi-protocol approach in ADNI3. Front. Neuroinform. 2019;13:2 doi: 10.3389/fninf.2019.00002.
6. Inglese M, Bester M. Diffusion imaging in multiple sclerosis: Research and clinical implications

Jensen JH, Helpern JA, editors. NMR Biomed. 2010;23:865–872 doi: 10.1002/nbm.1515.
7. De Santis S, Bastiani M, Droby A, et al. Characterizing Microstructural Tissue Properties in Multiple Sclerosis with Diffusion MRI at 7 T and 3 T: The Impact of the Experimental Design. Neuroscience 2019;403:17–26 doi: 10.1016/j.neuroscience.2018.03.048.
8. Kubicki M, McCarley R, Westin CF, et al. A review of diffusion tensor imaging studies in schizophrenia. J. Psychiatr. Res. 2007;41:15–30 doi: 10.1016/j.jpsychires.2005.05.005.
9. Cetin-Karayumak S, Di Biase MA, Chunga N, et al. White matter abnormalities across the lifespan of schizophrenia: a harmonized multi-site diffusion MRI study. Mol. Psychiatry 2019 doi: 10.1038/s41380-019-0509-y.
10. Sporns O, Tononi G, Kötter R. The human connectome: A structural description of the human brain. PLoS Comput. Biol. 2005;1:0245–0251 doi: 10.1371/journal.pcbi.0010042.
11. Bello L, Gambini A, Castellano A, et al. Motor and language DTI Fiber Tracking combined with intraoperative subcortical mapping for surgical removal of gliomas. Neuroimage 2008;39:369–382 doi: 10.1016/j.neuroimage.2007.08.031.
12. Golby AJ, Kindlmann G, Norton I, Yarmarkovich A, Pieper S, Kikinis R. Interactive diffusion tensor tractography visualization for neurosurgical planning. Neurosurgery 2011;68:496–502 doi: 10.1227/NEU.0b013e3182061ebb.
13. Raffa G, Conti A, Scibilia A, et al. The impact of diffusion tensor imaging fiber tracking of the corticospinal tract based on navigated transcranial magnetic stimulation on surgery of motor-eloquent brain lesions. Neurosurgery 2018;83:768–782 doi: 10.1093/neuros/nyx554.
14. Costabile JD, Alaswad E, D'Souza S, Thompson JA, Ormond DR. Current applications of diffusion tensor imaging and tractography in intracranial tumor resection. Front. Oncol. 2019;9:426 doi: 10.3389/fonc.2019.00426.
15. Jones DK. Diffusion MRI: theory, methods, and application. (Jones DK, editor.) Oxford ; Oxford University Press; 2011.
16. Le Bihan D, Poupon C, Amadon A, Lethimonnier F. Artifacts and pitfalls in diffusion MRI. J. Magn. Reson. Imaging 2006;24:478–488 doi: 10.1002/jmri.20683.
17. Li X, Morgan PS, Ashburner J, Smith J, Rorden C. The first step for neuroimaging data analysis: DICOM to NIfTI conversion. J. Neurosci. Methods 2016;264:47–56 doi: 10.1016/j.jneumeth.2016.03.001.
18. Veraart J, Novikov DS, Christiaens D, Ades-aron B, Sijbers J, Fieremans E. Denoising of diffusion MRI using random matrix theory. Neuroimage 2016;142:394–406 doi: 10.1016/j.neuroimage.2016.08.016.
19. Veraart J, Fieremans E, Novikov DS. Diffusion MRI noise mapping using random matrix theory. Magn. Reson. Med. 2016;76:1582–1593 doi: 10.1002/mrm.26059.
20. Cordero-Grande L, Christiaens D, Hutter J, Price AN, Hajnal J V. Complex diffusion-weighted image estimation via matrix recovery under general noise models. Neuroimage 2019;200:391–404 doi: 10.1016/j.neuroimage.2019.06.039.
21. Andersson JLR, Skare S, Ashburner J. How to correct susceptibility distortions in spin-echo echo-planar images: Application to diffusion tensor imaging. Neuroimage 2003;20:870–888 doi: 10.1016/S1053-8119(03)00336-7.
22. Smith SM, Jenkinson M, Woolrich MW, et al. Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage 2004;23:S208–S219 doi: 10.1016/j.neuroimage.2004.07.051.
23. Schilling KG, Blaber J, Hansen C, et al. Distortion correction of diffusion weighted MRI without reverse phase-encoding scans or field-maps. PLoS One 2020;15:e0236418.
24. Andersson JLR, Sotiropoulos SN. An integrated approach to correction for off-resonance effects and

subject movement in diffusion MR imaging. Neuroimage 2016;125:1063–1078 doi: 10.1016/j.neuroimage.2015.10.019.
25. Andersson JLR, Graham MS, Zsoldos E, Sotiropoulos SN. Incorporating outlier detection and replacement into a non-parametric framework for movement and distortion correction of diffusion MR images. Neuroimage 2016;141:556–572 doi: 10.1016/j.neuroimage.2016.06.058.
26. Jeurissen B, Leemans A, Sijbers J. Automated correction of improperly rotated diffusion gradient orientations in diffusion weighted MRI. Med. Image Anal. 2014;18:953–962 doi: 10.1016/j.media.2014.05.012.
27. Lauzon CB, Asman AJ, Esparza ML, et al. Simultaneous Analysis and Quality Assurance for Diffusion Tensor Imaging Alexander DC, editor. PLoS One 2013;8:e61737 doi: 10.1371/journal.pone.0061737.
28. Oguz I, Farzinfar M, Matsui J, et al. DTIPrep: Quality control of diffusion-weighted images. Front. Neuroinform. 2014;8:4 doi: 10.3389/fninf.2014.00004.
29. Bastiani M, Cottaar M, Fitzgibbon SP, et al. Automated quality control for within and between studies diffusion MRI data using a non-parametric framework for movement and distortion correction. Neuroimage 2019;184:801–812 doi: 10.1016/j.neuroimage.2018.09.073.
30. Tournier JD, Smith R, Raffelt D, et al. MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation. Neuroimage 2019;202:116137 doi: 10.1016/j.neuroimage.2019.116137.
31. Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM. FSL. Neuroimage 2012;62:782–790 doi: 10.1016/j.neuroimage.2011.09.015.
32. Tustison NJ, Cook PA, Klein A, et al. Large-scale evaluation of ANTs and FreeSurfer cortical thickness measurements. Neuroimage 2014;99:166–179 doi: 10.1016/j.neuroimage.2014.05.044.
33. Smith SM. Fast robust automated brain extraction. Hum. Brain Mapp. 2002;17:143–155 doi: 10.1002/hbm.10062.
34. Nelder JA, Mead R. A Simplex Method for Function Minimization. Comput. J. 1965;7:308–313 doi: 10.1093/comjnl/7.4.308.
35. Veraart J, Sijbers J, Sunaert S, Leemans A, Jeurissen B. Weighted linear least squares estimation of diffusion MRI parameters: Strengths, limitations, and pitfalls. Neuroimage 2013;81:335–346 doi: 10.1016/j.neuroimage.2013.05.028.
36. Basser PJ, Mattiello J, LeBihan D. MR diffusion tensor spectroscopy and imaging. Biophys. J. 1994;66:259–267 doi: 10.1016/S0006-3495(94)80775-1.
37. Westin CF, Peled S, Gudbjartsson H, Kikinis R, Jolesz FA. Geometrical Diffusion Measures for MRI from Tensor Basis Analysis. Proc. 5th Annu. Meet. ISMRM 1997:1742.
38. Tournier JD, Calamante F, Connelly A. MRtrix: Diffusion tractography in crossing fiber regions Lee J, editor. Int. J. Imaging Syst. Technol. 2012;22:53–66 doi: 10.1002/ima.22005.
39. Mori S, Wakana S, Van Zijl PCM, Nagae-Poetscher LM. MRI atlas of human white matter. Elsevier; 2005.
40. Wakana S, Caprihan A, Panzenboeck MM, et al. Reproducibility of quantitative tractography methods applied to cerebral white matter. Neuroimage 2007;36:630–644 doi: 10.1016/j.neuroimage.2007.02.049.
41. Hua K, Zhang J, Wakana S, et al. Tract probability maps in stereotaxic spaces: Analyses of white matter anatomy and tract-specific quantification. Neuroimage 2008;39:336–347 doi: 10.1016/j.neuroimage.2007.07.053.
42. Avants BB, Epstein CL, Grossman M, Gee JC. Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. Med. Image Anal. 2008;12:26–41 doi: 10.1016/j.media.2007.06.004.

43. Zhang Y, Brady M, Smith S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Trans. Med. Imaging 2001;20:45–57 doi: 10.1109/42.906424.
44. Harms MP, Somerville LH, Ances BM, et al. Extending the Human Connectome Project across ages: Imaging protocols for the Lifespan Development and Aging projects. Neuroimage 2018;183:972–984 doi: 10.1016/j.neuroimage.2018.09.060.
45. Shock NW. Normal human aging: the Baltimore longitudinal study of aging. (Shock NW (Nathan W, editor.) Baltimore, Md: U.S. Dept. of Health and Human Services, Public Health Service, National Institutes of Health, National Institute on Aging, Gerontology Research Center; 1984.
46. Resnick SM, Goldszal AF, Davatzikos C, et al. One-year age changes in MRI brain volumes in older adults. Cereb. Cortex 2000;10:464–472 doi: 10.1093/cercor/10.5.464.
47. Schilling KG, Yeh FC, Nath V, et al. A fiber coherence index for quality control of B-table orientation in diffusion MRI scans. Magn. Reson. Imaging 2019;58:82–89 doi: 10.1016/j.mri.2019.01.018.
48. Andersson JLR, Graham MS, Drobnjak I, Zhang H, Filippini N, Bastiani M. Towards a comprehensive framework for movement and distortion correction of diffusion MR images: Within volume movement. Neuroimage 2017;152:450–466 doi: 10.1016/j.neuroimage.2017.02.085.
49. Andersson JLR, Graham MS, Drobnjak I, Zhang H, Campbell J. Susceptibility-induced distortion that varies due to motion: Correction in diffusion MR without acquiring additional data. Neuroimage 2018;171:277–295 doi: 10.1016/j.neuroimage.2017.12.040.