Diffusion Tensor Imaging (DTI) White Paper (07-06-2012)

Release 5.1

Part 1 - Data

Release 5.0 is the first to contain DTI data. The low resolution (3mm) DTI data acquired with the “conventional” DTI protocol is included. Data collected with the high resolution (2.5mm) acquisition protocol are now included with Release 5.1. Datasets were assessed for adherence to protocol and for image quality. The DTI data in this release were included only if they met set quality criteria (498 scans from 274 unique subjects).

These are the types of data that are included for each dataset:

Tensor Derived Quantities. Provided as individual NIFTI format images, stored in the radiological convention (right side of brain on left of image), up to 9 files are included:

- Directionally encoded color maps (DEC) [1]
- Eigenvalues (EV) as a 4D image file with volumes in the order of λ1, λ2, λ3
- Fractional anisotropy (FA)
- Lattice index (LI)
- Relative anisotropy (RA)
- Trace of the diffusion tensor (TR) (equal to 3*Mean Diffusivity)
- Chi-Squared map of the fitting (CS). A measure of the goodness of fit of the tensor model.
- Outlier map (OUT). From RESTORE[2] or iRESTORE[3] fitting only. Indicates the percentage of data points identified as outliers and removed from the tensor fitting on a voxel-by-voxel basis.
- Brain mask (MS)

The computed tensor derived quantities are the files that most end-users would be interested in downloading. Several freely available software packages can open these nifti formatted images including (but not limited to) MIPAV (http://mipav.cit.nih.gov), FSLView, part of the FSL package (http://www.fmrib.ox.ac.uk/fsl/), MRicro, (http://www.mccauslandcenter.sc.edu/mricro/mricro/) and MRCron (http://www.mccauslandcenter.sc.edu/mricro/mrcron/).

If investigators are interested in doing additional processing or analyses, three TAR files are also included in the release:

1) Raw uncorrected diffusion weighted images (DWIs) (_RAW_UNCORRECTED.tar.gz). The DWI images in this directory are as they were acquired on the scanner, with no distortion or artifact corrections applied. The data has been re-organized to be as consistent as possible with the protocol for the study. This means that if extra data were acquired in order to compensate for bad data (for example if it was noticed at the scanner that the subject moved), then the bad images have been removed, and only the good ones are included here. The b-matrix files have been edited to account for any deviations from the protocol in the acquired data, i.e., to compensate for data being acquired out of order, incomplete, or
extra data. In addition, for Siemens data acquired with an oblique orientation, the b-matrix has been rotated to correctly compensate for the rotation of the gradient frame of reference that is applied by the scanner in this case. This correction is not necessary for GE data, as the scanner compensates automatically for oblique acquisitions.

The tarred directory contains the following 4 elements:

- Slice by slice images in raw floating point format (_RAW_UNCORRECTED directory)
- A text readable header file (.list) which describes the essential information about the images including resolution, slice thickness, etc.
- A text readable B-matrix file (.bmtxt) which contains 6 columns by n-rows, where n is the number of image volumes contained in the dataset (b0 images and DWI images). The 6 columns are (from left to right) \( b_{xz} \) \( 2*b_{xy} \) \( 2*b_{xz} \) \( b_{yz} \) \( 2*b_{yz} \) \( b_{zz} \)
- A text readable file of pointers to the raw data (.path) which contains the relative path to each slice in the order of 1st slice for all volumes, then 2nd slice for all volumes, ... nth slice for all volumes

The .list, .path and .bmtxt format is used in TORTOISE[4], which is a software package developed at the Eunice Kennedy Shriver National Institute of Child and Health and Human Development (NICHD) under the auspices of the Pediatric MRI project for the processing of diffusion tensor imaging data. TORTOISE is freely downloadable at: www.tortoisediti.org. TORTOISE has a number of export options for the end user who wishes to work in other software packages. See the brief tutorial at the end of this document for export instructions.

2) **Raw corrected DWIs plus tensor output (_RAW_CORRECTED_AND_TENSOR.tar.gz)**. The DWI images in this directory have undergone the standard TORTOISE diffusion tensor imaging data correction pipeline, consisting of the following steps (in this order):

i) Computation of transformation for motion and eddy distortion correction [5]

ii) Computation of transformation for susceptibility induced EPI distortion correction

A first step in this correction consists in registering with a second order model the first b0 image to the T2 weighted structural image described below at iii) in native DWI space. This step was applied to all datasets.

A second step in this correction involves using the b-spline image registration method described by Wu et al [6]. This method has different levels of success depending on subject age and amount of distortion. Due to this fact, this type of EPI correction was applied only to data with severe distortion. Generally, this was applied only on GE data for children above the age of 5 years. In all other data, b-spline correction was applied on a case-by-case basis and assessed for quality of improvement post correction. Use of this type of EPI distortion correction is indicated in the DTI metadata in the database under the tag name: mri_epi_correction_dti.

iii) Computation of transformation for reorientation to a common anatomical orientation,
This step was accomplished by registering the first non-diffusion weighted image of the dti dataset to a properly reoriented T2 weighted structural image of the same subject. (The image was produced by applying the transformation used to produce the scaled image in Talairach space to the native space T2 weighted structural image, but with the scaling terms set to zero. This image, which is called tal noscale, was provided by the data coordinating center which created it from structural images, from the Pediatric MRI project data release 3.) However, these structural target images with the same anatomical orientation as the Talairach images, without the warping into standard space, were not included in the structural MRI data releases. Thus, the DTI data does not match either the native space or tal space images provided in the structural data releases, but instead matches the tal noscale images which are now being released for the first time with the corresponding DTI data.

Additional structural images were used for DTI registration in cases where the structural scans failed to pass all stages of QC required for inclusion in the structural MRI portion of the database. In these cases, the T2W image was assessed for use in the DTI pipeline. If a T2W image was deemed un-usable or was otherwise unavailable, a T1W image was used in its place. In these cases, no b-spline EPI correction is used (step ii above). In these cases, the structural images were provided without defacing. As such, we performed a manual defacing before image registration was performed in TORTOISE.

iv) Application of all transformations, rotation of b-matrices [5, 7] and production of the “raw corrected” images. All transformations are combined and applied as a single transformation in order to have only a single interpolation.

For these “raw corrected” DWI images, noise variance is estimated and images of a variance modulating factor[8-9] are produced. In the data release, we provide transformation matrices (See [http://science.nichd.nih.gov/confluence/display/nihpd/Transformation+Files](http://science.nichd.nih.gov/confluence/display/nihpd/Transformation+Files) for full description of the transformation matrices), noise variance images, and the T2W structural images used as targets.

After corrections, the data is fit with either conventional nonlinear least squares tensor fitting, RESTORE robust tensor fitting[2], or iRESTORE robust tensor fitting[3]. The tensor, amplitude image and mask image are provided. Additional tensor derived metrics can be quickly computed from these three items using TORTOISE. See the brief tutorial on TORTOISE provided below. Additional detailed documentation is available on the TORTOISE website.

The tarred directory contains the following 11 elements:

- Slice-by-slice raw corrected DWI images in floating point format (_RAW_CORRECTED directory)
- .list
- .bmtxt
- .path
- Slice by slice images of the computed noise variance information (_noise_info directory)
- Path to the noise variance information images (_noise.path)
- Transformation matrix from native DWI space to the final reorientation to the tal noscale images (.transformations)
- Transformation matrix including initial registration to T2 weighted image, and motion and eddy corrections, (_up_rpd.transformations)
- Deformation field output for EPI correction (_deformation_field_output directory). This only exists if EPI correction was applied. If present, it contains 3 deformation field images (x, y and z deformations respectively) and two b0 images. Two of the deformation field images contain only zeros, while the third contains the deformations constrained to the Phase Encode direction (AP for all subjects in this study). The b0 images are the first b0 image from the dataset before (original_upsampled_b0.nii) and after EPI correction (warped_upsampled_b0.nii).
- T2 weighted structural image used as target for EPI distortion correction and/or reorientation template (_structural.nii)
- Computed tensor (_SAVE). This contains the amplitude image (_AM.nii), diffusion tensor (_DT.nii) and several mask images (_M*.nii), and a parameters file used by TORTOISE (.parms).
  Additional tensor derived metrics can be quickly computed from these items using TORTOISE.

3) **Tensor output** (_TENSOR_ONLY). The tarred directory contains the following 3 elements:

- .list
- .bmtxt
- Computed tensor (_SAVE). This contains the amplitude image (_AM.nii), diffusion tensor (_DT.nii) and several mask images (_M*.nii), and a parameters file used by TORTOISE (.parms).
  Additional tensor derived metrics can be quickly computed from these items using TORTOISE.

### Part 2 – Living Phantom

The living phantom data follows the same data format as described above. A detailed description of and assessment of the DTI phantom data, including both ACR and Living Phantom, can be found in Walker, et al. [10].

The living phantom data followed the Objective 1 DTI protocol (mainly 6 diffusion directions at b=1000s/mm² plus 1 b=0s/mm², repeated 4 times, totaling 28 brain volumes). However, four living phantom datasets also included 6 b=500s/mm² plus 1 b=2s/mm² images, repeated twice (42 brain volumes). These datasets have been included both with and without the b=500s/mm² images, with the following naming convention, where # refers to the data specific site number and ###### refers to the data specific identifier:

- Consistent with Objective 1 protocol: LIVING_PHANTOM_1_SITE_#_DTI_#####_REL02...
- Including b=500s/mm² images: LIVING_PHANTOM_1_SITE_#_full_DTI_#####_REL02...

### Part 3 – Quality Assessment
All DTI data included in the database underwent rigorous quality assessment for both adherence to protocol and for image quality. The different assessment categories are as follows.

**Deviations from Protocol.** Scoring: 0 – according to protocol, 1 – not according to protocol

- number of volumes, \( b_0 \) should = 4 (objective 1) or 6 (objective 2), \( b_{1000} \) should = 24 (objective 1 & 2), \( b_{500} \) should = 12 (objective 2)
- number of slices, between 48 and 60
- resolution, ~3mm isotropic
  - Some tolerance on resolution was allowed, allowable range was 2.97 - 3.13mm
- slice thickness, 3mm
- zero filling, no zero filling/zero padding at the scanner is expected
- averaging / NEX, no averaging is expected
- slice orientation, axial slice orientation is expected, oblique slices are considered not according to protocol
- echo time consistency, if the echo time (TE) is different between \( b=500 \) series and \( b=1000 \) series, the \( b=500 \) images were removed.
- scaling difference between image series – series with a scaling difference were identified and a score of 1 was indicated. Offending series volumes were removed.

**Data Quality.** Scoring: 0 – no issue, 1 – minor, 2 – moderate, 3 – severe

-a) Raw uncorrected data, before processing with the TORTOISE pipeline

- brain coverage, top of the brain
- brain coverage, bottom of the brain
- ghosting, severity
- artifacts affecting signal, such as spike noise, RF artifacts, or reconstruction artifacts
- motion, resulting in signal dropouts
- motion, occurrence of within volume (interleave) mis-registration, and severity of eddy current distortions
- susceptibility induced (geometric) EPI distortion, severity
- cardiac pulsation, pervasiveness

-b) Raw corrected data and tensor derived quantities, after processing with the TORTOISE pipeline

- quality of motion and eddy distortion correction
- quality of susceptibility induced EPI distortion correction
- visual assessment, performed for overall and regional quality of tensor derived quantities
  - frontal
  - parietal
  - occipital
  - temporal
Postprocessing to correct for artifacts remaining after correction pipeline.

As mentioned above, either conventional non-linear least-squares tensor fitting, RESTORE robust tensor fitting or iRESTORE robust tensor fitting were used on the data. The type of fitting selected was based on the quality assessment steps above. When artifacts resulting in signal increases and decreases were both seen in a single dataset, the RESTORE fitting was used. When artifacts resulting in only signal decreases were found, iRESTORE fitting was used. If insufficient data redundancy existed for either robust fitting method, or no artifacts were present, than non-linear fitting was used.

The visual assessment scores represent the scores of the best fitting method for each dataset.

Best efforts were made to use post-processing techniques to correct for all artifacts and distortions encountered in the DTI data. However, datasets were rejected in order to maintain a minimum level of quality in the database. A breakdown of included/rejected data is provided below.

Total of acquired scans: 883

Total of scans included in database: 498 (274 unique subjects)

Total Rejected: 385

i) For deviations from protocol: 169

ii) For uncorrectable artifacts/distortions: 171

iii) For deviations from protocol combined with artifacts/distortions: 45

Common reasons for rejection are as follows.

i) Deviations from protocol:

- Improper combination of Field of View (FOV) and acquisition matrix size. Occasionally, the larger recommended FOV (128x128mm) was used in combination with the smaller recommended matrix size (64x64). This resulted in 6mm isotropic voxels. This resolution was considered too low to meet minimum quality standard.

- Acquisition errors, such as improper maximum b-value, and applied gradient directions not according to protocol and not properly documented meant that a correct gradient table could not be determined, and data were rejected.
• Averaging (NEX>1) was not expected in the protocol, however, it was occasionally used. While this does not necessarily make the data un-usable, it is impossible to detect and correct many artifacts and distortions that exist in the data. Greater than 2 averages (i.e. if NEX > 2), the data were rejected. If NEX=2, data were assessed on a case by case basis (see section iii below).

• Data that was acquired from subjects with metal in their bodies, i.e. braces, were excluded due to major artifacts.

ii) Uncorrectable artifacts/distortions:

• Severe ghosting

• Signal increase or dropout artifacts which were pervasive. A judgment was made as to whether enough data remained after removal of artifacts. When too many volumes/slices were affected, data were rejected.

• Certain datasets contained severe eddy current distortions such that the distortions remained at a severe level after correction. These datasets were rejected.

iii) Deviations from protocol combined with artifacts/distortions:

• Data acquired with less than the protocol required number of volumes (i.e. 24 DWIs plus 4 b0 images) were generally included if the data quality were adequate. However, when artifacts were present, there was insufficient data to allow proper correction or removal of bad volumes, and data were rejected.

• In cases where NEX=2, if the data quality were poor, this generally indicated that there was potential motion between series acquisitions. Due to the averaging at the scanner, this motion was uncorrectable, and data were rejected.

Part 4 - Tutorial for installing and using the data release in TORTOISE

TORTOISE can be downloaded from www.tortoisediti.org. The software runs on Mac and Linux. The software requires either an IDL virtual machine (similar to the way Java runs on most computer systems) or an IDL license. TORTOISE can be downloaded with a pre-compiled version of the IDL virtual machine, or if desired, the end user can choose to install IDL by downloading it from http://www.ittvis.com/ProductServices/IDL.aspx. At the date when this document was written, the current version of TORTOISE is v1.2.2, and was created under IDL version 8.1.

Step 0: Download DTI data from the data release in the Pediatric MRI data repository, via NDAR. If you only wish to download the computed NIFTI quantities, then you do not need to download TORTOISE. Simply download the desired data, and use one of the many freely available software packages to view your NIFTI images.
If you would like to view the tensor itself, and/or export the tensor to other software packages, for example, to perform tractography, then you need to download the additional compressed files described above (Raw Uncorrected, Raw Corrected with Tensor, or Tensor Only). You must uncompress them by using the following command at a command line (but substituting your filename in the example below):

```bash
tar -xzvf example_REL01.tar.gz
```

Or by right clicking on the file and using an archive manager to extract the files.

**Step 1: Register and download the software.** Go to [www.tortoisedti.org](http://www.tortoisedti.org). You will see the download link in the menu on the top left hand side of the page.

This will take you to a page where you can register as a new user:
Fill out the required information, and read and accept the license agreement and submit. A message will appear as shown below that indicates an email will be sent to your email address.

The email will contain your username and password, and links to the download page and the support email address.

You can use your username and password to log in. Then you will see the list of available downloads as shown below. Please download the newest file, version 1.2.2.
To download the file, click “Accept Agreement” and accept the license agreement. It will reload the page, and it will now have a “Download” option. Click on Download, and then select “save file”. The file you have downloaded will be a “.tar.gz” compressed file.

**Step 2: Uncompress the software.** Either open the file with an archive manager to extract the files, or, at the command line in a terminal, navigate to the location of the downloaded file, and type the following command.

```
tar -xzvf filename.tar.gz
```

Once it completes the file extraction you are ready to use the software.

**Step 3: Run DIFF_CALC to load a dataset.** TORTOISE is composed of two modules, DIFF_CALC for tensor estimation and visualization, and DIFF_PREP for preprocessing and registration steps. For more detailed instructions on the two modules, see the software guides available on [www.tortoisedti.org](http://www.tortoisedti.org). For simplicity we provide instructions only in loading, exporting and visualizing the datasets with DIFF_CALC using the IDL virtual machine in this brief tutorial.

Running the software is quite simple.
- Navigate to the directory that has your newly uncompressed files.
- Go into DIFF_CALC/diffcalc_main
- In diffcalc_main you should see an executable with the name calcvm. If you work from command line in a terminal, type calcvm and hit enter. If you work from a graphical interface, then double click on calcvm.
- A splash screen will appear. Click on it and the software will load.
- The first time you run the software, you may get some messages indicating that working directories have been created for you. Simply click OK and the software will continue normally. This should only happen the first time running.

The software should look something like this:

Click on load listfile.
In the dialogue that pops up, navigate to the downloaded and uncompressed data (see Step 0). Select the file ending in .list, and click OK. More buttons have now sensitized, and the software looks like this. If you downloaded a version including the computed tensor, then click restore session to load or computed the tensor derived quantities, and then follow the directions below. If you downloaded raw diffusion weighted images, please follow the tutorial for tensor fitting with TORTOISE at https://science.nichd.nih.gov/confluence/display/nihpd/diff_calc+fit.

After the session is restored, click on compute eigenvalues.
Then click derive variables.

To view the tensor derived quantities, click on roi utilities.
This will bring up a single slice viewer, and a couple of menus. All details are described in the online documentation. In general, use the slider on the ROI FUNCTION menu to look at all of the slices of the brain volume. Select the quantity you would like to view on the Image Type menu.

For a three plane viewer, select triplanar (the button above roi utilities in DIFF_CALC). This will display a sagittal, coronal and axial view of most computed quantities.
Step 4: Export to other formats. Click on export images. This will provide a list of export options.

Most commonly used export options are RAW FSL-LIKE UNSORTED which will create a 4D analyze file (readable by FSL) along with a mask image and the bvals and bvecs files required for use in FSL. This is useful if the end user wishes to do probabilistic tractography. See hints below if you wish to do a TBSS style analysis. And, for deterministic tractography, PROCESSED FOR TRACKVIS, which provides output for the software package TrackVis (http://www.trackvis.org/).

IMPORTANT HINTS:

1) ALWAYS use the “DONE” or “quit” buttons instead of clicking the X’s in the top left or right hand corner of the windows. Using the X can create instabilities with the software. If this happens, simply quit and start DIFF_CALC again (step 3).

2) If you wish to do any analyses on the data use the CORRECTED images. The RAW UNCORRECTED are provided in case the end user wishes to apply their own correction pipeline to the raw data. It is not recommended to do any analysis on uncorrected data.

3) For TBSS style analysis, we highly recommend using the FA NIFTI images provided in the database. These were created from the corrected DWI images, with appropriate b-matrix reorientations, with a correction factor for noise modulation due to interpolation effects, and using a non-linear least squares method of tensor estimation which is a better model than the linear fit traditionally used in many software packages. For TBSS analysis, skip step 0 (http://www.fmrib.ox.ac.uk/fsl/tbss/index.html) and simply start with step 1 (tbss_1_preproc).
4) If you run into any error messages, please check the FAQ for common issues.  
http://science.nichd.nih.gov/confluence/display/nihpd/FAQ. If your question does not appear in the list, contact TORTOISE support at tortoisedti@gmail.com.

References


