**GOGNESTIC 2023: Diffusion Tensor Imaging Analysis with FSL**

This document is an edited version of the FSL Diffusion Toolbox Practical to match what we will cover during GOGNESTIC. The full official FSL course tutorials, as well as the FSL FDT user’s guide, are available here:

<https://fsl.fmrib.ox.ac.uk/fslcourse/2019_Beijing/lectures/FDT/fdt1.html>

<https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FDT/UserGuide>

**FSL Diffusion Toolbox**

In this session we will walk through the steps needed to prepare your diffusion data for analysis. We will also cover diffusion tensor model fitting and group analysis of DTI data using tract-based-spatial-statistics (TBSS).

**Download the data and set up**

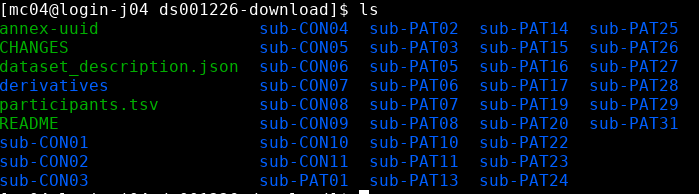
First of all, you need to install FSL if you haven’t done so already. For instructions please see:

<https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FslInstallation>

For this tutorial we will use the Brain Tumor Connectomics Data which is available from OpenNeuro:

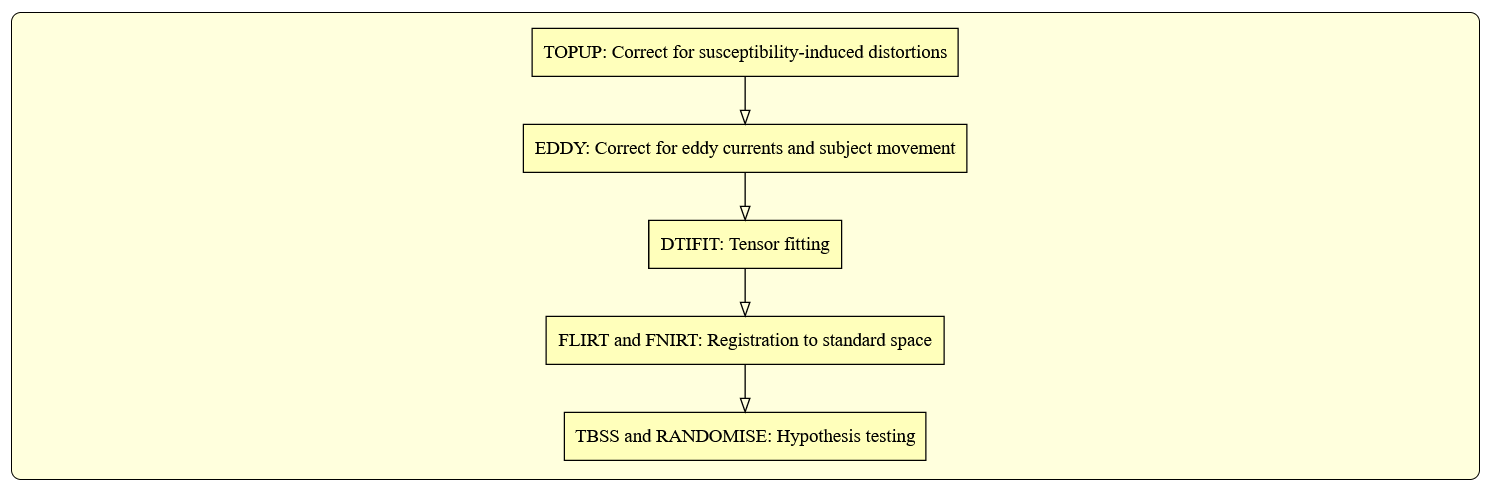
<https://openneuro.org/datasets/ds001226/versions/00001>

Once you have downloaded the data, the directory including all of the individual subject datasets will be your working directory. In my case this is called ‘ds0011226\_download’ and this is what it contains:



All of the steps described below have been implemented in the scripts **FDT\_DTI\_pipeline.sh** and **FDT\_DTI\_TBSS.sh** which you can download from <https://imaging.mrc-cbu.cam.ac.uk/methods/COGNESTIC2022> .

**The FDT DTI Pipeline**

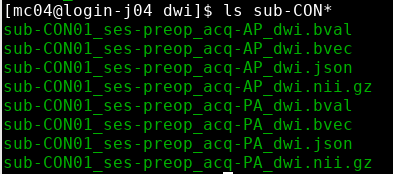


**Diffusion Data**

In this first section of the practical we will familiarise ourselves with diffusion data. If you are comfortable with working with diffusion data, feel free to run through this section quickly or [skip to TOPUP.](https://fsl.fmrib.ox.ac.uk/fslcourse/2019_Beijing/lectures/FDT/fdt1.html#topup)

cd /Your/data/folder/ds001226-download/sub-CON01/ses-preop/dwi

List this directory and you should see a set of files that are obtained from a typical diffusion MRI acquisition. This includes a nii.gz dwi data file, as well as .bval and .bvec files that contains the information on the diffusion-sensitising magnetic field gradients:



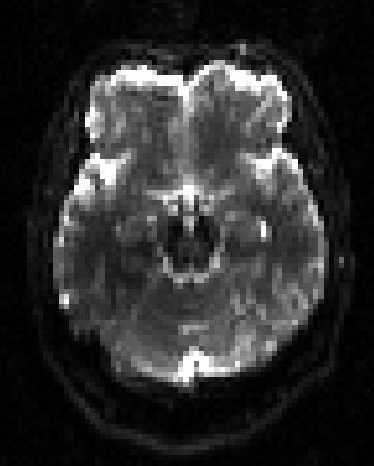
The bval file contains a scalar value for each applied gradient, corresponding to the respective b-value. bvec contains a 3x1 vector for each gradient, indicating the gradient direction. The entries in bval and bvec are as many as the number of volumes in the nii.gz file. So the ith volume in the data corresponds to a measurement obtained after applying a diffusion-sensitising gradient with a b-value given by the ith entry in bval and a gradient direction given by the ith vector in bvec. You can quickly see the contents of these two files, by typing:





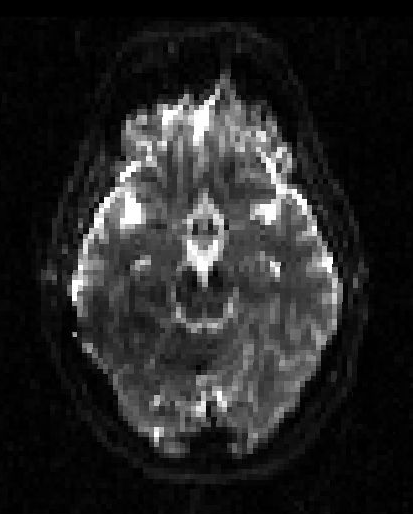
Bring up fsleyes and open sub-CON01\_ses-preop\_acq-AP\_dwi.nii.gz - you will need to reset the maximum display range to around 1000. This is the diffusion data before correction for distortions and it is a 4D image. Turn on movie mode (https://fsl.fmrib.ox.ac.uk/fslcourse/2019_Beijing/lectures/FDT/FSL%20Diffusion%20Toolbox%20Practical_files/movie_icon.png) to see the effect on the signal of the different b-values and gradient directions. Turn off the movie and navigate through the different slice to notice the susceptibility-induced distortions at the frontal part of inferior slices.

Susceptibility-induced distortions



Select *File > Add from file* and select sub-CON01\_ses-preop\_acq-PA\_dwi.nii.gz – this is the data acquired with the reversed phase encode direction (P >A). Notice how the susceptibility induced distortions now go the opposite way.

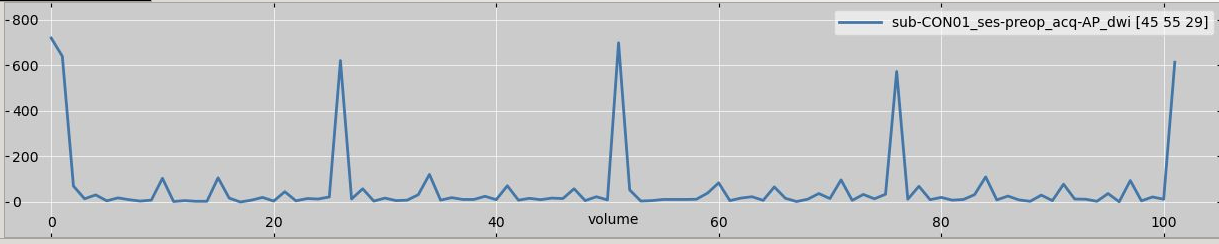
Susceptibility-induced distortions



Bottom of Form

Now, let's have a look at all the data associated with a voxel. Choose a CSF voxel (e.g. [60, 63, 39]) and observe how the signal changes with gradient direction. To do that choose from the menu *View > Time series*. A new window will appear with a plot of the signal intensity at the chosen location for the different diffusion-weighted volumes. Notice the few high intensity values and the very low intensities in most of the data points. The former correspond to the b=0 images. The latter to diffusion-weighted images, for which maximal attenuation of the CSF signal has occurred.

Notice that the higher the b-value, the lower the signal. Can you identify time points that correspond to the three shells used in this acquisition? Check you are right by checking the bval file.



Bottom of Form

Look again at the images. For volumes with diffusion gradients applied, see if you can work out the direction of the gradient. Remember that diffusion data appear darker in places where there is more diffusion along the gradient direction. You should be able to work out the gradient direction by looking at a diffusion weighted image, and comparing darker areas with your knowledge of white matter orientation.

**TOPUP - Correcting for Susceptibility-induced Distortions**

FSL wiki: <https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/topup>

A minimum requirement for using topup for correcting distortions is that two spin-echo (SE) EPI images with different PE-directions have been acquired. The best is to acquire them with opposing PE-directions (i.e. A→P AND P→A or L→R AND R→L). An SE-EPI image is the same as a b=0 image acquired as a part of a diffusion protocol. Just as for fieldmaps this pair must be acquired at the same occasion as the full diffusion protocol and the subject must not leave the scanner in between and no re-shimming can be done.

**Identifying susceptibility-induced distortions**

From the nii.gz files choose a volume without diffusion weighting (e.g. the first volume). You can now extract this as a standalone 3D image, using fslroi. Call the extracted files nodif\_AP/PA.

fslroi sub-CON01\_ses-preop\_acq-AP\_dwi.nii.gz nodif\_AP 0 1

fslroi sub-CON01\_ses-preop\_acq-PA\_dwi.nii.gz nodif\_PA 0 1

Open nodif\_AP in FSLeyes and have a look at different slices. Notice that as you go to more inferior slices, the frontal part of the brain starts to appear distorted (e.g. "squashed" or "elongated"). These distortions are always present in EPI images and are caused by differences in the magnetic susceptibilities of the areas being imaged.

Bottom of Form

Now open and superimpose in FSLeyes the image nodif\_PA. This is an image without diffusion-weighting (i.e. b=0) of the same subject that has been acquired with the opposite PE direction (Posterior→Anterior). Switch on and off the visibility of this image to see how the distortions change sign between nodif\_AP and nodif\_PA. Regions that are squashed in the first appear elongated in the second and vice versa. Unsurprisingly the areas that were very obviously distorted when viewed in the nodif\_AP image changes a lot as you switch back and forth between nodif\_AP and nodif\_PA. We will correct these distortions by combining the two b=0 images using the TOPUP tool. We will then pass the results on to the EDDY tool where it will be applied to the correction of all diffusion data.

**Running topup**

First you need to merge the AP and PA images into a single image using fslmerge. Merge the files along the 4th 'timeseries' axis. Call the merged image AP\_PA\_b0.

fslmerge -t AP\_PA\_b0 nodif\_AP nodif\_PA

Then create a text file that contains the information with the PE direction, the sign of the AP and PA volumes and some timing information obtained by the acquisition. This is a text-file called acqparams.txt that contains the lines:

0 -1 0 0.0266003

0 1 0 0.0266003

The first three elements of each line comprise a vector that specifies the direction of the phase encoding. The non-zero number in the second column means that is along the *y*-direction. A -1 means that *k*-space was traversed Anterior→Posterior and a 1 that it was traversed Posterior→Anterior. The final column specifies the "total readout time", which is the time (in seconds) between the collection of the centre of the first echo and the centre of the last echo. In the FAQ section of the online help for topup there are instructions for how to find this information for Siemens scanners.

The file should contain as many entries as there are volumes in the image file that is passed to topup.

We are now ready to run topup which we would do with the command:

topup --imain=AP\_PA\_b0 \

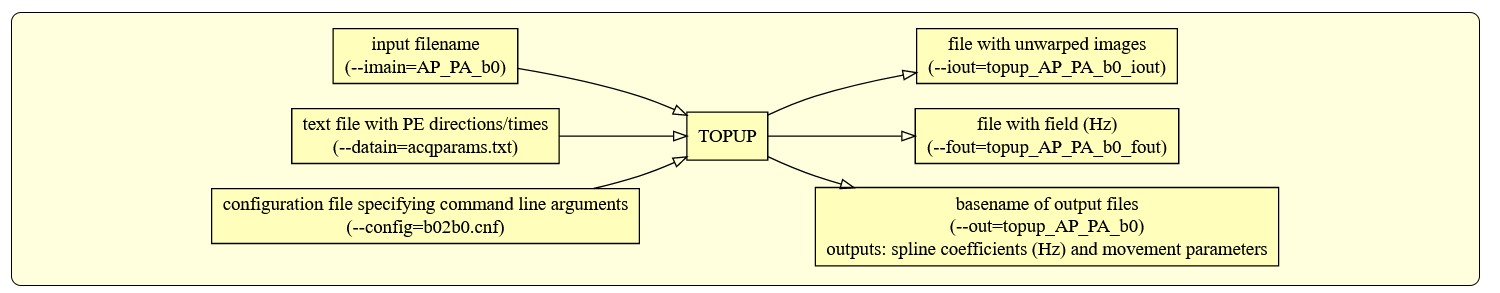
--datain=acqparams.txt \

--config=b02b0.cnf \

--out=topup\_AP\_PA\_b0 \

--iout=topup\_AP\_PA\_b0\_iout \

--fout=topup\_AP\_PA\_b0\_fout



This command will take a long time to run, depending on the size of the dataset.

There are four result files. topup\_AP\_PA\_b0\_fieldcoef.nii.gz contains information about the off-resonance field and topup\_AP\_PA\_b0\_movpar.txt specifies any movement between nodif\_AP and nodif\_PA. Open FSLeyes and load topup\_AP\_PA\_b0\_fieldcoef.nii.gz. It looks like a low resolution fieldmap, and it contains the spline coefficients for the field that TOPUP has estimated. Close FSLeyes and re-open it, but this time take a look at the actual field (topup\_AP\_PA\_b0\_fout). Moreover, to check that TOPUP has done its job properly, load topup\_AP\_PA\_b0\_iout and compare its two volumes to those we provided as input (AP\_PA\_b0.nii.gz).

**EDDY - Correcting for Eddy Currents**

FSL wiki: <https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/eddy>

We will first generate a brain mask using the corrected b0. We compute the average image of the corrected b0 volumes using fslmaths and calling the output file hifi\_nodif.

fslmaths topup\_AP\_PA\_b0\_iout -Tmean hifi\_nodif

And then we use BET on the averaged b0. Again figure out the command and call the output file hifi\_nodif\_brain. Hint: create a binary brain mask, with a fraction intensity threshold of 0.2.

bet hifi\_nodif hifi\_nodif\_brain -m -f 0.2

You can then use FSLeyes to ensure that bet has done a good job.

**Running EDDY**

Now we would typically run eddy using the command:

eddy --imain=sub-CON01\_ses-preop\_acq-PA\_dwi.nii.gz \

--mask=hifi\_nodif\_brain\_mask \

--index=index.txt \

--acqp=acqparams.txt \

--bvecs=bvecs \

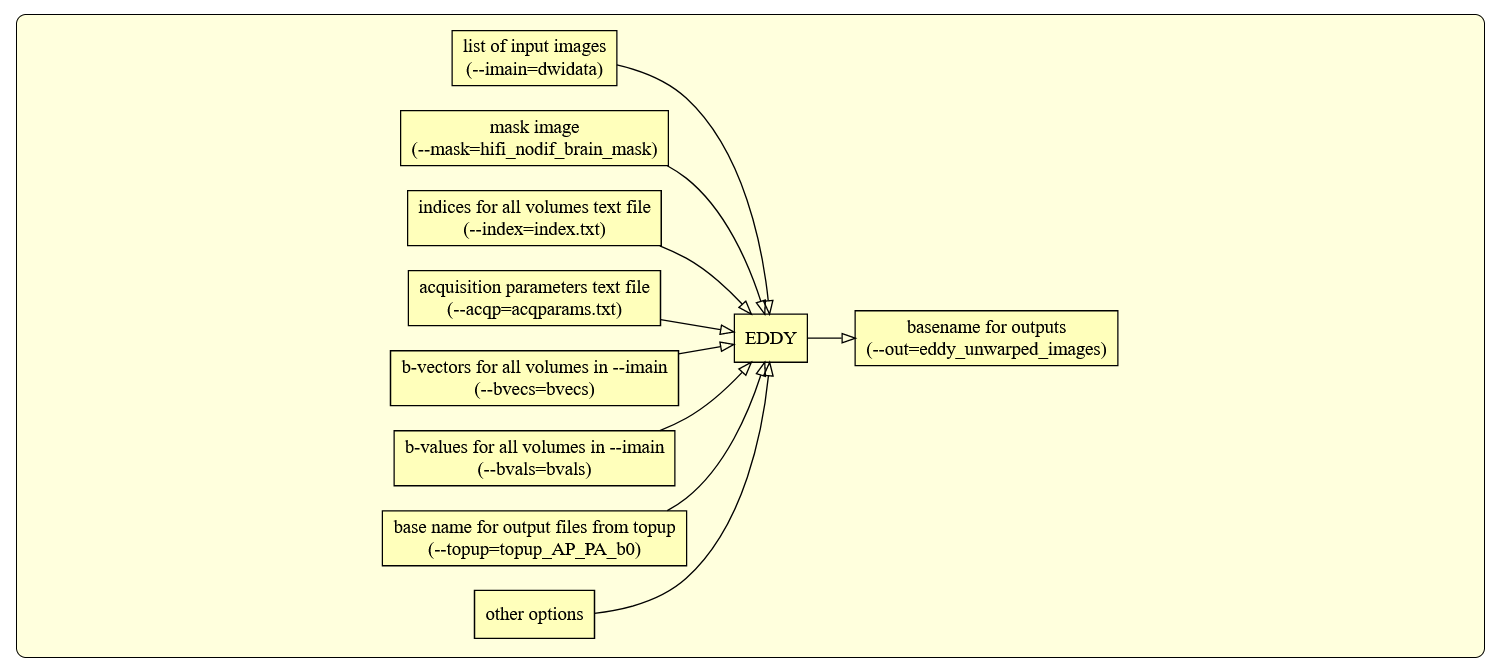
--bvals=bvals \

--fwhm=0 \

--topup=topup\_AP\_PA\_b0 \

--flm=quadratic \

--out=eddy\_unwarped\_images



In this command:

* sub-CON01\_ses-preop\_acq-PA\_dwi.nii.gz is the full diffusion data (including *b*=0 volumes), all acquired with a Anterior→Posterior phase-encoding.
* The text-file index.txt contains a column of ones, one for each volume in dwidata, specifying that all volume were acquired with the parameters specified by the first row in acqparams.txt.
* topup\_AP\_PA\_b0 was the name given as the --out parameter when we ran topup and will cause eddy to look for the files topup\_AP\_PA\_b0\_fieldcoef.nii.gz and topup\_AP\_PA\_b0\_movpar.txt.
* The parameters --fwhm=0 and --flm=quadratic specify that no smoothing should be applied to the data and that we assume a quadratic model for the EC-fields. These are our current recommendations and you are unlikely to ever have to use any other settings.

To create the index.txt file for a dataset with 102 volumes (such as the one we are using in this tutorial) you can use the following lines within a bash environment:

indx=""

for ((i=1; i<=102; i+=1));

do indx="$indx 1";

done

echo $indx > index.txt

To find out the number of volumes in your dataset you can use fslinfo data.nii.

As eddy performs a simultaneous registration of all volumes in the data set it is quite memory and CPU hungry. Therefore this step will take a very long time.

You will find that eddy has produced three output files: eddy\_unwarped\_images.nii.gz, eddy\_unwarped\_images.rotated\_bvecs and eddy\_unwarped\_images.eddy\_parameters. The former of those is the "main result" and contains the data in sub-CON01\_ses-preop\_acq-PA\_dwi.nii.gz corrected for susceptibility, eddy currents and subject movements.

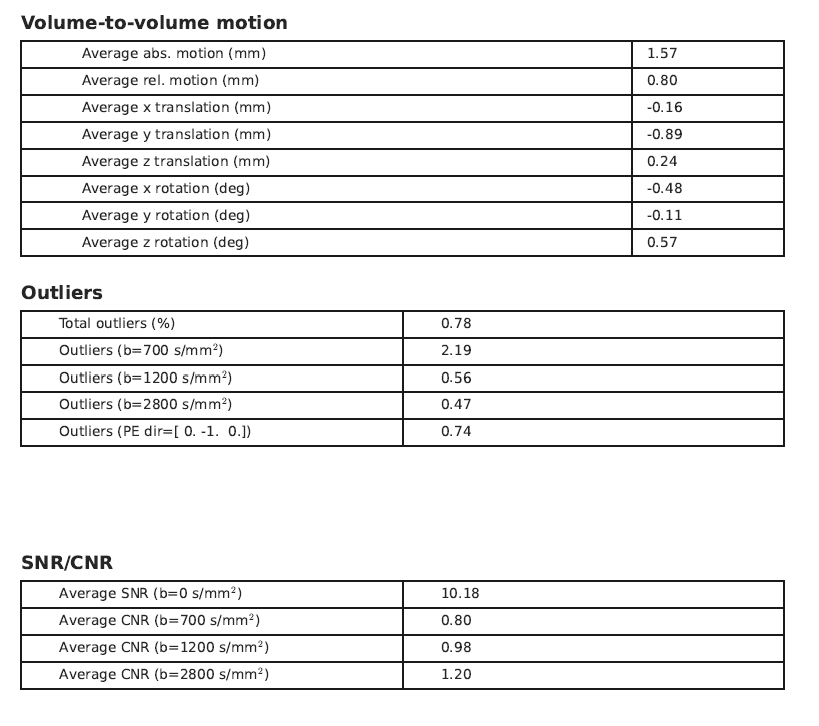
**Quality control: EDDY QC**

As we have seen, dMRI data can be affected by many hardware or subject-specific artefacts. If undetected, these artefacts can bias downstream analysis. Quality control is therefore very important - always look at your data! In large population studies, manual quality control may not be practical.

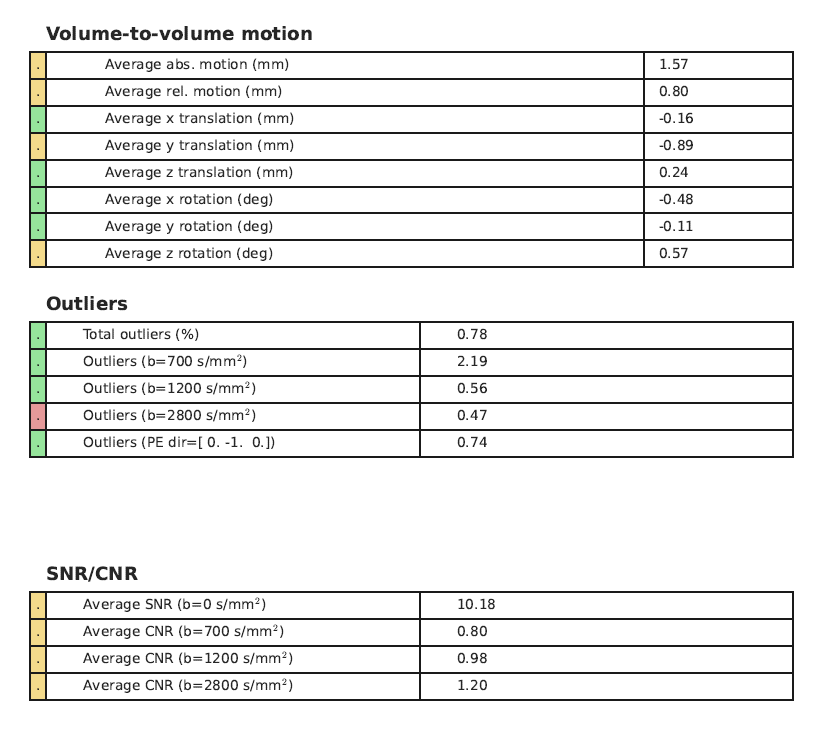
A new FSL tool eddy\_quad provides automatic quality control at both the single subject and group level. For more info see

FSL wiki: <https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/eddyqc> .

For each subject, EDDY QC will generate a pdf file with a data quality report, including information about subject movement, outliers, signal-to-noise ratio, amongst other metrics.



Quality control can also be performed at the group level, using the eddy\_squad command. This command reads all the single subject outputs from eddy\_quad, generates study-wise reports and can optionally update single subject reports, indicating how the subject’s dataset compares to other data, using a ‘traffic light’ system (see below and example of an updated QC report). Lastly, eddy\_squad also allows to report QC indices based on user-provided variables.



**DTIFIT - fitting diffusion tensors**

The data we are working with comprises of four b-values per dataset: b=0, b=700, b=1200 and b=2800 s/mm2. The diffusion tensor model is not accurate for b-values greater than approximately 1300 s/mm2 (see Jones et al, 2004, <https://doi.org/10.1002/mrm.20283>) and therefore for the next steps we will include only the data acquired with b=0, 700 and 1200 s/mm2.

select\_dwi\_vols is the tool to extract volumes with specific b-values from a 4D diffusion-weighted dataset. For example:

select\_dwi\_vols \

eddy\_unwarped\_images.nii.gz \

sub-CON01\_ses-preop\_acq-AP\_dwi.bval \

data\_for\_dtifit \

0 -b 700 –b 1200 \

-obv eddy\_unwarped\_images.eddy\_rotated\_bvecs

This command will create a new 4D file containg only those volumes with b-values ~=0 and ~=700 and ~=1500. It will also generate two new bval and bvec files containing only the selected b-values and b-vectors.

You can run DTIFIT (the FSL diffusion tensor fitting program) by calling dtifit from the command line:

dtifit --data=data\_for\_dtifit.nii.gz \

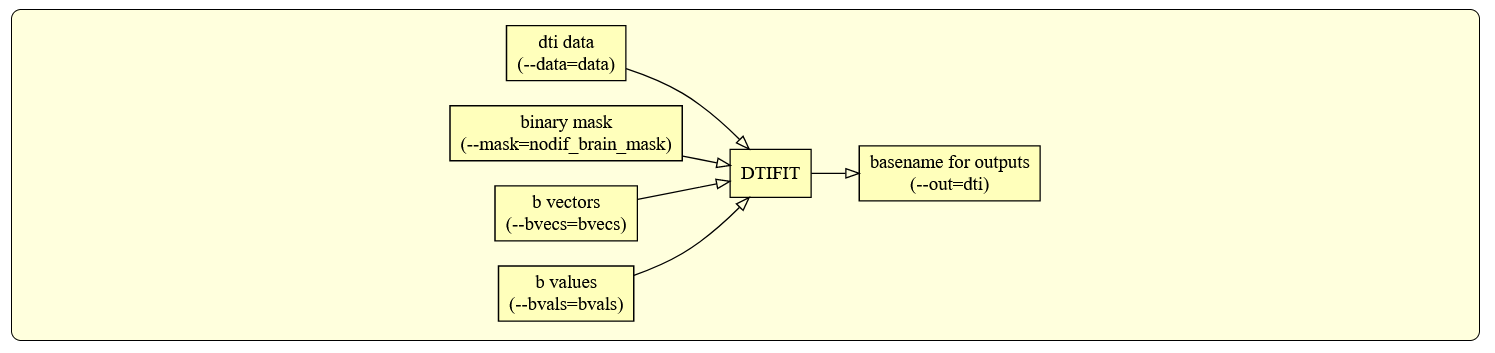
--mask=nodif\_brain\_mask \

--bvecs=data\_for\_dtifit.bvec \

--bvals=data\_for\_dtifit.bval \

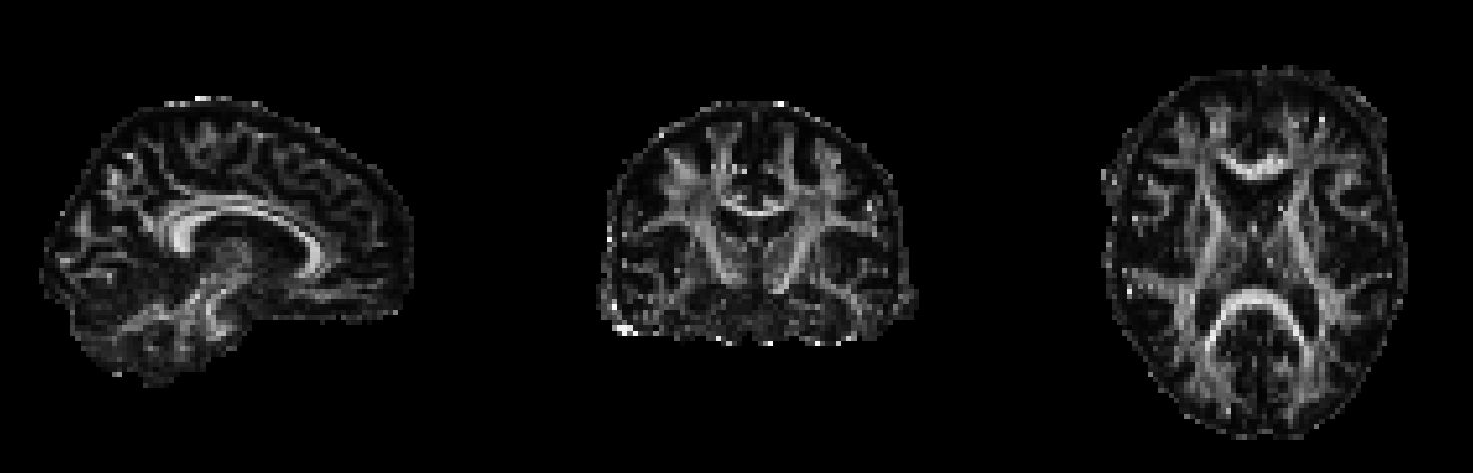
--out=dti \

--wls --sse



**DTI output**

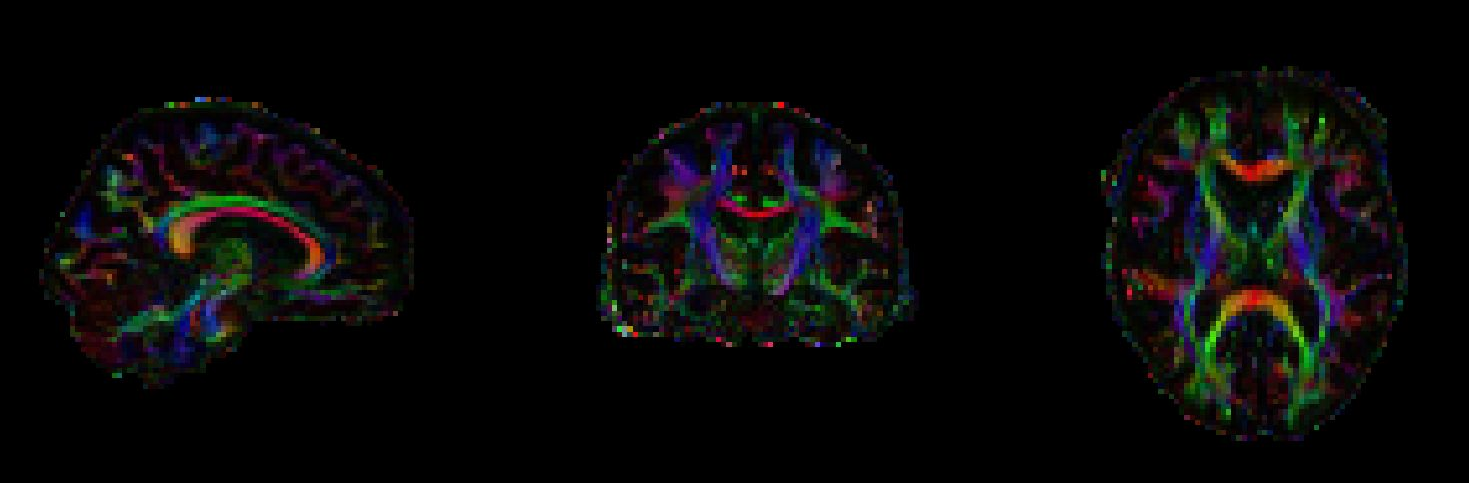
When you have run dtifit, load the anisotropy map dti\_FA into FSLeyes.



Bottom of Form

Add the principal eigenvector map to your display: *File > Add from file > dti\_V1.*

FSLeyes should open the image as a *3-direction vector image (RGB)*. Diffusion direction is now coded by colour. For a more interpretable image, we can modulate the colour intensity with the FA map so anisotropic voxels appear bright. In the display panel (https://fsl.fmrib.ox.ac.uk/fslcourse/2019_Beijing/lectures/FDT/FSL%20Diffusion%20Toolbox%20Practical_files/gear_icon.png) change the **Modulate by** setting to dti\_FA. You may wish to adjust the brightness and contrast settings.



Change the **Modulate by** setting back to *None*, and then change the **Overlay data type** to *3-direction vector image (Line)*. Zoom in and out of your data. You should see clear white matter pathways through the vector field.



Finally, change the **Overlay data type** to *3D/4D volume*. The image should now look quite messy - you are looking at the first (X) component (the first volume) of the vector in each voxel. The second and third volumes contain the Y and Z components of the vector.

Now load the Mean Diffusivity map (dti\_MD) into FSLeyes. Change the minimum/maximum display range values to 0 and 0.001 respectively.



**TBSS - Tract-Based Spatial Statistics**

FSLwiki: <https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/TBSS>.

So far, we have stepped through the steps required to pre-process a typical diffusion-weighted data. The final part of the tutorial focusses on running TBSS to compare FA values between two groups of participants.

We will now run a TBSS analysis of a small dataset – 11 controls and 11 patients with grade I Meningioma. We will attempt to find whether there are any FA differences between the two groups.

**1. Creating FA data from a diffusion study**

We have already created the FA images from the 22 subjects, using dtifit. So now we just need to create a new directory called TBSS and copy the relevant FA images to that folder. We will use all 11 control subjects and the 11 patients with grade I Meningioma (PAT01, PAT02, PAT03, PAT06, PAT08, PAT10, PAT13, PAT14, PAT15, PAT17 and PAT19).

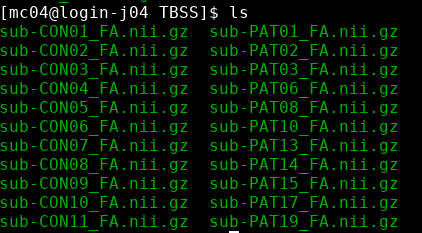
cd /Your/data/folder/ds001226-download/

mkdir TBSS

Once this is done, change into the TBSS directory and use cd to check all files were copied correctly.

cd TBSS

ls



Remember that all of the main scripts (until you reach the randomise stage near the end) need to be run from within this directory.

**LOOK AT YOUR DATA**

* Open one image with FSleyes to get a feel for the raw data quality and resolution. You may need to adjust the intensity display range. Look at the image histogram (*View -> Histogram*).
* Run slicesdir \*.nii.gz and open the resulting web page report; this is a quick-and-dirty way of checking through all the original images.

**2. Preparing your FA data for TBSS**

We must first run an additional pre-processing script which will erode your FA images slightly to remove brain-edge artifacts and zero the end slices (again to remove likely outliers from the diffusion tensor fitting). Type:

tbss\_1\_preproc \*.nii.gz

The script places its output in a newly-created sub-directory called FA. It will also create a sub-directory called origdata and place all your original images in there for posterity.



L**OOK AT YOUR DATA**

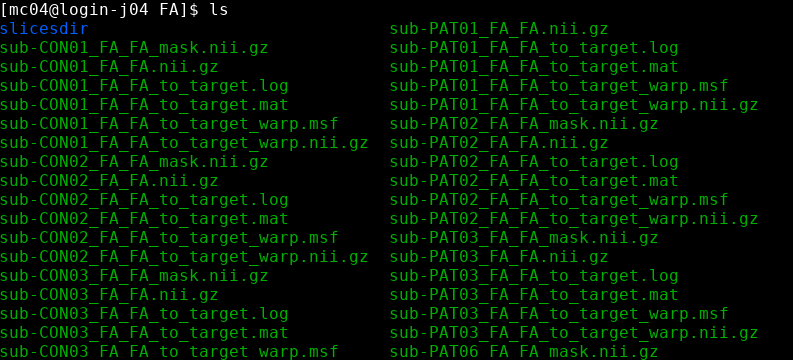
Check the images created in the FA directory. The tbss\_1\_preproc script will have re-run slicesdir on the preprocessed FA maps - open this report (you can find it in FA/slicesdir/index.html) and compare it to the slicesdir report you created earlier.

**3. Registering all the FA data**

The next TBSS script runs the nonlinear registration, aligning all the FA data across subjects. The recommended approach is to align every FA image to the FMRIB58\_FA template. This process can take a long time, as each registration takes around 10 minutes. You can easily speed this up if you have multiple computers running cluster software such as SGE (Sun Grid Engine).

tbss\_2\_reg -T

Once this step has finished, you will see lots of new files in TBSS/FA/:



**4. Post-registration processing**

The previous script (tbss\_2\_reg) only got as far as registering all subjects to the chosen template. The tbss\_3\_postreg script applies these registrations to take all subjects into 1x1x1mm standard space.

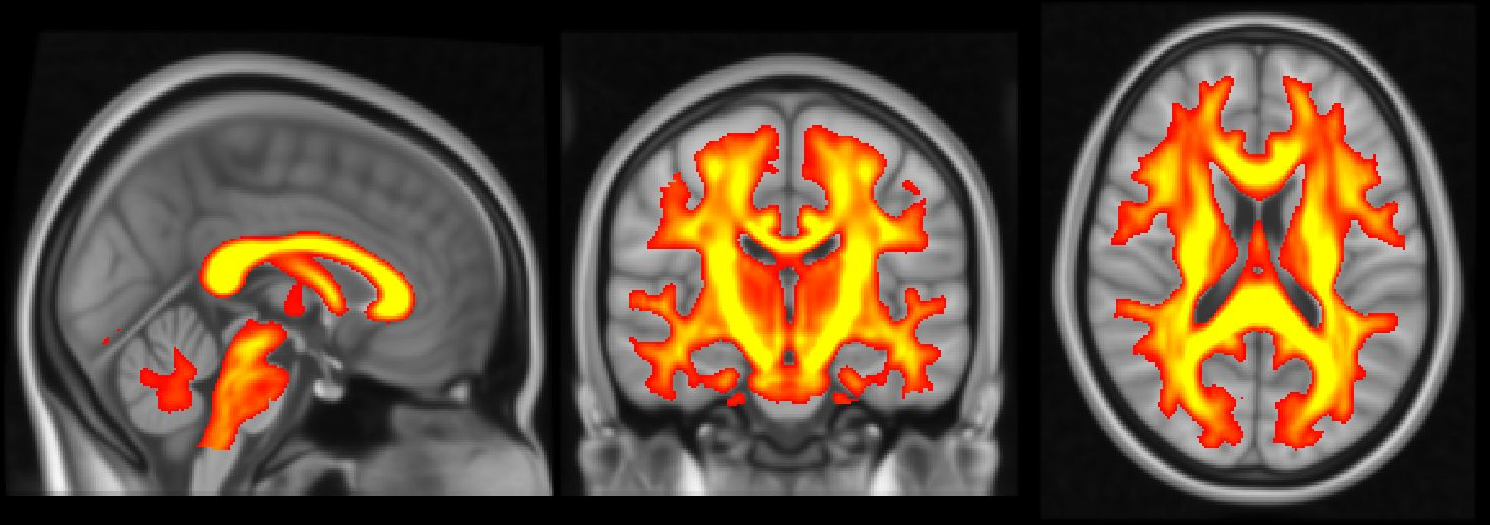
tbss\_3\_postreg -S

This script also merges all of the subjects' standard space nonlinearly aligned images into a single 4D image file called all\_FA, created in a new subdirectory called stats. The mean of all FA images is created, called mean\_FA, and this is then fed into the FA skeletonisation program to create mean\_FA\_skeleton.

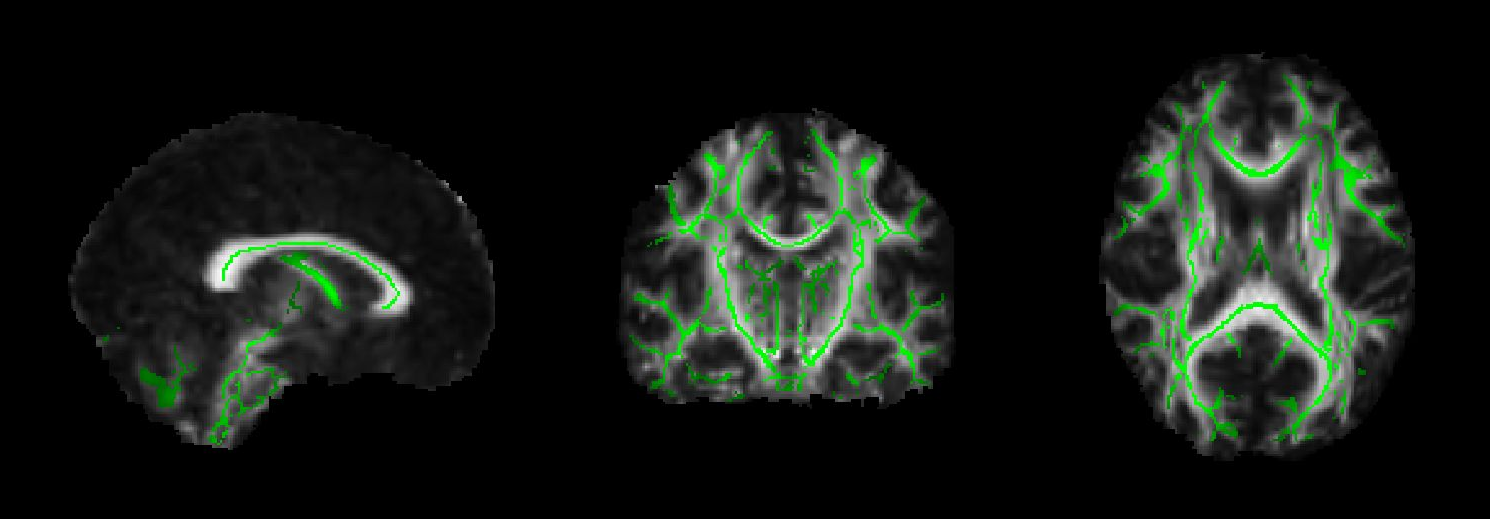
Once the script has finished running, check that the mean FA image looks reasonable, and is well aligned with the MNI152 image:

cd stats

fsleyes -std1mm mean\_FA -cm red-yellow -dr 0.2 0.6



As you move around in the image you should see that the mean FA image is indeed well aligned to standard space and corresponds to white matter in the MNI152 image. Remove these images (*Overlay -> Remove all*), then open the aligned FA maps for all subjects and the mean FA skeleton (*File -> Add from file*, and select all\_FA and mean\_FA\_skeleton). Change the colour map for mean\_FA\_skeleton to *Green*, and the display range to 0.2 - 0.6.



Select all\_FA and turn on the movie loop (https://fsl.fmrib.ox.ac.uk/fslcourse/2019_Beijing/lectures/FDT/FSL%20Diffusion%20Toolbox%20Practical_files/movie_icon.png); you will see the mean FA skeleton on top of each different subject's aligned FA image. If all the processing so far has worked, the skeleton should look like the examples shown in the lecture. If the registration has worked well you should see that in general each subject's major tracts are reasonably well aligned to the relevant parts of the skeleton.

**5. Projecting all pre-aligned FA data onto the skeleton**

The last TBSS script carries out the final steps necessary before you run the voxelwise cross-subject stats. It thresholds the mean FA skeleton image at the chosen threshold:

If you're still in the stats directory:

cd ..

Then:

tbss\_4\_prestats 0.2

This takes 4-5 minutes to run. The thresholding creates a binary skeleton mask that defines the set of voxels used in all subsequent processing.

Next a "distance map" is created from the skeleton mask. This is used in the projection of each subject's FA onto the skeleton; when searching outwards from a skeleton voxel for the local tract centre, the search only continues while the distance map values keep increasing - this means that the search knows to stop when it has got more than halfway between the starting skeleton point and another separate part of the skeleton.

Finally, the script takes the 4D pre-aligned FA images in all\_FA and, for each "timepoint" (subject ID), projects the FA data onto the mean FA skeleton. This results in a 4D image file containing the (projected) skeletonised FA data. It is this file that you will feed into voxelwise statistics in the next section.

Once the script has finished, cd into stats and have a look at all\_FA\_skeletonised in FSLeyes - turn on movie mode to see the different timepoints of the skeletonised data.



**6. Voxelwise statistics on the skeletonised FA data**

The previous step resulted in the 4D skeletonised FA image. It is this that you now feed into voxelwise statistics, that, for example, tells you which FA skeleton voxels are significantly different between two groups of subjects.

The recommended way of doing the stats is to use the randomise tool. For more detail see the [Randomise manual](https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Randomise). Before running randomise you will need to generate design matrix and contrast files (e.g., design.mat and design.con). Since we are only interested in the comparison between the two groups, with no covariates, we will use the design\_ttest2 command:

cd stats

design\_ttest2 design 11 11

This step creates two files named design.mat and design.con. In the terminal you can use less to look at the design.mat and design.con files.

You are now ready to run the stats using randomise. TBSS developers recommend you should use --T2 option for TFCE instead of cluster-based thresholding for correction of multiple comparisons. Typically you would run 5000 permutations, but in the interest of time we will run just 500:

randomise -i all\_FA\_skeletonised \

-o tbss \

-m mean\_FA\_skeleton\_mask \

-d design.mat \

-t design.con \

-n 500 \

--T2 -V

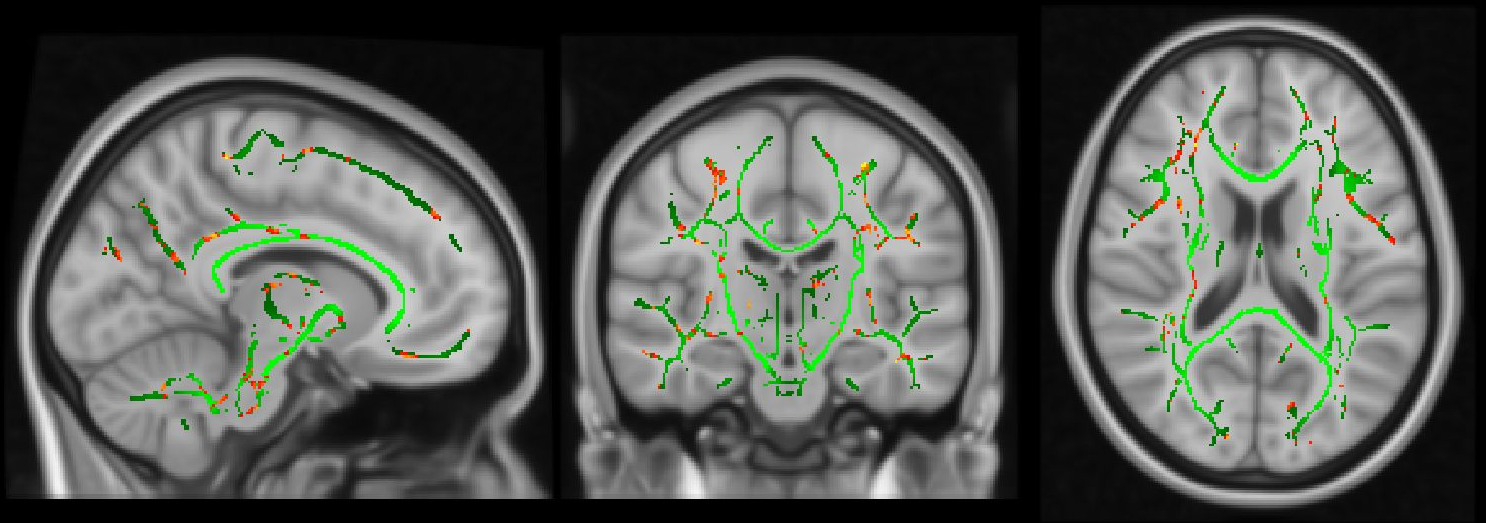
Contrast 1 gives the controls>patients test. The raw unthresholded tstat image is tbss\_tstat1 and the corresponding (p-values corrected for multiple comparisons) cluster image is tbss\_tfce\_corrp\_tstat1.

Thresholding clusters at 0.95 (corresponding to thresholding the p-values at 0.05, because randomise outputs p-values as 1-p for convenience of display - so that higher values are more significant). The following shows unthresholded t-stats in red-yellow and corrected t-stats in blue:

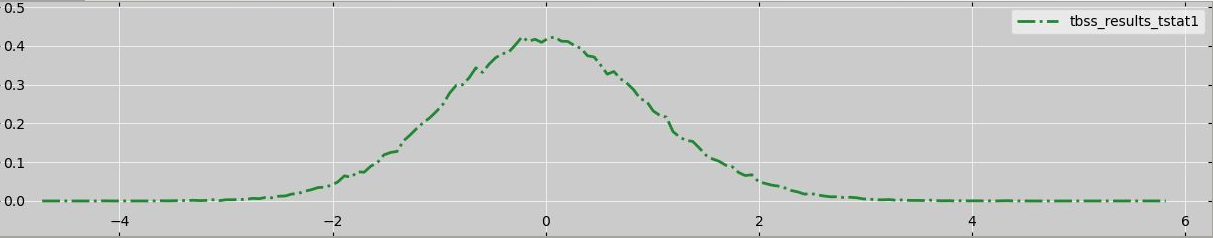
fsleyes -std1mm mean\_FA\_skeleton -cm green -dr .3 .7 \

tbss\_tstat1 -cm red-yellow -dr 1 3 \

tbss\_tfce\_corrp\_tstat1 -cm blue-lightblue -dr 0.95 1 &



In this case we find that there are no differences that survive correction for multiple comparisons. This is further confirmed by looking at the histogram of tstat1 (*View -> Histogram*): it is perfectly cantered around zero, suggesting no difference in FA between controls and patients.



This is also the case for contrast 2, which gives the patients>controls comparison:

