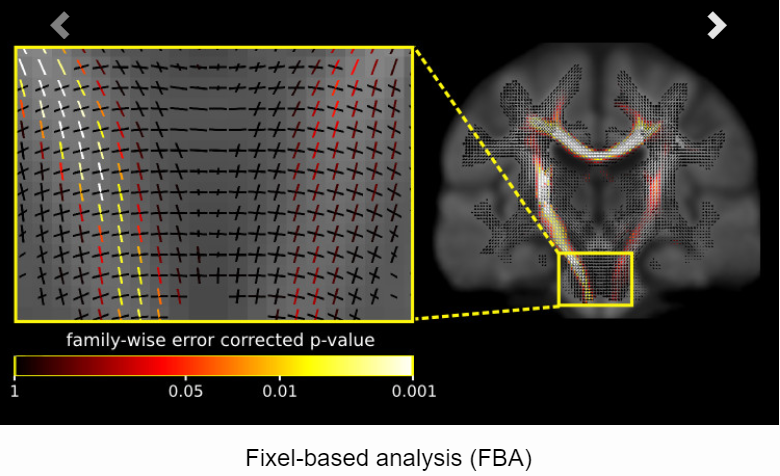
DPRC Diffusion FBA Pipeline

This pipeline takes the data from the CSD pipeline. Now, we will conduct *fixel-based analysis (FBA)* in order to derive 3 diffusion fibre metrics, known as *apparent fibre density (AFD)* aka *fibre density (FD), fibre cross-section (FC)*, and a combination of both of them, *fibre density and cross-section (FDC)*. We will perform statistical analysis on these metrics. We will also perform whole-brain fibre tractography. Script and files needed can be downloaded from here: <https://github.com/Ltah72/DPRC-analysis>.



<https://mrtrix.readthedocs.io/en/latest/fixel_based_analysis/mt_fibre_density_cross-section.html>

\*Steps:

1. Compute a white matter template analysis fixel mask
2. Estimate participants’ fixels and FBA metrics
3. Warp FOD images to template space
4. Segment FOD images to estimate fixels and their FD metric
5. Reorient fixels
6. Assign subject fixels to template fixels
7. Compute FC metric
8. Compute FDC metric
9. Perform whole-brain fibre tractography on the FOD template
10. Reduce biases in tractogram densities (using SIFT)
11. Generate fixel-fixel connectivity matrix
12. Smooth fixel data using fixel-fixel connectivity
13. Perform statistical analysis of FD, FC, and FDC
14. Visualise results
15. Display results with streamlines
16. Reduce number of streamlines to 200k
17. Create .tsf file (map fixel values to streamlines)
18. Visualise .tsf files in mrview
19. Smooth the .tsf file
20. FBA post-statistical inference
21. Calculate whole-brain FBA metrics per each participant and put onto a text file
22. Express the effect size relative to controls

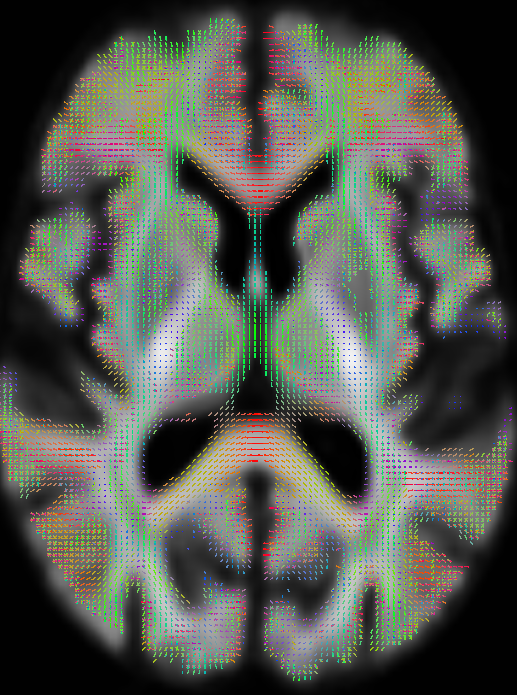
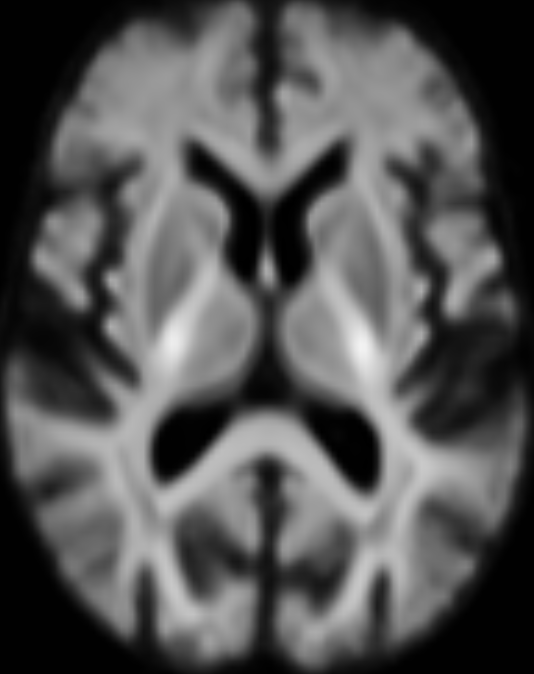
On the dementia vm (8 CPU cores with 8 GB RAM), steps 1-9 take approximately ~36 hours for 30 participants. All timings are elapsed from 1 participant, unless otherwise stated.

**\*Note that before these running these steps in this pipeline, you must decide upon the statistical analysis tests that you will do. You will need to create the design and contrast matrix files for this. See** [**here**](#Create_matrices) **in step 7 for more info.**

1. **Compute a white matter template analysis fixel mask**

In this step, we segment fixels from the FOD template. The result is the *fixel mask* that defines the fixels for which statistical analysis will later on be performed (and hence also which fixels’ statistics can support others via the mechanism of connectivity-based fixel enhancement (CFE) [[Raffelt2015]](https://mrtrix.readthedocs.io/en/latest/reference/references.html#raffelt2015))

Command: fod2fixel

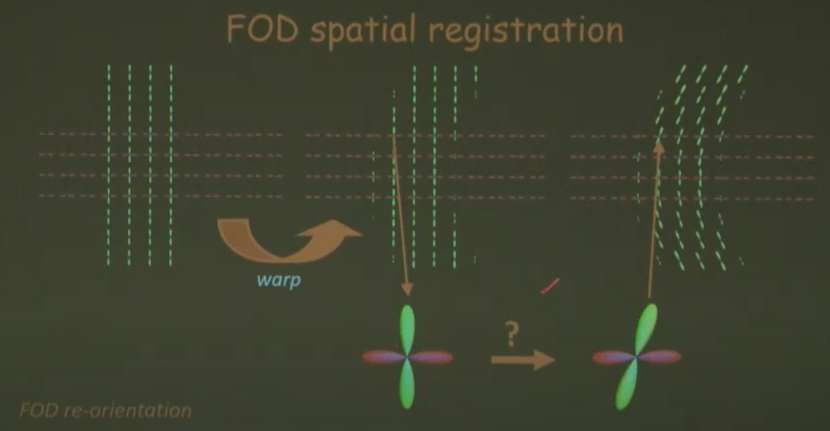


wmfod\_template.mif fixel\_mask (index and direction values)

\*You can view fixels by using the *fixel plot tool* on *mrview*. To view the fixels, you need to load in the directions.mif, index.mif, and any metric file (e.g. fd.mif), if available.

Elapsed time: 36 sec

1. **Estimate participants’ fixels and FBA metrics**

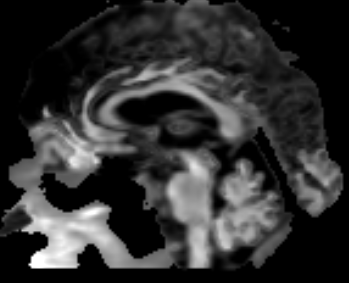
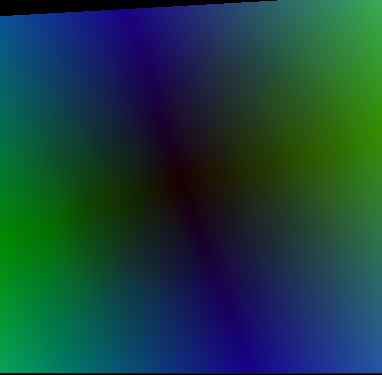


This is FOD reorientation that needs to be applied after warping to template space.

1. **Warp FOD images to template space**

Here, we warp FOD images into template space without FOD reorientation, as reorientation will be performed in a separate subsequent step (after fixel segmentation).

Command: mrtransform

wmfod\_norm.mif subject2template.mif (the warp)

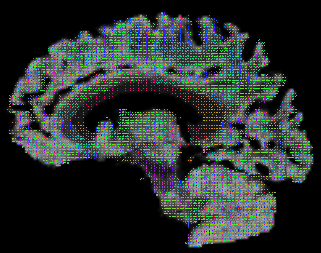
fod\_in\_template\_space\_NOT\_REORIENTED.mif

Elapsed time: 16 sec

1. **Segment FOD images to estimate fixels and their FD metric**

Here we segment each FOD lobe to identify the number and orientation of fixels in each voxel. The output also contains the apparent fibre density (AFD) value per fixel (estimated as the FOD lobe integral).

Command: fod2fixel

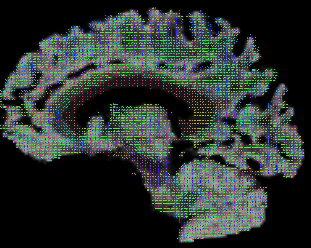
fod\_in\_template\_space\_NOT\_REORIENTED.mif NOT REORIENTED

Elapsed time: 42 sec

1. **Reorient fixels**

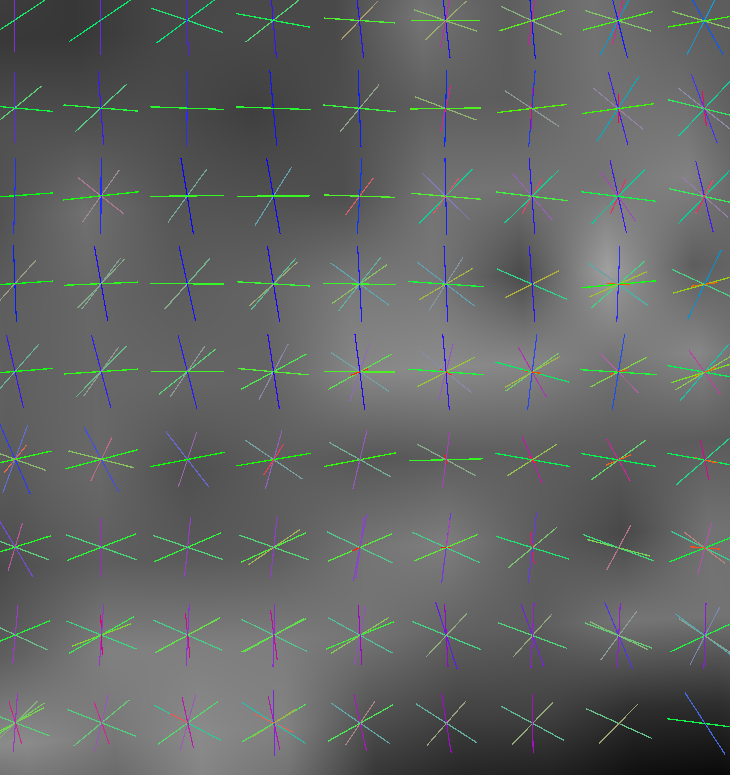
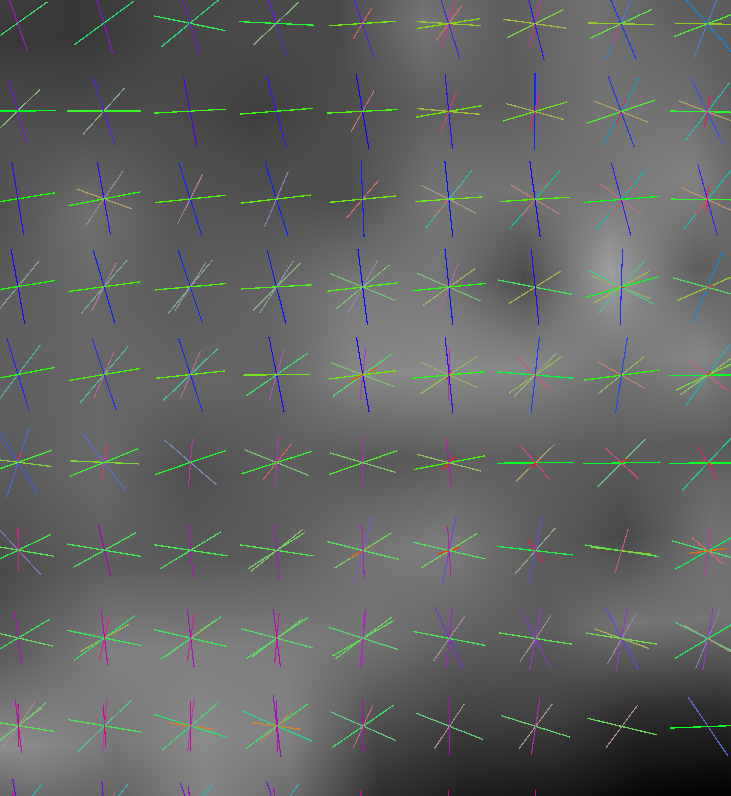
Here we reorient the fixels of all subjects in template space based on the local transformation at each voxel in the warps used previously.

Command: fixelreorient

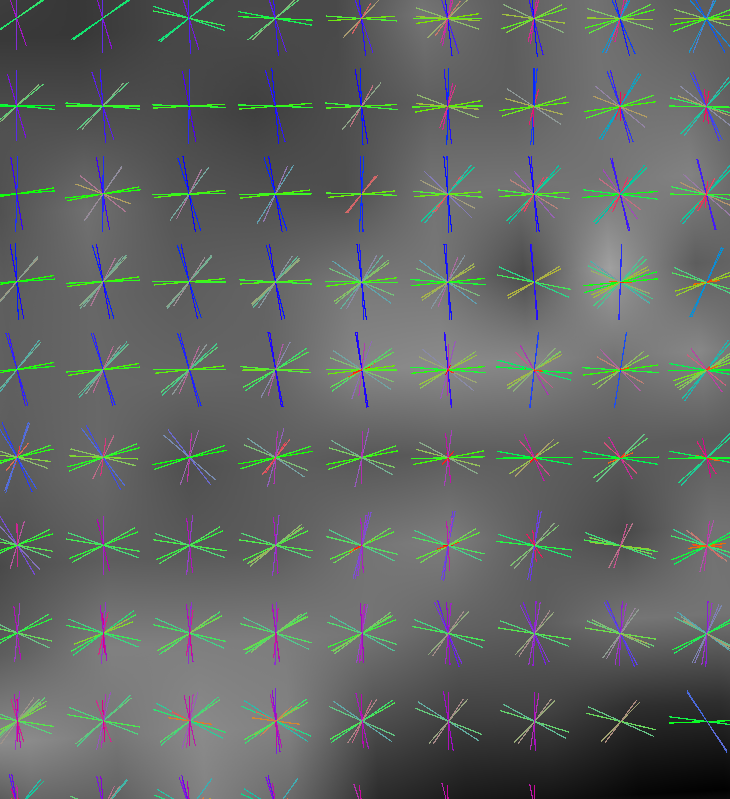


REORIENTED

Zoom into the fixels to see the rotation of the fixels:

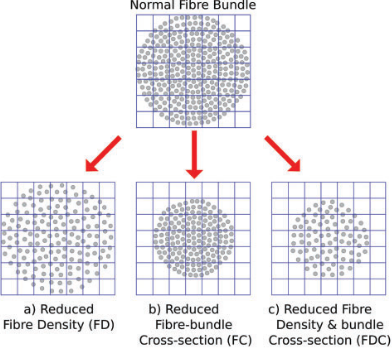
 

NOT REORIENTED REORIENTED



You can see that there is a DIFFERENCE between these fixels (NOT REORIENTED vs. REORIENTED) when you overlay them together. Tournier talks in great detail about how this works in orientating fixels here, at around 6 min + <https://www.youtube.com/watch?v=adACBScwBJ4>

Elapsed time: 3 sec



*This figure shows the fixel-based metrics. FD is showing reduction in the fibres within the voxels – within-voxel fibre density (more microstructural), while FC is showing a reduction across the fibre bundles (more macroscopic). FDC is a combination of both of these measurements.*

1. **Assign subject fixels to template fixels (compute FD)**

While each subject’s data has already been (spatially) warped to the common template space, and subject fixels have been reoriented accordingly, there is still no specification of which fixels match (across subjects, and between the subject and template fixels). This step establishes exactly that, by matching the fixels of each individual subject to the single common set of template fixels (which then inherently also defines how they match across subjects). This is achieved by, for each fixel in the template fixel mask, identifying the corresponding fixel in the matching voxel of the subject image and assigning the FD value of this corresponding subject fixel to that fixel in template space. If no fixel exists or can be found in a subject that corresponds to a given template fixel then it is assigned a value of zero (as the absence of a subject fixel at this stage is most likely due to a very low, or even zero, FD).

Command: fixelcorrespondence

Note that the output fixel directory ../fixel\_directory/PAR\_NAME/fd is the same for all subjects. This makes sense, since after this operation, there is only a single remaining set of fixels (i.e. the template fixels), with corresponding FD values as obtained from each subject. This resulting directory ../fixel\_directory/PAR\_NAME /fd now stores these data as individual fixel data files: one for each subject, and all with respect to a single set of corresponding template fixels. This way of storing the entire population’s FD data is then ready for input to fixelcfestats later on.

Elapsed time: 3 sec

1. **Compute FC metric**

The fibre density metric, mapped directly without any modulation to the fixel template space as above, is only sensitive to the original density of intra-axonal space in each voxel (see figure above). In other words, it ignores the cross-sectional size of the bundle, which is another property that would factor into the bundle’s total intra-axonal space across its full cross-sectional extent, and hence influence its total capacity to carry information. In certain cases, for example, atrophy may impact this cross-sectional size, but not per se the local fibre density metric.

In this step, we compute a fixel-based metric related to morphological differences in fibre cross-section (FC), where information is derived entirely from the warps generated during registration (see [[Raffelt2017]](https://mrtrix.readthedocs.io/en/latest/reference/references.html#raffelt2017) for more information)

Command: warp2metric

Elapsed time: 3 sec

You need to also take the log of the FC because FC alone is a *relative* metric, expressing the local fixel-wise cross-sectional size *relative* to this study’s population template. While this makes it possible to interpret differences of FC *within* a single study (because only a single unique template is used in the study), the FC values should not be compared across different studies that each have their own population template. Reporting absolute quantities of FC, or absolute effect sizes of FC, also provides little information; as again, it is only meaningful with respect to the template.

Command: mrcalc -log

Elapsed time: < 1 sec

1. **Compute FDC metric**

The total capacity of a fibre bundle to carry information, is modulated both by the local fibre density at the voxel (fixel) level, as well as its cross-sectional size (see figure above). Here we compute a combined metric, which factors in the effects of both FD and FC, resulting in a fibre density and cross-section (FDC) metric.

Command: mrcalc -mult

Elapsed time: < 1 sec

1. **Perform whole-brain fibre tractography on the FOD template**

Statistical analysis using connectivity-based fixel enhancement (CFE) [[Raffelt2015]](https://mrtrix.readthedocs.io/en/latest/reference/references.html#raffelt2015) exploits local connectivity information derived from probabilistic fibre tractography, which acts as a neighbourhood definition for threshold-free enhancement of locally clustered statistic values. To generate a whole-brain tractogram from the FOD template (note the remaining steps from here on are executed from the template directory). Tckgen’s default algorithm is iFOD2 (probabilistic algorithm).

1. We will generate 20 million tracts, using a 0.06 cut-off value (as suggested by MRtrix). You can also start out with 100K tracts to generate a smaller number of streamlines in order to visually confirm that the generated streamlines exhibit an appropriate extent of propagation at the ends of white matter pathways, before committing to generation of the dense tractogram.

Command: tckgen

**(image too big to show)**

%Sample output: tracks\_20\_million.tck

Elapsed time: 6 hours and 33 min, for 36 participants

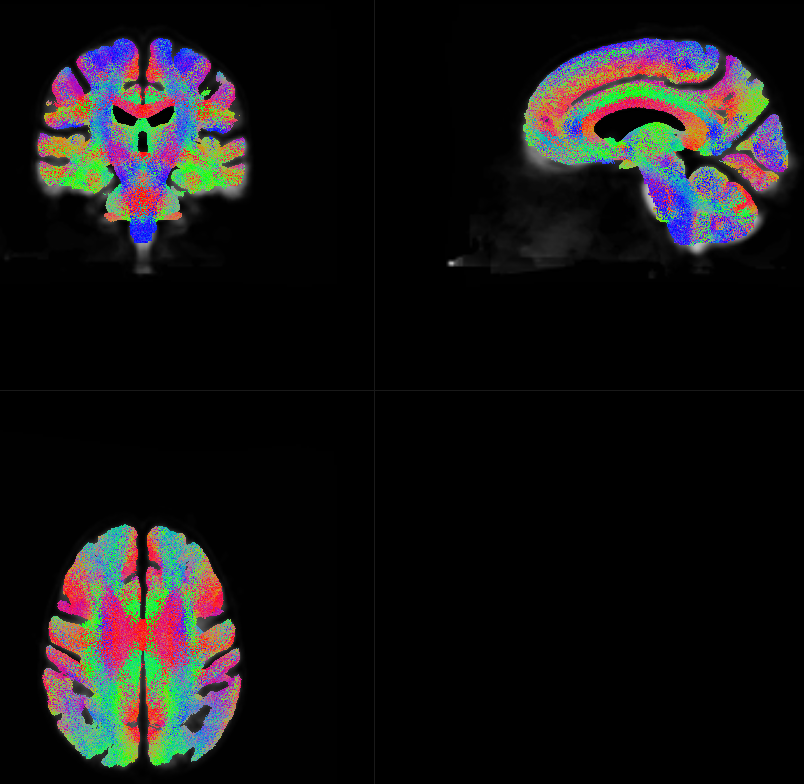
Note that the anatomically constrained tractography (ACT) option (which utilises anatomical images, e.g. T1, FLAIR) is only done when generating a connectome, which is a separate pipeline / analysis that you can do later on.

1. **Reduce biases in tractogram densities (using SIFT)**

Perform Spherical-deconvolution Informed Filtering of Tractograms (SIFT) to reduce tractography biases in the whole-brain tractogram. Use this to make improve the quantitative nature of whole-brain streamlines reconstructions. By producing a reconstruction where the streamlines densities are proportional to the fibre densities as estimated by spherical deconvolution throughout the white matter, the number of streamlines connecting two regions becomes a proportional estimate of the cross-sectional area of the fibres connecting those two regions.

This will weigh the streamlines selectively to match the contribution to the FODs. Good for false positives – kind of ‘cleans up’ the data. Removes the non-relevant streamlines. Makes it more representative, and this is especially good for when generating a connectome. We will filter with 2 million tracts (as suggested by MRtrix).

Command: tcksift



From 20 million to 2 million tracks using SIFT, overlayed on wmfod\_template.mif

%Sample output: tracks\_2\_million\_sift.tck

Elapsed time: 1 hour and 39 min, for 36 participants

1. **Generate fixel-fixel connectivity matrix**

Generation of the fixel-fixel connectivity matrix based on the whole-brain streamlines tractogram.

Command: fixelconnectivity

The output directory should contain three images: index.mif, fixels.mif and values.mif; these are used to encode the fixel-fixel connectivity that is by its nature sparse.

Elapsed time: 17 min, for 36 participants

1. **Smooth fixel data using fixel-fixel connectivity**

Smoothing of fixel data is performed based on the sparse fixel-fixel connectivity matrix.

Command: fixelfilter

Elapsed time: 18 min, for 36 participants

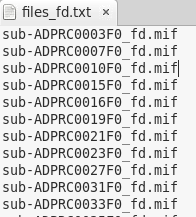
1. **Perform statistical analysis of FD, FC, and FDC**

Statistical analysis using connectivity-based fixel enhancement (CFE) is performed separately for each metric (FD, log(FC), and FDC).

Command: fixelcfestats

The input files.txt (e.g. *files\_fd.txt, files\_log\_fc.txt*, and *files\_fdc.txt*) are text files containing the filename of each file (i.e. not the full path) to be analysed inside the input fixel directory, each filename on a separate line. The line ordering should correspond to the lines in the file design\_matrix.txt.

There are 3 input files.txt (for the 3 metrics) that contain the list of the participants on a separate line in your study. Using the **CreateParticipantFixelList.m** function, my script will create these lists for you based upon the participants you ran, automatically.



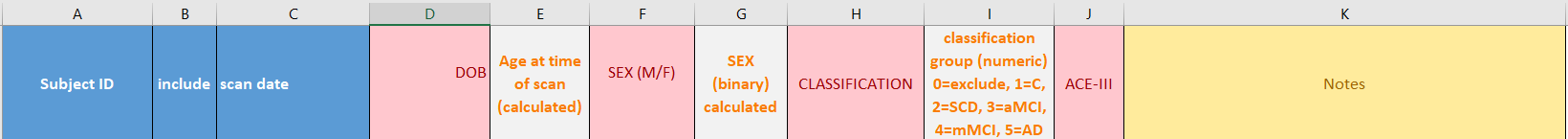
**\*Note that you must create these matrices previous to running the pipeline. This will determine the statistical tests that you will run.**

User has option to automate matrices creation, with creator’s customised study (i.e. compare 5 groups – HC, SCD, aMCI, mMCI, and AD) with 3 covariates (age, sex, and overall ACE-III score). If user is running a different study with different groups/covariates, then they will need to create their own study matrices, and place them into the stats\_matrices folder. If user is using the automated version, the given DPRC excel file is organised in this manner, with these 11 columns (A:K):

*Subjects(A) Include(B) Scan Date(C) DOB(D) Age at time of scan(E)*

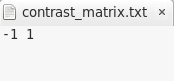
*Sex (M/F)(F) Sex (binary)(G) Classification(H) Classification Group(I)*

*ACE-III(J) Notes(K)*

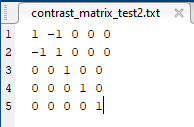
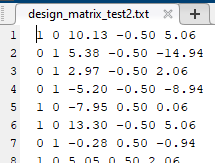


The *design\_matrix.txt* file contains the group (i.e. status) of each of the participants. This is dummy-coded (e,g. 0, 1, 2, 3…etc). This must be manually done (so far).The *contrast\_matrix.txt* file is specified as rows of weights (e.g. -1, 1 or -1, 1, 0, 0). This must be manually done (so far). The matrices ***must match*** in the number of columns.

Below is an example of running a two-sample unpaired t-test:

If you included 3 covariates (e.g. age, sex, ACE-III), it may look something like this:



*Design matrix: Each row corresponds to Contrast matrix: c1 = Where is*

*a participant. group one’s mean greater than group 2’s mean? c2 = Where is group two’s mean greater than group 1’s mean? c3 = Do FBA metrics correlate with age? c4 = Do FBA metrics correlate with sex? c5 = Do FBA metrics correlate with the ACE-III score?*

Another example: If you had 4 groups/conditions (i.e. 1-way between-subjects ANOVA), with the question, *do the group averages account some significant effect?* Then, your design matrix and your contrast matrix would look like this:

design\_matrix.txt contrast\_matrix.txt

1 0 0 0 1 0 0 0

1 0 0 0 0 1 0 0

0 1 0 0 0 0 1 0

0 1 0 0 0 0 0 1

0 0 1 0

0 0 1 0

0 0 0 1

0 0 0 1

….

(2 participants per group) (4 groups)

MRtrix uses the same GLM format as FSL, so be sure to check out their manual for more information: <https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/GLM>

Your output will be in the stats folder of each apparent fibre density (AFD) metric (e.g. stats\_fd). Use this for step 8, to visualise the results.

Elapsed time: 27 hours and 12 min, for 29 participants (with 5 contrasts, i.e. 5 groups/conditions) \*

\*Do note that this can vary depending on the number of contrasts you have set up. Also, this step takes a lot of RAM, and so, it is best to have other processes closed in order for this to work properly, as sometimes the process can get ‘killed’, if overwhelmed. As far as I know, it is okay for more than one user to run fixelcfestats simultaneously, but I found that when one user is running tcksift (RAM hungry), then fixelcfestats would get killed, if it was in its initial process (permutations step < 3%). I think fixelcfestats runs an initial test to see whether there is ‘enough memory’ available on the system, and if not, then it gets killed. This can be tricked, however, in that if fixelcfestats is run and passes the initial threshold (permutations step > %3), then other processes can be run simultaneously without problems.

*Raffelt, D. A., Smith, R. E., Ridgway, G. R., Tournier, J. D., Vaughan, D. N., Rose, S., … Connelly, A. (2015). Connectivity-based fixel enhancement: Whole-brain statistical analysis of diffusion MRI measures in the presence of crossing fibres.*

The non-parametric permutation statistics on the GLM is based on this paper:

*Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M., & Nichols, T. E. (2014). Permutation inference for the general linear model.*

Another thing to note is that if you have several hypotheses, it might be good to try and control for the increasing family-wise error (FWE) and implement the -strong option in this command. This will make sure that the FWE will stat at 5% for the number of hypothesis tests.

Since I am running 10 contrasts across the 3 FBA metrics….

1. - (0.95)30 = 79% 

* 79% chance of reporting at least one false positive, if you don’t control for FWE.

<https://community.mrtrix.org/t/fba-design-contrast-matrices-for-three-groups/1791/11>

Some more info on how to set your design and contrast matrices, depending on the type of statistical tests that you are running, and if you want to include co-variates:

<https://community.mrtrix.org/t/fba-design-contrast-matrices-for-three-groups/1791>

<https://community.mrtrix.org/t/statistical-analysis-of-fd-fc-and-fdc/840/15>

Longitudinal study design advice (use mrcalc and pre-calculate individual difference-over-time images):

<https://community.mrtrix.org/t/longitudinal-fixel-based-analysis/910/2>

<https://community.mrtrix.org/t/replicating-longitudinal-fixel-based-analysis-approach/2071/15>

<https://community.mrtrix.org/t/2x2-mixed-design/4072/3>

Helpful YouTube channel for neuroimaging statistics: <https://www.youtube.com/channel/UCZ7gF0zm35FwrFpDND6DWeA>

1. **Visualise results**

To view the results load the population FOD template image in mrview, and overlay the fixel images using the vector plot tool. Note that p-value images are saved as (1 - p-value; i.e. “fwe\_1mpvalue”). Therefore to visualise all results at a threshold of p < 0.05, within the mrview fixel plot tool, apply a lower threshold at a value of 0.95.

There is no ‘hard fix’ on what you should use to view your results – but for the most basic visualisation, you can load the *index.mif* file into the fixel. Then, use the *directions.mif* file for the ‘coloured by’ option (this is set as the default, anyways). Then, use the *fwe\_1mpvalue.mif* as the ‘threshold’ option, and set the lower threshold (on the left side, make sure to tick the box) to 0.95, so that it will display the significant fixels with a p < 0.05.

Other values (and what I think they are):

***abs\_effect\_t1.mif***

***beta0.mif***– the mean of the first group (e.g. control group), and the y-intercept of your linear model. Each subsequent beta value (e.g. beta1.mif…beta4.mif, etc) can be subtracted from beta0 to get the mean of that group. You can also use this value to calculate percentage differences (displaying the magnitude of the differences) between groups, other than from the control group (see my notes for this).

***cfe\_t1.mif*** – CFE stands for connectivity-based fixel enhancement, which MRtrix’s method for permutation testing. It provides FWE correction, and is more stringent than FDR and more powerful than Bonferroni. Formal definition -- ‘…contains the values resulting from the CFE statistical enhancement according to the equation presented in the paper.’ ‘What CFE provides (by default) is what’s known as *weak* FWE control. This is where, for each hypothesis test individually, you have a e.g. 5% likelihood of reporting a false positive anywhere; but each hypothesis is treated entirely independently.’

***directions.mif*** - these are the directions of the fixels, as shown in colour.

***fwe\_1mpvalue\_t1.mif*** - a probability value from 0-100, which shows significant differences across whole-brain fixels, with family-wise error correction. Can view this by setting the threshold at 0.95.

***index.mif*** – this is the template file of the fixels to load in under fixel plot.

***null\_contributions\_t1.mif***

***null\_dist\_t1.txt***

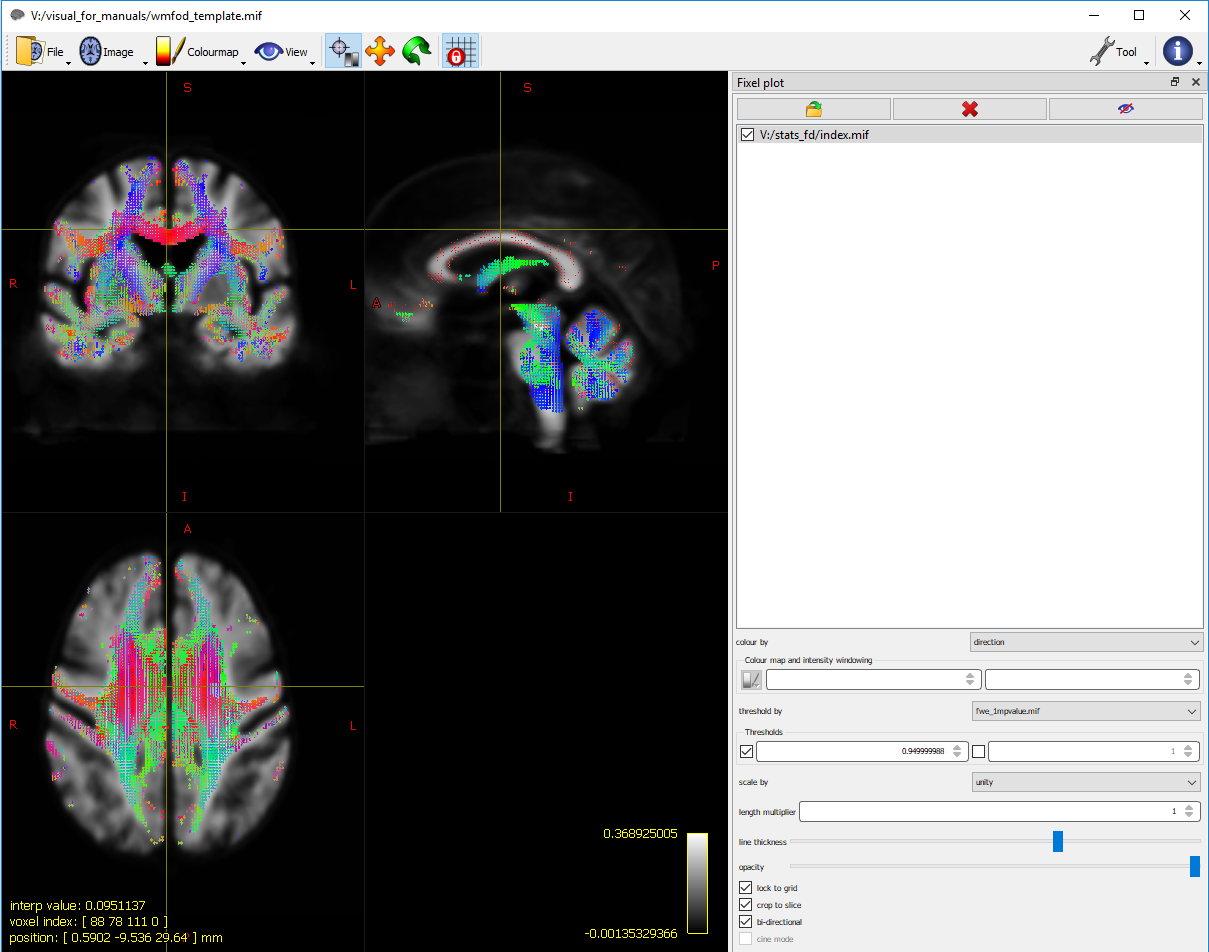
***std\_dev.mif***

***std\_effect\_t1.mif***

***tvalue\_t1.mif***- the t-value is associated with the p-value in statistics – a larger t-value (whether positive or negative) means that there would more likely be a significant difference (but this does not have family-wise error correction).

***uncorrected\_pvalue\_t1.mif*** - a probability value from 0-100, which shows significant differences across whole-brain fixels, but it is not being corrected with family-wise error. Can also view this by setting the threshold at 0.95. More formal definition – ‘the fraction of permutations where the enhanced statistic was greater than that of the default permutation for just that fixel, rather than using the maximal enhanced statistic across the entire image to build the null distribution.’

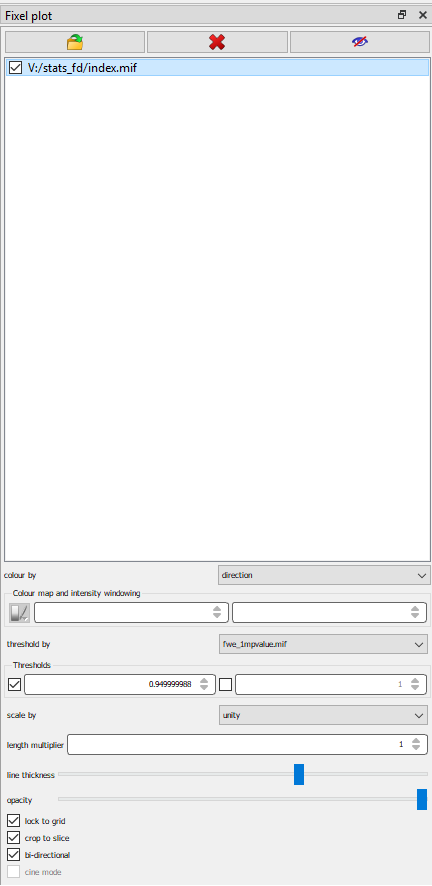
***Zstat\_t1.mif***



Whole-brain FBA with the FD metric seeing if there is a *correlation* between healthy controls with aMCI from the 29 ADPRC participants.



FD metric difference in whole-brain FBA between healthy older adults with no cognitive impairments (HC – 16 and SCD – 3) vs. cognitively impaired older adults (aMCI – 11, mMCI – 5, and AD – 1) (n = 36), but only when age is negatively correlated with white matter integrity (e.g., as age increases, white matter integrity decreases), meaning that these differences are only significant due to the correlation of age. Significant difference of p < 0.05, shown in the hippocampus and fornix regions.



You can change other options, like the line thickness and opacity for visualisation.

See here for more information on which files to choose from to view and which files to use for setting the threshold: <https://community.mrtrix.org/t/fba-displaying-significant-results/993>

1. **Display results with streamlines**

Display the streamlines from the whole-brain tractography analysis which correspond to the significant fixels (with FD, FC, and FDC metrics).

1. **Reduce number of streamlines to 200,000.**

Command: tckedit

%Sample output: tracks\_200k\_sift.tck

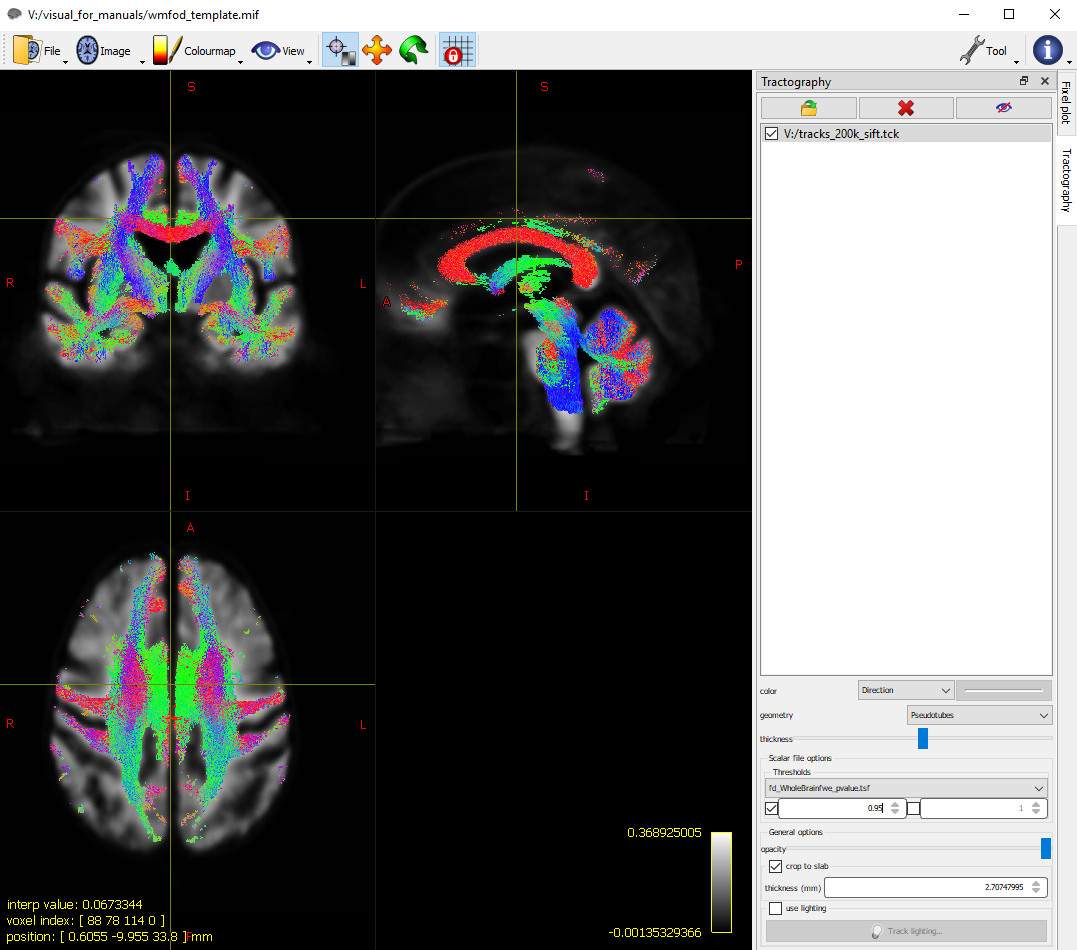
1. **Map fixel values to streamline points, and save them as a ‘track scalar file’ (.tsf).**

Command: fixel2tsf

%Sample output: fd\_WholeBrainfwe\_pvalue\_t4.tsf

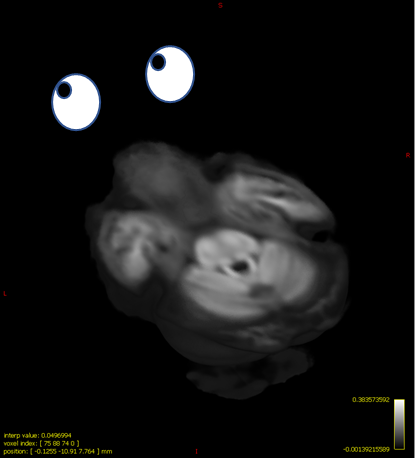
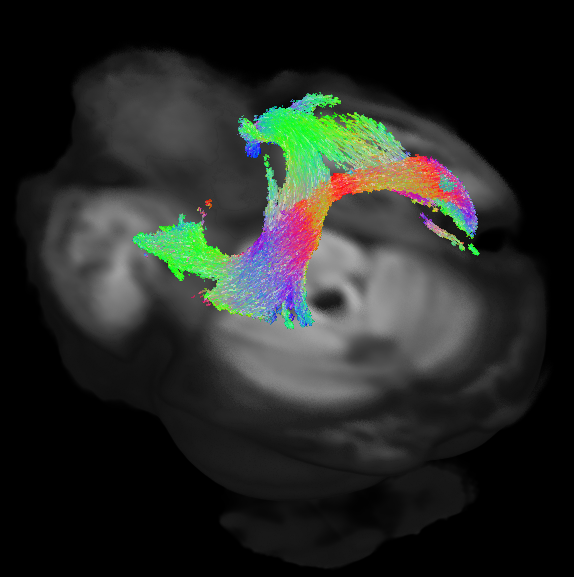
1. **Visualise the track scalar files (.tsf) with the tractogram tool, using mrview.**

First, load the streamlines (tracks\_200k\_sift.tck). Then to dynamically threshold (remove) streamline point by p-value select the “Thresholds” dropdown and select “Separate Scalar file” and set to 0.95.





See MRtrix documentation here for further information: <https://mrtrix.readthedocs.io/en/3.0_rc3/fixel_based_analysis/displaying_results_with_streamlines.html>

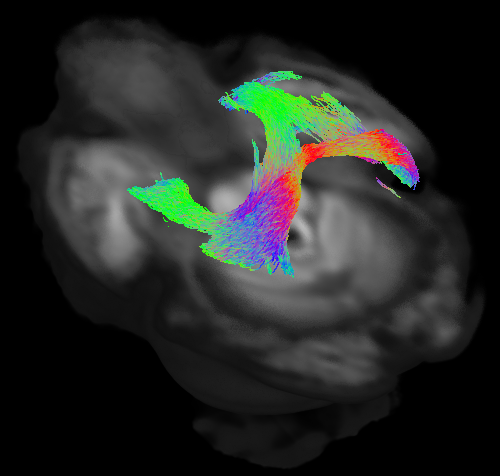
 

3D volume rendering of a track scalar file (.tsf) showing a significant difference (p < .05) in the fibre density (FD) metric with the contrast of controls being greater than individuals with Alzheimer’s disease (C > AD), with the covariate of clinical site, in whole-brain FBA. This shows that the fornix and part of the corpus callosum has a greater FD in C compared to AD.

1. **Smooth the .tsf file**

You can also apply a Gaussian smoothing filter on the track scalar file (.tsf).

Command: tsfsmooth



%Sample output: smoothed\_fd\_WholeBrainfwe\_pvalue\_t4.tsf

1. **FBA post-statistical inference**
2. **Calculate whole-brain FBA metrics per each participant and put onto a text file.**

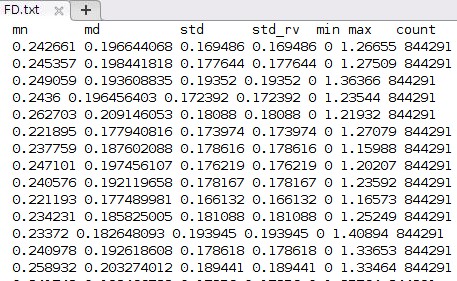
If you want to calculate a single measurement of the FBA metrics, of either the whole-brain tract analysis or a specific tract, you can do so by taking the average of all participants. We will use the in-house function, **CreateWholeBrainFBAMetricFiles.m**, which will call upon MRtrix’s mrstats command to do this. Specifically, this function will generate the mean, median, std, std\_rv (standard error), min, max, and count per each metric (FD, log\_FC, and FDC) per each participant. This will be used for group comparisons. However, this may be more prevalent and relevant in tract of interest (TOI) analysis.

Command: mrstats

%Sample output: FD.txt

FC\_log.txt

FDC.txt



If you want to do this with specific tracts, you can by comparing the fixels against the fixel mask of that region / tract of interest.

<https://community.mrtrix.org/t/calculating-average-fba-metrics-of-specific-tracts/1805>

See more on MRtrix wiki post here: <https://community.mrtrix.org/t/fba-post-statistical-inference-tricks/2255>

1. **Express the effect size relative to controls**

The apparent Fibre Density (FD) and Fibre Density and Cross-section (FDC) are relative measures and have arbitrary units. Therefore, the units of *abs\_effect.mif* are not directly interpretable. In a patient-control group comparison t-test, one way to present results is to express the absolute effect size as a percentage relative to the control group mean.

Command: mrcalc

Sample command: mrcalc stats\_results/All\_groups/stats\_fd/abs\_effect\_t3.mif stats\_results/All\_groups/stats\_fd/beta0.mif -div 100 -mult stats\_results/All\_groups/stats\_fd/percentage\_effect\_t3.mif

Another thing that you can do, to express the effect size relative to other groups (not just controls, which would be beta0.mif), is that you can add the beta estimates together (e.g., beta0.mif + beta1.mif = SCD group (group 2)), to get the group mean, and then re-run the above command on that. For example:

mrcalc stats\_results/All\_groups/stats\_fd/beta0.mif stats\_results/All\_groups/stats\_fd/beta1.mif -add stats\_results/All\_groups/stats\_fd/mean\_group2.mif

\*Use mean\_group2.mif for the calculation of the percentage difference when comparing between other groups (not the control) vs. the SCD (this would be group 2) group.

See more on MRtrix documentation here: <https://mrtrix.readthedocs.io/en/latest/fixel_based_analysis/computing_effect_size_wrt_controls.html>

<https://community.mrtrix.org/t/fba-post-statistical-inference-tricks/2255>

References

Raffelt, D. A., Smith, R. E., Ridgway, G. R., Tournier, J. D., Vaughan, D. N., Rose, S., … Connelly, A. (2015). Connectivity-based fixel enhancement: Whole-brain statistical analysis of diffusion MRI measures in the presence of crossing fibres. *NeuroImage*, *117*, 40–55. https://doi.org/10.1016/j.neuroimage.2015.05.039

Raffelt, D. A., Tournier, J. D., Smith, R. E., Vaughan, D. N., Jackson, G., Ridgway, G. R., & Connelly, A. (2017). Investigating white matter fibre density and morphology using fixel-based analysis. *NeuroImage*, *144*, 58–73. https://doi.org/10.1016/j.neuroimage.2016.09.029

Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M., & Nichols, T. E. (2014). Permutation inference for the general linear model. *NeuroImage*, *92*, 381–397. https://doi.org/10.1016/j.neuroimage.2014.01.060