DPRC Diffusion Preprocessing Pipeline

Preprocessing script for DPRC diffusion data. Preprocessing is done on each participant for cleaning the raw data. Preprocessing based upon UKB data recommendation:

*Maximov & Westlye (2019). Towards an optimised processing pipeline for diffusion magnetic resonance imaging data: Effects of artefact corrections on diffusion metrics and their age associations in UK Biobank.*

*Alfaro-Almagro, F., Jenkinson, M., Bangerter, N. K., Andersson, J. L. R., Griffanti, L., Douaud, G., … Smith, S. M. (2018). Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank.*

Software programs, such as FSL, MRtrix3, ExploreDTI(?) and ANTs are utilised for these steps. Assumes BIDS formatting and organisation. This script calls upon 6\* preprocessing steps which are, in order:

**Pre-steps** (organise files and things, etc.)

**1. Noise correction** (denoising -- MP-PCA, *Veraart et al., 2016*)

**2. Gibbs ringing correction** (local sub-voxel shift*, Kellner et al., 2016*)

**Edit gradient files** (LastB0AddOn – in-house function)

**3. Field distortion** (TOPUP -- FSL)

2 steps together, aka ‘Geometric distortion correction' (GDC)

**a) BestB0 pair selection** (BestB0 – in-house function)

**4. Eddy current distortions** (Eddy -- FSL)

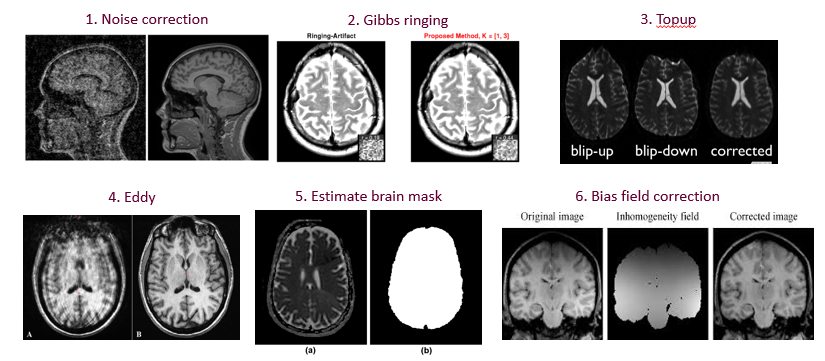
**5. Estimate a brain mask** (iterate this with bias field correction (step 6))

**6. Bias field correction** (ANTs -- N4BiasFieldCorrection, *Tustison et al., 2010*)

**a) generate second brain mask**

**\* Most(?) preprocessing steps are deployed through MRtrix3. MRtrix operates only through command-line usage and images can be viewed via *mrview* or if converted, through *fsleyes*. MRtrix preprocessing steps are detailed here:** [**https://mrtrix.readthedocs.io/en/0.3.16/workflows/DWI\_preprocessing\_for\_quantitative\_analysis.html**](https://mrtrix.readthedocs.io/en/0.3.16/workflows/DWI_preprocessing_for_quantitative_analysis.html)

[**https://mrtrix.readthedocs.io/en/latest/fixel\_based\_analysis/mt\_fibre\_density\_cross-section.html**](https://mrtrix.readthedocs.io/en/latest/fixel_based_analysis/mt_fibre_density_cross-section.html)



An overview of all of the MRtrix steps are here: <https://mrtrix.readthedocs.io/en/latest/>

*Tournier, J. D., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch, M., … Connelly, A. (2019). MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation.*

On the dementia vm, steps 1-6 take 1 hour and 32 minutes per subject. Once data is cleaned, then user may choose to run another pipeline to conduct further diffusion imaging analysis, such as for fitting diffusion tensors (e.g. VBA, TBSS) or fibre orientation models (e.g. BEDPOSTX, CSD).

PREPROCESSING

**Pre-steps:**

* First, create a derivatives folder for the output.
* Combine all diffusion data (the main dwi, the blip-up (BU), and the blip-down (BDs) – BU and BD are scans which are used for distortion correction)) to run steps 1 + 2 on them.
* Convert .nifti files to .mif files (mrtrix3 formatting) for processing.
* The input data for this step are the original source data files.

%Sample input: sub-ADPRC0012F0\_acq\_data\_dwi.nii 🡪 ‘main dwi’ has 99 dwi + 5 non dwi, B0s/BUs (phase encoding anterior to posterior (a-p)) = 104 vol

sub-ADPRC0012F0\_acq\_BU\_dwi.nii 🡪 1 non dwi , B0/BU (phase encoding

a-p)

sub-ADPRC0012F0\_acq\_BD1\_dwi.nii

sub-ADPRC0012F0\_acq\_BD2\_dwi.nii 🡪 3 BDs (phase encoding p-a)\*

sub-ADPRC0012F0\_acq\_BD3\_dwi.nii

\*Typically, participants will have 3 BD files, but some earlier participant scans may have less.

Command: mrcat Combine all volumes (main dwi, BUs, and BDs) to run steps 1

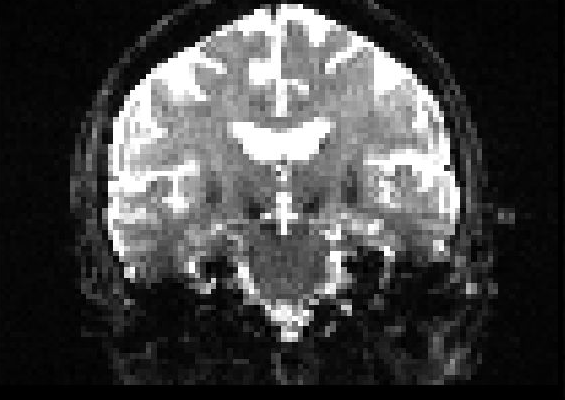
+ 2 on them

mrconvert Convert .nii to .mif (mrtrix3 formatting) for processing

%Sample output: combined\_sub-ADPRC001F0\_acq\_data\_dwi.nii

combined\_sub-ADPRC001F0\_acq\_data\_dwi.mif

*fsleyes contrast setting for images below: min 100, max 3000*

Non dwi - B0/BU (vol 0) dwi (vol 1)

1. **Denoising**

Noise correction is based upon principle component analysis of Marchenko-Pastur (MP-PCA) method. This thermal noise correction is the first step in data analysis due to an assumption about uncorrelated noise both spatially and across the diffusion space. The noise in the MR images can be described as a Rician distribution. MP-PCA estimates the standard deviation at each voxel for evaluation of the true signal.

*Veraart et al., (2016). Diffusion MRI noise mapping using random matrix theory.*

*Veraart et al., (2016). Denoising of diffusion MRI using random matric theory.*

%Sample input: combined\_sub-ADPRC001F0\_acq\_data\_dwi.mif

Command: dwidenoise Denoise the entire dataset (main dwi, B0/BUs and BDs).

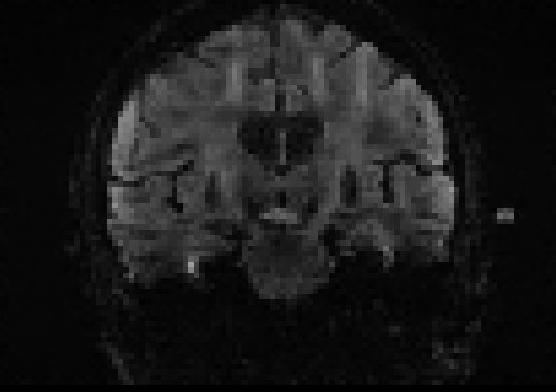
%Sample output: dsub-ADPRC001F0\_acq\_data\_dwi.mif 🡪 denoised data

noise\_sub-ADPRC001F0\_acq\_data\_dwi.mif 🡪 noise

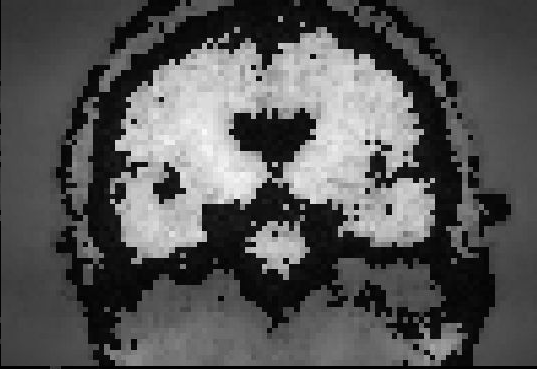
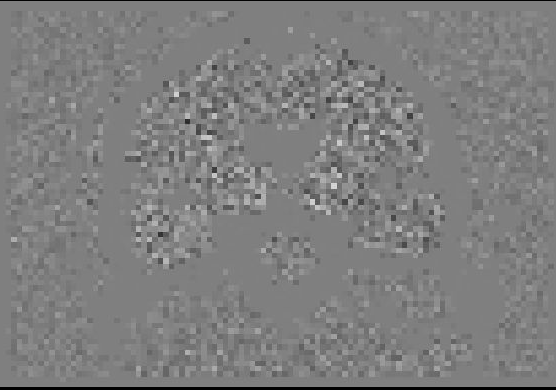
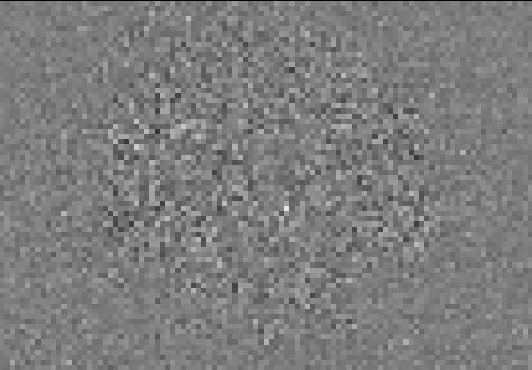
res\_sub-ADPRC001F0\_acq\_data\_dwi.mif 🡪 residual data

The noise.mif file is the estimated spatially-varying noise level. Eyeball your residuals as part of quality control; compare raw data with the denoised data. If denoising did a good job, there should be little or no anatomy in the residual maps or the diffusion weighted images (this does not apply to the initial B0 – denoising should apply more to the higher shells). The lack of anatomy in the residual maps is a marker of accuracy and signal-preservation during denoising. Ideally, these residuals are Gaussian distributed with zero-mean and contain no anatomical structure.

Elapsed time: 2 min 20 sec

raw denoised

noise residual (b0) residual (b2000)



Checking your residuals is good for quality control!

<https://community.mrtrix.org/t/dwidenoise-residuals-map/1293/2>

1. **Gibb’s ringing**

Corrections for various artefacts, such as table vibration, radio-frequency-based distortions, and incorrect magnetic field gradient calibration, which can significantly degrade diffusion data. The Gibbs ringing artefact appears in k-space truncation along finite image sampling, and it can be suppressed by post hoc methods.

This method is based on local sub-voxel shifts. The strength of ringing (caused by truncation of Fourier transform space) is dependent on the location of the edge relative to the sampling grid. Method works by finding optimal sub-voxel shift in the neighborhood of shard edges.

*Kellner et al., (2016). Gibbs-ringing artefact removal based on local subvoxel shifts*

<https://community.mrtrix.org/t/processing-of-volumes-prior-to-dwipreproc/2112>

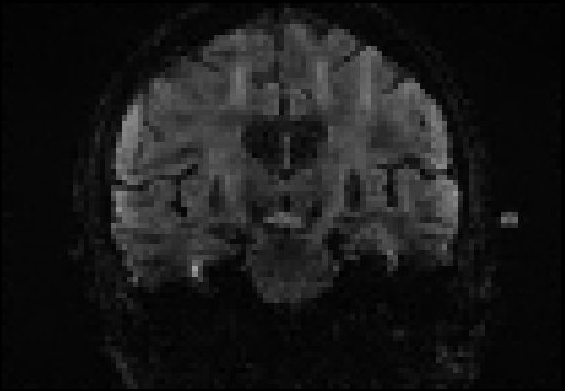
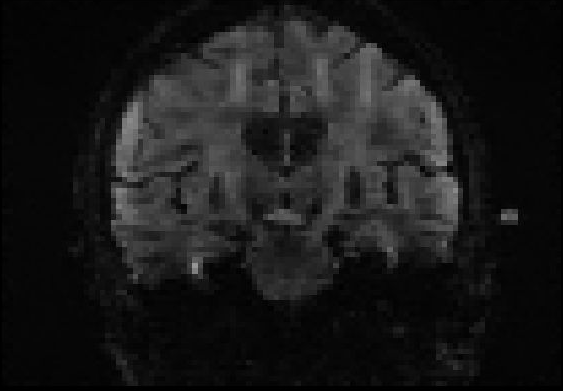
%Sample input: dsub-ADPRC001F0\_acq\_data\_dwi.mif

Command: mrdegibbs Like noise correction, perform Gibbs ringing correction on

the entire dataset (main dwi, BUs and BDs).

%Sample output: gdsub-ADPRC001F0\_acq\_data\_dwi.mif

Elapsed time: 20 sec

Denoised Gibbs corrected

**Edit gradient files:**

After the first 2 steps (denoising, gibbs) have been completed, we need to separate the data back out into the main dwi dataset and the p-a (BDs) to prepare for the next step, topup.

%Sample input: gdsub-ADPRC001F0\_acq\_data\_dwi.mif

Command: mrconvert

%Sample output: cgdsub-ADPRC001F0\_acq\_data\_dwi.mif 🡪 main dwi

PA\_sub-ADPRC001F0\_acq\_data\_dwi.mif 🡪 p-a (BDs); can be up to 3 vols

Then, for the dprc data specifically, we must edit the .bvec and .bval gradient files to add in the last B0 (BU) file (volume 106). The assumption is that the last BU will have the same parameters as the other BUs embedded in the main dwi dataset. These values (both in the .bval and .bvec files) should typically be around 0, and I have created a function for the last BU file to copy the first BU values as its own.

* Shell = firstB0, ~0 (.bval file)
* Vectors = ~0, ~0, ~0 (.bvec file)

Add the modified gradient files, now with the extra B0/BU with the main dwi data and to its header file.

%Sample input: PAR\_NAME = participant ID, e.g. ‘sub-ADPRC001F0’

datafile = dwi datafile, e.g. ‘\_acq\_data\_dwi’

Command: **LastB0AddOn.m** In-house function to edit gradient files

mrconvert with -fslgrad option Adds modified gradient value to the header

file of the data

%Sample output: bcgdsub-ADPRC001F0\_acq\_data\_dwi.mif

1. **Topup**

Data acquisition is based on echo planar imaging (EPI), so must correct for distortions from magnetic field inhomogeneity. Reversed phase encoding method is used. Utilises TOPUP function from FSL, called through MRtrix. Opposite phase encoding directions for non dwi are posterior-anterior (PA).

We are preparing the appropriate B0s for topup/eddy. Note that MRtrix only accepts the number amount of a-p (BUs) and p-a (BDs) volumes (therefore, an even number in total) for this correction. However, again, note that not all participants have the same number of BD files – most have 3, but some have only 1 or 2, so the script accounts for this. There are also 6 BUs to choose from. My script implements the UKB ‘best B0’ (see below), with a function I created.

%Sample input: bcgdsub-ADPRC001F0\_acq\_data\_dwi.mif

Command: dwiextract extract all B0s from the dataset

mrcat place a-p and p-a images into one file

%Sample output: AP\_sub-ADPRC001F0\_acq\_data\_dwi.mif 🡪 all a-p (BUs); 6 vols total

allB0s\_ sub-ADPRC001F0\_acq\_data\_dwi.mif 🡪 all a-p and p-a in one fie

1. **BestB0 pair selection:**

One best B0 pair (a-p, p-a) is selected and combined together for topup correction. First, all of the B0 images in both phase encoding directions (AP and PA) are aligned with one another with a rigid-body registration with 6 degrees of freedom (dof) (using fsl's FLIRT tool).

Next, the correlation is calculated between each of the b0 images to all of the others (fsl's FSLCC tool). Typically, the first BU and BD file are selected, and if their correlation is 0.98 or greater (Jesper's criteria), then we use that. The function checks that at least one of the B0 images has a cross-correlation greater than 0.95 - if not, then that participant is flagged, but will still use the highest (and earliest, if tied) cross-correlation B0 for it. If the first B0 has sufficient quality (a correlation of 0.98 or higher - Jesper's criterion), we would select this as the 'best B0 image' to use. If this is not the case, then the second B0 is checked, and so on, so forth. If none of the B0s have a higher correlation than 0.98, but there is at least one B0 greater than 0.95, then we would select for the highest correlation B0 (and earliest, if tied) for it.

A text file (BestB0.txt) is generated which will show the participant ID, the B0 status, and the BU and BD number used in the sequence.

In addition, the best BU volume, which is selected, will now be the first volume in the dwi sequence - the first volume and selected volume will switch places (if this is done). The gradient files (.bval and .bvec) should also be edited for this accordingly (in another function called **GradientEdit\_forBestB0.m**). As suggested by Jesper, in a fsl community post: <https://www.jiscmail.ac.uk/cgi-bin/webadmin?A2=ind1703&L=FSL&P=R62904>

And confirmed from Jesper: <https://www.jiscmail.ac.uk/cgi-bin/wa-jisc.exe?A2=ind2007&L=FSL&O=D&X=65845096BD798100FD&Y=ltah262%40aucklanduni.ac.nz&P=94911>

\*Note that any time you edit the gradient files (.bval and .bvec), you will need to re-import and provide the gradient files to the diffusion files by using the -fslgrad option. My script accounts for this. Only do this if the BU used is not the first one in the diffusion sequence.

%Sample input: PAR\_NAME = participant ID, e.g. ‘sub-ADPRC001F0’

datafile = dwi datafile, e.g. ‘\_acq\_data\_dwi’

NumBDs = number of BD files participant has (typically 3)

startdir = start directory that you defined in the script - where the data will be stored.

Command: **BestB0.m** In-house function to carry select bestB0 pair

mrconvert Convert .mif files to .nii files for fsl processing

flirt Align all a-p and p-a images with rigid-body registration with 6 dof

fslcc Run cross-correlation with each volume and the registered image

%Sample output: bbcgdsub-ADPRC001F0\_acq\_data\_dwi.mif 🡪 dwi sequence with best a-p

as the first volume

TUB0s\_sub-ADPRC001F0\_acq\_data\_dwi.mif 🡪 best B0 pair (1 a-p & 1 p-a) in one file

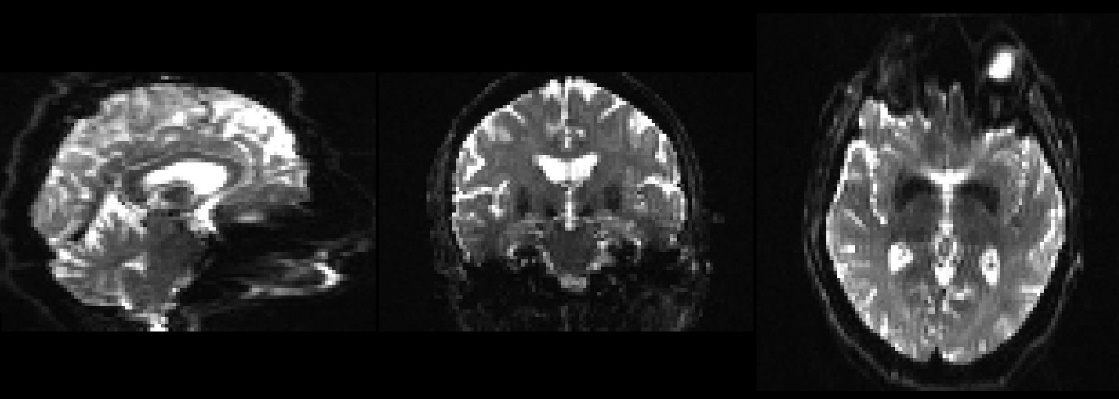
BestB0.txt 🡪 text file detailing the B0 status and B0s used

*Andersson & Sotiropoulos, (2016). An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging*

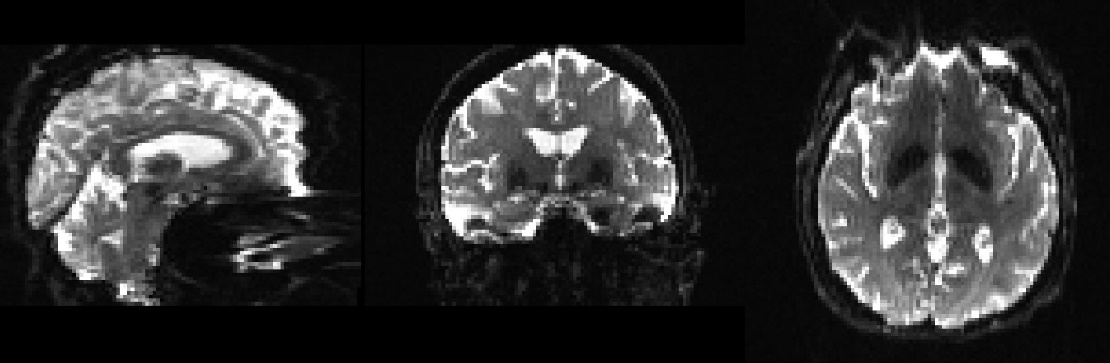
*Alfaro-Almagro, F., Jenkinson, M., Bangerter, N. K., Andersson, J. L. R., Griffanti, L., Douaud, G., … Smith, S. M. (2018). Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank.*

Elapsed time: \_\_\_\_ sec

*fsleyes contrast setting for images below: min 0, max 5000*



BU (blip-up)



BD (blip-down)

<https://community.mrtrix.org/t/dwipreproc-register-fieldmap-to-dwi-other-questions/621>

reversed phase encoding input options:

<https://community.mrtrix.org/t/dwipreproc-option-2/1308>

dwifslpreproc only accepts and equal number of a-p and p-a volumes as inputs:

<https://community.mrtrix.org/t/dwipreproc-rpe-pair-why-equal-number-of-up-down-b0-encodings/957>

1. **Eddy Currents**

Eddy and TOPUP work together to correct for distortions which appeared due to eddy currents, such as head motion and susceptibility originated artefacts. Also, for the rapid gradient field changes. This will make predictions about how the signal should look and uses ‘error signals’ to estimate and correct for the EC-induced field and subject movement.

Run FSL's eddy to correct for eddy currents and subject motion. Eddy will take the inputs from TOPUP and apply correction to all dwi images.

*Andersson & Sotiropoulos (2016) . An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging*

All DWI volumes are acquired with precisely the same phase encoding direction and EPI readout time. In addition, one or more pairs of spin-echo b=0 EPI volumes are provided, where half of these volumes have the same phase encoding direction and readout time as the DWIs, and the other half have precisely the opposite phase encoding direction (but the same readout time). These additional images are therefore used to estimate the inhomogeneity field, but do not form part of the output DWI series.

\*note that the whole process of topup + eddy current correction is also known as **geometric distortion correction (GDC)**.

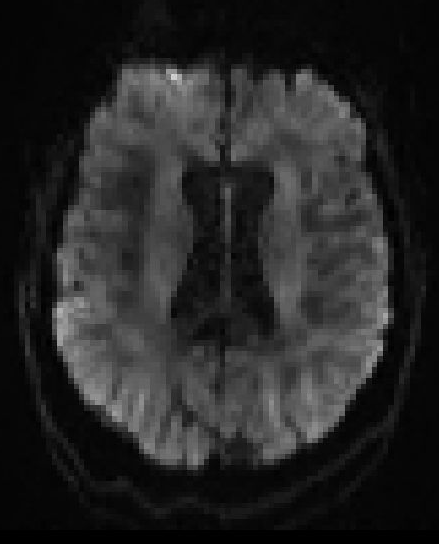
%Sample input: bbcgdsub-ADPRC001F0\_acq\_data\_dwi.mif 🡪 main dataset (dwi)

TUB0s\_sub-ADPRC001F0\_acq\_data\_dwi.mif 🡪 best B0 pair (1 a-p & 1 p-a) in one file

Command: dwifslpreproc run topup, eddy (w/ reversed phase encoding and -repol flag

option)

%Sample output: ebbcgdsub-ADPRC001F0\_acq\_data\_dwi.mif

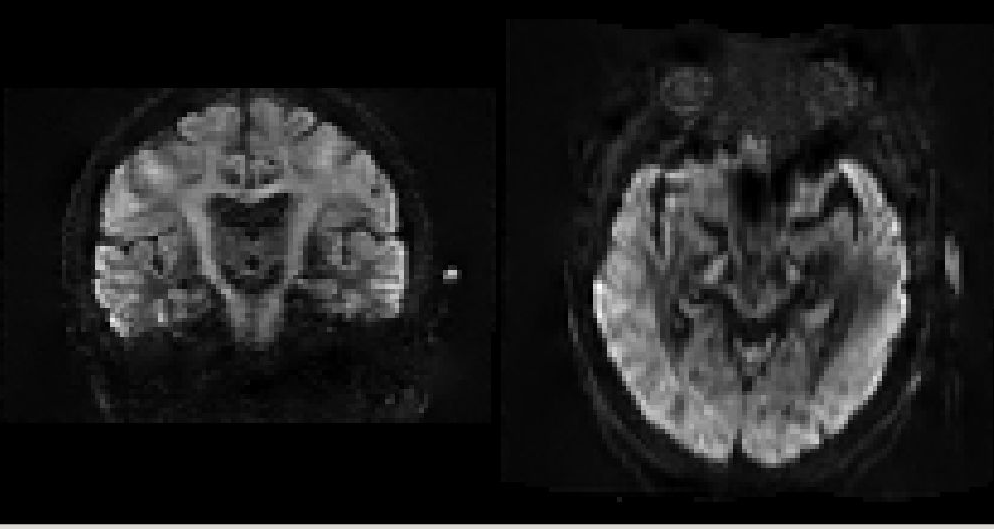
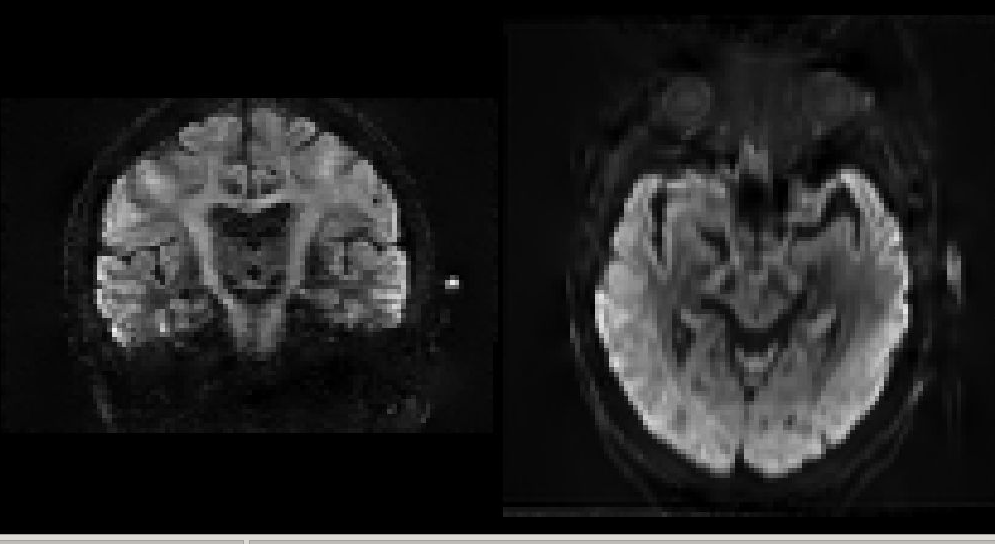
Gibbs ringing (pre-eddy) topup/eddy corrected

Fsl user guide wiki for using additional correction options in eddy: <https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/eddy/UsersGuide#WARNING_this_page_is_being_edited_in_preparation_of_a_new_release_and_may_be_in_an_inconsistent_state>

<https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/eddy#What_is_new_in_6.0.1.3F>

Correction options are: --repol –ol\_nstd=4

The --repol flag removes any slices deemed as outliers and replace them with predictions made by the Gaussian Process. 4 standard deviations has been chosen because it is a good compromise between type 1 and 2 errors for a "standard" data set of 50-100 directions.

No repol option With repol option on

Elapsed time: 1 hour 32 min

1. **Estimate a brain mask**

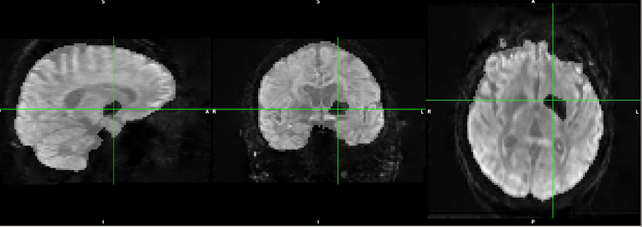
Restricts analysis only to relevant voxels – i.e. keeps brain matter and removes non-brain matter (skull). A whole-brain mask is required for most analysis, and as an input to the subsequent bias field correction step and the global normalisation step. An initial (first-pass) brain mask is first generated, and then bias field correction (next step) is ran, and then another brain mask is generated from that. This iterative process helps generate a more accurate brain mask. So, this first mask gives you an imperfect, but ok-ish mask.

%Sample input: ebbcgdsub-ADPRC001F0\_acq\_data\_dwi.mif

Command: dwi2mask

%Sample output: initial\_mask\_sub-ADPRC001F0\_acq\_data\_dwi.mif

Elapsed time: 12 sec



It is crucial to check that your brain masks for your participants are alright. Here, we can see that there is a ‘hole’ in this brain mask, where the pointer is. Understandably, since this is the first pass. So, to fix this, we will run the next step (bias field correction) and then generate another brain mask from that.

<https://community.mrtrix.org/t/dwi2mask-holes-in-mask-images/484/13>

Note that BET (from fsl) and dwi2mask (from mrtrix) are different - you can read more about this hear from a community forum: <https://community.mrtrix.org/t/dwi2mask-holes-in-mask-images/484/13>. The gist is that they are both completely different algorithms, but are both suitable methods for brain extraction; however, dwi2mask is more suitable for diffusion data.

1. **Bias Field correction**

To correct for field inhomogeneity caused by MR images possessing a low frequency intensity shift appearing as intensity inhomogeneity over the image. Utilises **ANTs** source code (N4BiasFieldCorrection), employed through MRtrix3.

*Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., & Gee, J. C. (2010). N4ITK: Improved N3 Bias Correction.*

%Sample input: ebbcgdsub-ADPRC001F0\_acq\_data\_dwi.mif 🡪 main dataset (dwi)

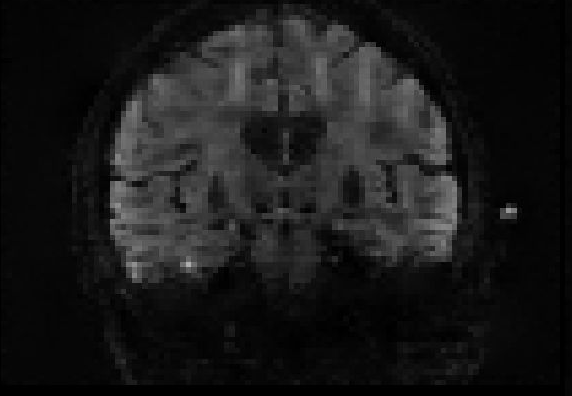
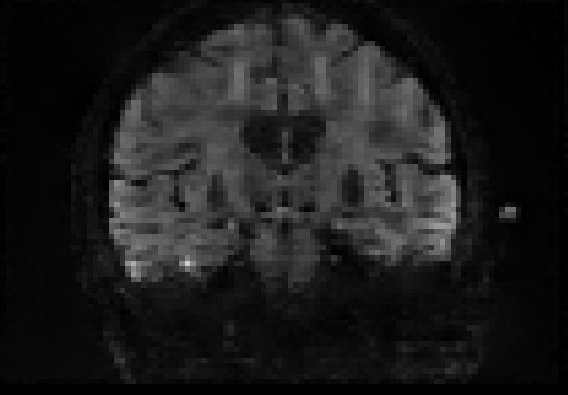
initial\_mask\_sub-ADPRC001F0\_acq\_data\_dwi.mif

Command: dwibiascorrect ants Correction is applied to the B0 images. Then the

corrected B0 images get applied to the rest of the dwi images

%Sample output: febbcgdsub-ADPRC001F0\_acq\_data\_dwi.mif

Elapsed time: 17 sec

  eddy corrected inhomogeneity field corrected

1. **Generate second brain mask**

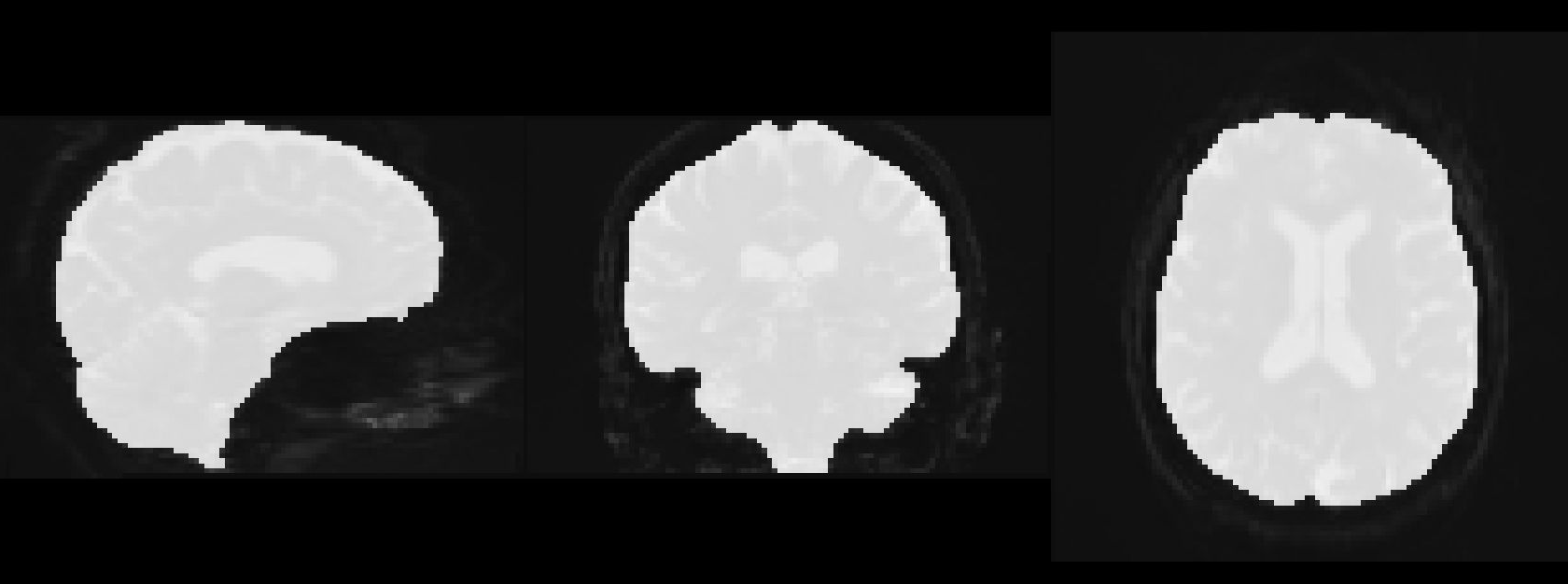
Now, generate a brain mask after having run this step. Should hopefully, look better (more full) and capturing the brain parts. Do check this with every participant.

%Sample input: febbcgdsub-ADPRC001F0\_acq\_data\_dwi.mif

Command: dwi2mask -clean\_scale value = 3

I made the brain threshold a bit more conservative by setting the value to 3 (default in mrtrix is 2). This is because after examining the first 10 participants much non-brain parts were included, so this should help a bit. I tested this threshold out, and it worked out okay, so this will be the default for all participants.

%Sample output: brain\_mask\_sub-ADPRC001F0\_acq\_data\_dwi.mif



New brain mask generated from after running bias field correction step.

Elapsed time: 12 sec

Check your brain masks! Make sure it covers the ‘brain parts’ only.



You need to check your final brain masks before proceeding onto further analysis. There should be no ‘holes’ in your brain mask, and it should generally cover all brain parts.

Note that it is okay if the brain mask is covering non-brain parts (e.g. eyeballs, sinus regions, etc). It is more problematic if there are holes in your brain mask, and you should go back and edit this in manually. Two posts from the MRtrix forum state this:

<https://community.mrtrix.org/t/problem-with-dwi2mask-result/3036>

<https://community.mrtrix.org/t/dwi2mask-creates-mask-with-insufficient-overage/3766/3>

\*I have noted in a text file (manual\_brain\_masks.txt) which require manual input to fill in the ‘holes of the brain mask.’

**References**

Alfaro-Almagro, F., Jenkinson, M., Bangerter, N. K., Andersson, J. L. R., Griffanti, L., Douaud, G., … Smith, S. M. (2018). Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank. NeuroImage, 166(April 2017), 400–424. https://doi.org/10.1016/j.neuroimage.2017.10.034

Andersson, J. L. R., & Sotiropoulos, S. N. (2016). An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. NeuroImage, 125, 1063–1078. https://doi.org/10.1016/j.neuroimage.2015.10.019

Kellner, E., Dhital, B., Kiselev, V. G., & Reisert, M. (2016). Gibbs-ringing artifact removal based on local subvoxel-shifts. Magnetic Resonance in Medicine, 76(5), 1574–1581. https://doi.org/10.1002/mrm.26054

Perona, P., & Malik, J. (1990). Scale-Space and Edge Detection Using Anisotropic Diffusion. IEEE Transactions on Pattern Analysis and Machine Intelligence, 12(7), 629–639. https://doi.org/10.1109/34.56205

Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., & Gee, J. C. (2010). N4ITK: Improved N3 Bias Correction. IEEE Transactions on Medical Imaging, 29(6), 1310–1320. https://doi.org/:10.1109/TMI.2010.2046908

Vellmer, S., Tonoyan, A. S., Suter, D., Pronin, I. N., & Maximov, I. I. (2018). Validation of DWI pre-processing procedures for reliable differentiation between human brain gliomas. Zeitschrift Fur Medizinische Physik, 28(1), 14–24. https://doi.org/10.1016/j.zemedi.2017.04.005

Veraart, J., Fieremans, E., & Novikov, D. S. (2016). Diffusion MRI noise mapping using random matrix theory. Magnetic Resonance in Medicine, 76(5), 1582–1593. https://doi.org/10.1002/mrm.26059.

Veraart, J., Novikov, D. S., Christiaens, D., Ades-Aron, B., Sijbers, J., & Fieremans, E. (2016). Denoising of diffusion MRI using random matrix theory. NeuroImage, 142, 394–406. https://doi.org/10.1016/j.neuroimage.2016.08.016

Veraart, J., Sijbers, J., Sunaert, S., Leemans, A., & Jeurissen, B. (2013). Weighted linear least squares estimation of diffusion MRI parameters: Strengths, limitations, and pitfalls. *NeuroImage*, *81*, 335–346. https://doi.org/10.1016/j.neuroimage.2013.05.028