**GOGNESTIC 2023: Voxel Based Morphometry (VBM) analysis with FSL**

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

[**https://fsl.fmrib.ox.ac.uk/fslcourse/2019\_Beijing/lectures/Structural/seg\_struct.html**](https://fsl.fmrib.ox.ac.uk/fslcourse/2019_Beijing/lectures/Structural/seg_struct.html)

[**https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FSLVBM/UserGuide**](https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FSLVBM/UserGuide)

**FSL-VBM**

In this session we look at a small ageing study comparing two groups (data\_2groups, young and older healthy controls) for local differences in grey matter volume, using FSL-VBM. We will also look at another ageing dataset (data\_Age) to see how we can use FSL-VBM to look for correlations between grey matter volume and a continuous measure (age, cognitive test score, etc). Most of the steps have already been carried out, as there isn't enough time in this session to run all of the registrations required to carry out a full analysis from scratch, but you can re-run these later in your own time.

**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 a subset of the CamCAN dataset (<https://www.cam-can.org/>). To download please read the user data agreement you can find here <https://imaging.mrc-cbu.cam.ac.uk/methods/COGNESTIC2022> and email [marta.correia@mrc-cbu.cam.ac.uk](mailto:marta.correia@mrc-cbu.cam.ac.uk) to confirm you accept the terms and conditions.

All of the steps described below have been implemented in the script **FSLVBM\_cognestic\_all.sh** which you can download from <https://imaging.mrc-cbu.cam.ac.uk/methods/COGNESTIC2023>.

# **FSL-VBM Pipeline - Overview**

Running FSL-VBM involves a few simple steps:

* prepare your T1-weighted images in the right format
* fslvbm\_1\_bet - carry out brain extraction on all T1 images
* fslvbm\_2\_template - create the study-specific symmetric grey matter template
* fslvbm\_3\_proc - register all the grey matter images to the template, modulate and smooth them with different kernel sizes and finally runs an initial GLM analysis for qualitative evaluation
* randomise - carry out voxelwise GLM analysis using permutation testing

## A - Prepare your data for the FSL-VBM study

a) Place all your T1-weighted data in your FSL-VBM directory. Start by creating a new directory, for example:

mkdir FSLVBM

Then copy into your FSL-VBM directory all of your subjects' T1 images, giving each subject's T1 image a different name, preferably with a prefix corresponding to each group, for example:

OLD\_CBU110220\_MPRAGE.nii

OLD\_CBU110519\_MPRAGE.nii

OLD\_CBU110564\_MPRAGE.nii

OLD\_CBU110653\_MPRAGE.nii

YNG\_CBU110468\_MPRAGE.nii

YNG\_CBU110547\_MPRAGE.nii

YNG\_CBU110752\_MPRAGE.nii

YNG\_CBU110799\_MPRAGE.nii

b) Select subjects to create a study-specific template. If you have more than one group and the number of subjects in each is not the same, choose (at random) among the biggest group(s) the images that you will use to create the study-specific template, with the same number as of the smallest group (in order to create an unbiased template - see below for further explanation). Once you've chosen which T1 images to keep to build the template, put all the selected names of exams in a file called template\_list in your FSL-VBM directory.

All your different populations included in this study MUST be represented in the template construction.

If you are using all datasets, simply run the following:

for f in \*.nii

do

echo $f >> template\_list

done

c) At this point you should have a quick look at all your data to check that all subjects' structural images are what you expected:

slicesdir `imglob \*`

The imglob command lists all of your input images. The slicesdir command takes the list of images and creates a simple web-page containing snapshots for each of the images. Once it has finished running it tells you the name of the web page to open in your web browser, to view the snapshots. Have a careful look.

d) It's a good idea to consider your cross-subject statistical model **before** you run the FSL-VBM analysis. So you should at this point create your design.mat and design.con in your FSL-VBM directory; see some FSL GLM examples here <https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/GLM>. For simple experimental design with just two groups you can also use the script design\_ttest2:

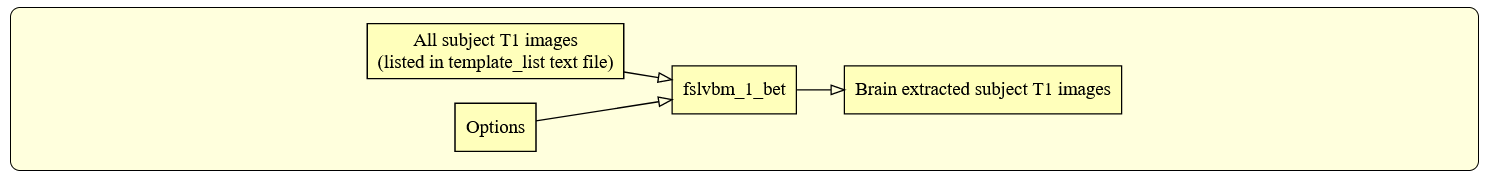
design\_ttest2 design n1 n2 -m

where n1 and n2 represent the sizes of each group. The ‘-m’ flag indicates that the design matrix will include two rows for individual group means.

**WARNING!!!** The order of the rows in your design.mat model MUST match the order of your images when doing an 'imglob \*' command in your FSL-VBM directory.

## B - Extracting brain information: fslvbm\_1\_bet

The first FSL-VBM script moves all your input images into a new struc subdirectory (and adding "\_struc" to the end of each filename). It then runs brain extraction on the images. You can either use the -b option to get default BET behaviour, or use the -N option if your images include a lot of neck (which most of the time confounds the BET preprocessing).



To run this first script, just type:

fslvbm\_1\_bet -b

or

fslvbm\_1\_bet -N

in your FSL-VBM directory.

At the end of this step, it is once again worth **CHECKING** the brain images (\*\_brain.\*) in your struc directory by loading the new slicesdir output into a web browser. Brain extraction is the step which is the most likely to need tweaking in the FSL-VBM protocol. It might not be too much of an issue if you get "more" than the grey matter (eyes, dura etc.) though this will need careful checking before running your statistics. If you do not get good results with either option (i.e., if some images are missing some grey matter), you can try adding other [bet](https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/BET) options after the -b or -N option.

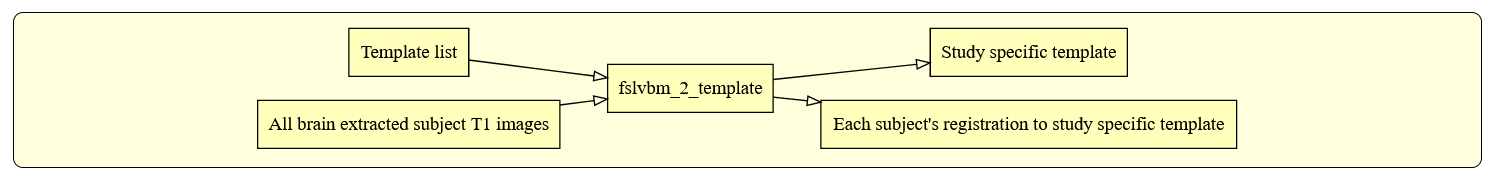
If you later want to add more subjects to your analysis then just put the new subjects' images inside the top level directory (e.g. FSLVBM) and re-run fslvbm\_1\_bet. Don't forget to update template\_list if necessary.

## C - Creating the template: fslvbm\_2\_template

The second step of the FSL-VBM protocol creates the study-specific grey matter (GM) template.

First, all brain-extracted images are segmented into GM, WM and CSF. Then, GM images selected in the template\_list file (\*\_struc\_GM) are affine-registered to the GM ICBM-152 template, concatenated and averaged. This averaged image is then flipped along the x-axis and the two mirror images then re-averaged to obtain a first-pass, study-specific "affine" GM template ("template\_GM\_init"). Second, the template\_list GM images are re-registered to this "affine" GM template using non-linear registration, concatenated into a 4D image called "template\_4D\_GM", averaged, flipped along the x-axis. Both mirror images are then averaged to create the final symmetric, study-specific "non-linear" GM template at 2x2x2mm3 resolution in standard space.

If you have different populations, they should all be represented in your template. You should use the same number of subjects from each in the construction of the study-specific template. This is to avoid any bias during the registration step that would have consisted in favouring one of the groups. For example, if you have only controls in your template, or more controls than patients, it is likely that the non-linear registration would be more accurate for your control subjects than for your patients. Then you cannot distinguish, in your results showing differences in the GM volume distribution between the two groups, what is actually disease-related from what is registration-related!



For this step, you have two options: either you want to create a template based on an affine registration (-a option) of GM images to the GM ICBM-152 template, or on a non-linear registration (-n option).

So to run this second step script, just type:

fslvbm\_2\_template -a

or

fslvbm\_2\_template -n

in your FSL-VBM directory.

Once this is completed, CHECK the "template\_GM\_4D" image in struc with the movie loop in fsleyes.

fsleyes struc/template\_4D\_GM

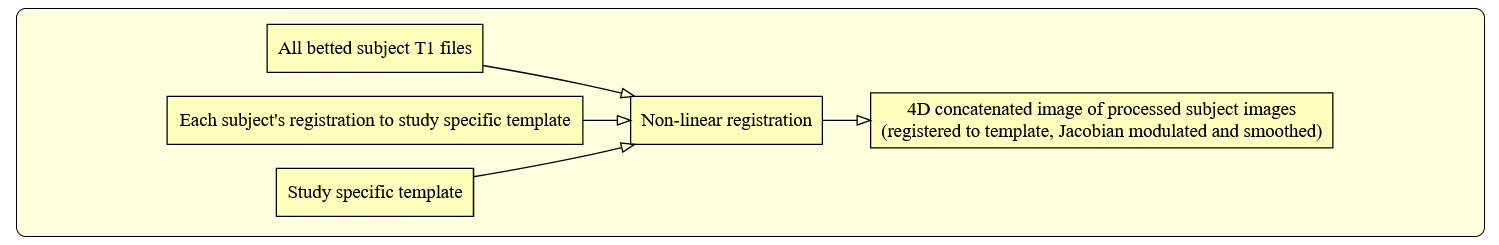


## D - Processing the native GM images: fslvbm\_3\_proc

The final script will non-linearly register all your GM images to the study-specific template and concatenate them into a 4D image ("GM\_merg") in the stats directory in your working FSL-VBM directory. The FSL-VBM protocol also introduces a compensation (or "modulation") for the contraction/enlargement due to the non-linear component of the transformation: each voxel of each registered grey matter image is multiplied by the Jacobian of the warp field (see Good et al., 2001). All the modulated registered GM images are concatenated into a 4D image in the stats directory ("GM\_mod\_merg") and then smoothed ("GM\_mod\_merg\_s3" for instance) by a range of Gaussian kernels; sigma = 2, 3, 4mm, i.e., approximately from FWHM = 2x2.3 = 4.6mm to FWHM = 9mm.

Finally, this last step gets everything ready for you to run permutation-based non-parametric inference using the design.mat and design.con which you supplied, a mask of the GM ("GM\_mask") and the 4D multi-subject concatenated processed data (e.g. "GM\_mod\_merg\_s3"). The script runs randomise with inference (generation of p-value maps) turned off, so that it very quickly creates just the raw tstat maps. These tstats maps should help you decide which smoothing is the most relevant to feed into a full run of randomise, and which threshold to use for the cluster-based thresholding (option -c in the [randomise command](https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Randomise)); however, in general we would recommend using the TFCE option (-T) instead of the cluster-based thresholding.

**WARNING!!!** By default fslvbm\_3\_proc concatenates the images in alphabetical order (following the names that they started with); make sure this matches the subject ordering assumed in your design.mat model.



All of the above is done simply by running the script:

fslvbm\_3\_proc

in your FSL-VBM directory.

Please do not forget the final **CHECK** of the 4D image of modulated registered GM images "GM\_mod\_merg" using the movie loop in fsleyes.

fsleyes stats/GM\_mod\_merg



## E - Obtaining and displaying your FSL-VBM results

We strongly recommend using randomise (permutation testing) for inference in VBM-style analysis and not Gaussian random field theory (GRF), as the approximations underlying the latter are not generally appropriate in such analyses.

### **Running randomise and displaying TFCE-based thresholding results**

Choose the most appropriate smoothing (e.g., sigma=3mm) for the TFCE-based analysis. If you want to apply a different smoothing than already applied, you can do so (e.g., sigma=3.5mm) with:

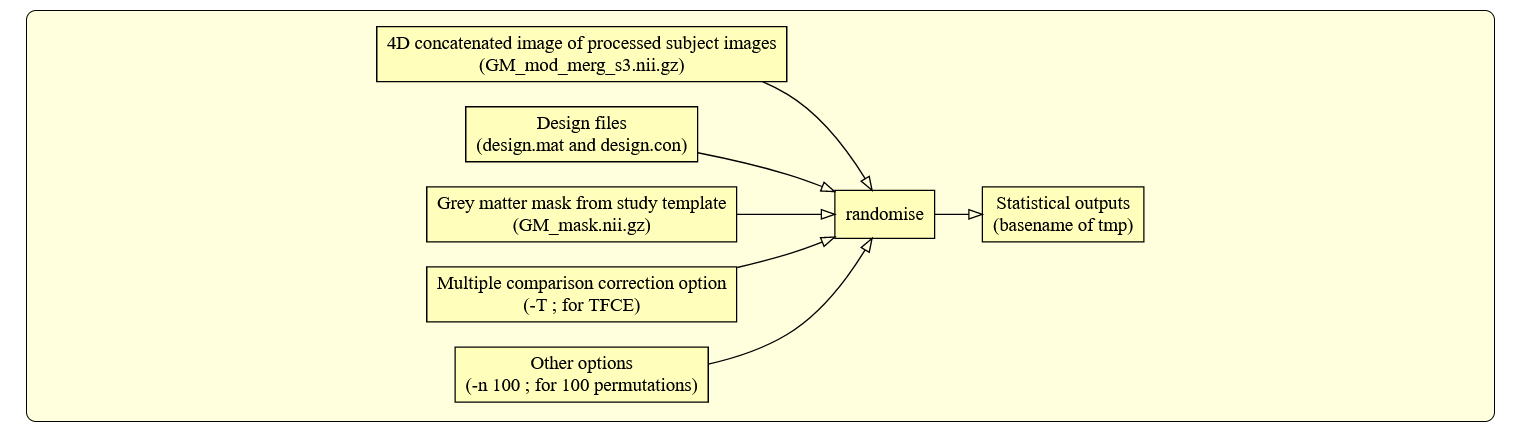
fslmaths GM\_mod\_merg -s 3.5 GM\_mod\_merg\_s3.5

Having chosen the most appropriate smoothing (e.g. sigma = 3mm), run randomise (see [randomise usage](https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Randomise)), for instance:

cd stats

randomise -i GM\_mod\_merg\_s3 -m GM\_mask -o fslvbm -d design.mat -t design.con -T

-n 5000

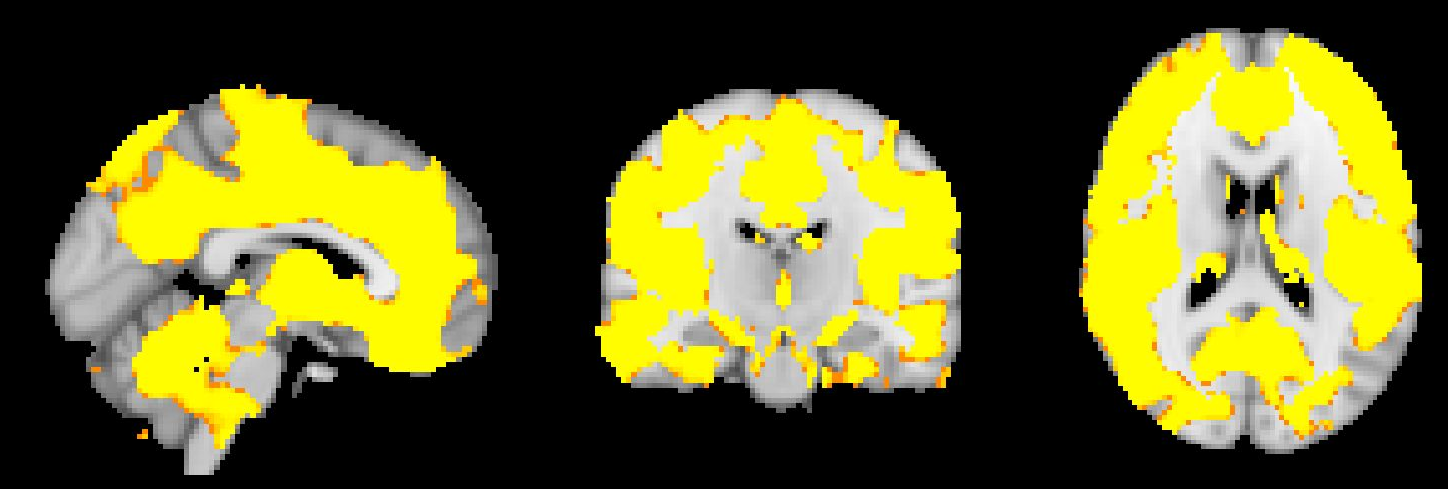


You can then view the (1-p) corrected p-value images in FSLEyes:

fsleyes $FSLDIR/data/standard/MNI152\_T1\_2mm fslvbm\_tfce\_corrp\_tstat1 –cm red-yellow –dr 0.95 1

### 

fsleyes $FSLDIR/data/standard/MNI152\_T1\_2mm fslvbm\_tfce\_corrp\_tstat2 –cm red-yellow –dr 0.95 1



**Correlation between GM density and Age**

For this analysis we will use the second dataset you downloaded, data\_Age. All of the steps above (A to E) will be repeated for this dataset, we will just need a new design matrix and contrasts file. For this analysis we will include one covariate of interest (Age) and one nuisance variable (sex), and therefore the design\_ttest2 script will no longer be useful to create the design.mat and design.con files. Instead we will create this files by hand. Detailed instructions on how to do this can be found here:

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

There are three steps involved:

1. Enter the data for your design matrix, contrasts, or F-tests, using your data-editing tool of choice (e.g. Microsoft Excel, Google docs, Notepad).
2. Save your data as a plain text file, with columns separated by spaces or tabs - in Excel, saving your file as **Text (Tab delimited)** should do the trick.
3. Use the Text2Vest tool, to convert the data into the format used by FSL.

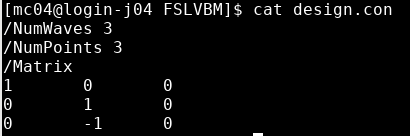
Steps 1 and 2 have already been performed for you, and you can find the text files together with the data you downloaded: design\_Age.txt and contrasts\_Age.txt. So all we have to do is run the following commands:

Text2Vest design\_Age.txt design.mat

Text2Vest contrasts\_Age.txt design.con

If you look at the design.mat file you can see we have three explanatory variables: one for the group (all 1s because this time the data is being analysed as a single group), one for demeaned age, and one for gender (also demeaned).

And the contrast file looks like this:



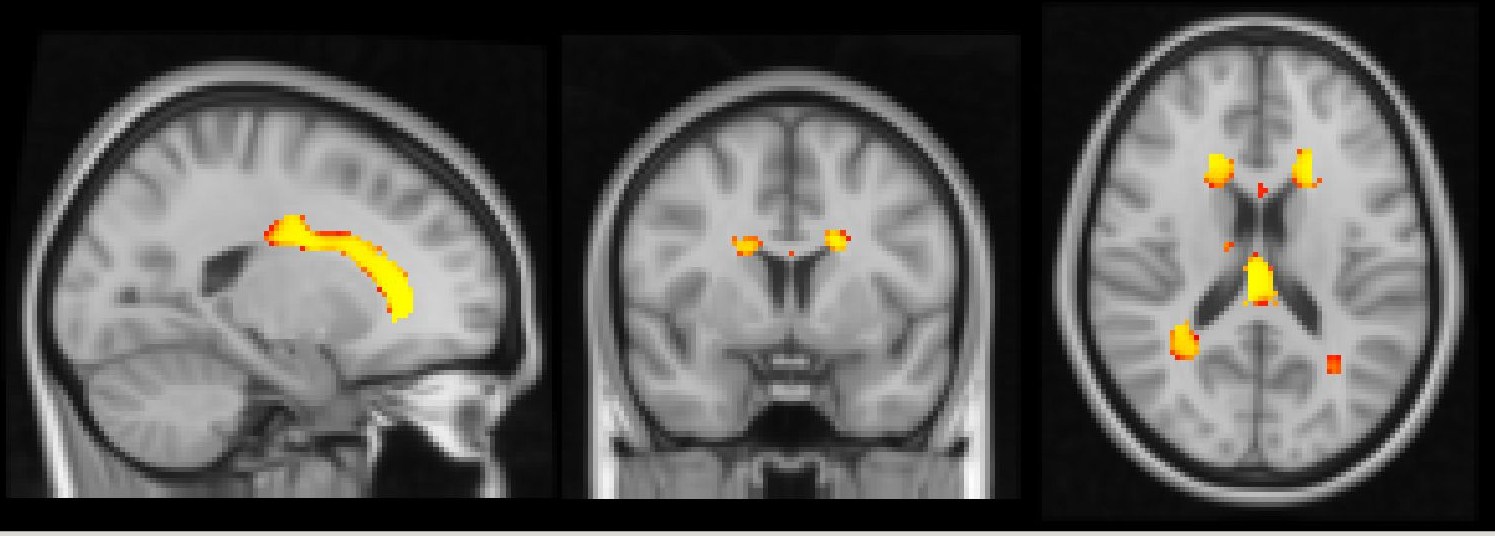
Which means step E in the analysis pipeline will generate three statistical maps:

* tstat1: group mean
* tstat2: positive correlation with age
* tstat3: negative correlation with age

When step E has finished you can look at the results using fsleyes.

Positive correlation with age:

fsleyes $FSLDIR/data/standard/MNI152\_T1\_1mm fslvbm\_tfce\_corrp\_tstat2 –cm red-yellow –dr 0.95 1



Negative correlation with age:

fsleyes $FSLDIR/data/standard/MNI152\_T1\_1mm fslvbm\_tfce\_corrp\_tstat3 –cm blue-lightblue –dr 0.95 1

