**Chapter 5 FreeSurfer Imaging Analysis Procedure**

For more information on analysis procedures see: <https://surfer.nmr.mgh.harvard.edu/fswiki/FsTutorial/GroupAnalysis>

This can be run for both correlational and group analyses (i.e., ANOVA or ANCOVA)

1. **MRI DATA PROCESSING PROCEDURE**
   1. All files are assumed to be NifTI files.
      1. If files need to be converted from PAR/REC to NifTI, navigate to the mricron directory open **dcm2niigui**. Ensure the output is compressed FSL (4D NifTI nii) and then drag and drop data.REC file into the window. Repeat for all participants
   2. Set subjects directory using the following commands:
      1. **$ tcsh**
      2. **$ setenv SUBJECTS\_DIR “*Subjects directory”***
   3. Convert NifTI to mgz file- using the following command:
      1. **mri\_convert “*subject folder/input nii.gz file*” *“subject folder/ output .mgz file”***
      2. This can also be scripted for all participants in the subject directory using the following Bash command

# This script requires that the respective subject's ".nii.gz" file is in their subject folder.

# Example: Subject 001.MR1 in the PARTICIPANT folder must have a file ending with .nii.gz in the directory /usr/freesurfer/subjects/PARTICIPANT/001.MR1/

# How to use: Place this file in the project folder (e.g. /usr/freesurfer/subjects/PARTICIPANT

# run using:

# bash FS\_mriConvert.sh

# add or edit list of participants as needed

for subj in 001.MR1 001.MR2 002.MR1 002.MR2

do

for file in ./${subj}/\*.nii.gz

do

mri\_convert ./${subj}/${file} ./${subj}/mri/orig/001.mgz

done

done

\*Note: Ensure that the output file of MRI convert is in the MRI/orig folder for the participant

\*This will place a 001.mgz file in the mri/orig directory for each participant

* 1. Next, perform recon-all procedure which will provide the fully automated analysis of structural data. The command line should be:
     1. **$ “recon-all -s \*subject\* -all”** where \*subject\* is the name of their directory contained in the subject folder.
     2. This can also be automated for all participants using the following commands in Bash:

# Place this in the project folder of the respective project you wish to edit.

# Example: place this file in the /usr/freesurfer/subjects/PARTICIPANT/

# Navigate to the Project folder and run using:

# bash FS\_reconAll.sh

# Make sure that the $SUBJECTS\_DIR is specified to the project that you are working on

# Add participants as needed. List out ALL participants and timepoints

for subj in 001.MR1 001.MR2 002.MR1 002.MR2

do

recon-all -s ${subj} -all -openmp 8

done

* + 1. Recon-all on a regular system can take anywhere between 12-24 hours. Because of this, a simple project with 60 participants (small amount of data) can take up to 2 months, not including any troubleshooting or corrections.
    2. The “**openmp**” option can replace this and is faster than processing each subject by itself. In CentOS, using a single Terminal with a recon-all list utilizing the openMP flag will use all cores available. For a computer to use the ‘openMP’ flag, determine the number of cores that the PC has and multiply it by two.

1. **SMOOTHING MRI IMAGES**
   1. mri\_glmfit needs each subject to have pre-computed smoothed data for the target surface (fsaverage is the default) for each measure (thickness, sulc, area, curv, etc.). Your SUBJECTS\_DIR should contain either a link or a copy of the **'fsaverage'** subject found in your subjects directory. Presmoothing the data onto the target surface is not part of the normal recon processing stream, but you can easily create this data with recon-all, using the command:
      1. **recon-all -s <subjid> -qcache**
      2. This can also be automated using the following script in Bash (similar to 1.d.ii.):

# Place this in the project folder of the respective project you wish to edit.

# Example: place this file in the /usr/freesurfer/subjects/PARTICIPANT/

# Navigate to the Project folder and run using:

# bash FS\_reconAll.sh

# Make sure that the $SUBJECTS\_DIR is specified to the project that you are working on

# Add participants as needed. List out ALL participants and timepoints

for subj in 001.MR1 001.MR2 002.MR1 002.MR2

do

recon-all -s ${subj} -qcache

done

1. **CREATE FSGD File**
   1. Physical activity, sleep quality, and cognitive data for all participants was collected in an excel file. However, FreeSurfer requires the reformulation of this data as a FSGD file. For our main analyses, we created FSGD files for each analysis of interest. Below is a simplified example for our FSGD file for the relationship between physical activity and cortical thickness:

GroupDescriptorFile 1

Title ExampleVer

MeasurmenetName thickness

Class Main

Variables MVPA.Minutes.1 Age Sex Education

Input Subject\_ID\_1 Main 81.875 69.000 0.000 1.000

Input Subject\_ID\_2 Main 66.812 58.000 1.000 0.000

Input Subject\_ID\_3 Main 66.812 64.000 1.000 0.000

1. **ASSEMBLE DATA (MRI PREPROC)**
   1. When you run recon-all with the -qcache option (**See Part 2**), recon-all will resample data onto the average subject (called fsaverage) and smooth it at various FWHM (full-width/half-max) values, usually 0, 5, 10, 15, 20, and 25mm. This can speed later processing. Run this command to assemble your previously cached data (smoothed to 10 mm, needs to be run for both hemispheres):

mris\_preproc --fsgd *fsgd\_file\_name.fsgd* \

--cache-in thickness.fwhm10.fsaverage \

--target fsaverage \

--hemi lh \

--out *output\_filename*.*mgh*

* 1. For this correlational analyses, contrast statements need to be created for the intercept and each variable to be included in the analysis. Using the example FSGD file (**See Part 3**), create the following contrast files (.mtx) for the intercept and each predictor:
     1. Intercept.mtx

**+**1.0000 +0.0000 +0.0000 +0.0000 +0.0000

* + 1. MVPA.mtx

**+**0.0000 +1.0000 +0.0000 +0.0000 +0.0000

* + 1. Age.mtx

**+**0.0000 +0.0000 +1.0000 +0.0000 +0.0000

* + 1. Sex.mtx

**+**0.0000 +1.0000 +0.0000 +0.0000 +0.0000

* + 1. Education.mtx

**+**0.0000 +1.0000 +0.0000 +0.0000 +0.0000

* + 1. These contrast statements will be used to create the following linear regression:

Cortical thickness = β0 + β1\**MVPA +* β2\**Age +* β3\**Sex +* β4\**Education*

1. **MRI GLM\_FIT**
   1. To perform analysis, you will need:
      1. The output file from **Part 4.a** (an .mgh file which will be the hemisphere of the brain you want to analyze) as your dependent variable (this was *output\_filename*.*mgh* in **Part 4.a,** now referred to as *dependent\_variable.mgh*)
      2. The FSGD file you created from **Part 3** (*fsgd\_file\_name.fsgd*)
      3. The Contrast files you crated in **Part 4.b** (*Intercept.mtx,* *MVPA.mtx, Age.mtx, Sex.mtx, Education.mtx*)

mri\_glmfit \

--y *dependent\_variable.mgh* \

-- fsgd *fsgd\_file\_name.fsgd* dods\

--C *Intercept.mtx*\

--C *MVPA.mtx*

--C *Age.mtx*

--C *Sex.mtx*

--C *Education.mtx*

--surf fsaverage lh \

--glmdir *output\_directory.glmdir*

* + 1. As mentioned in **Part 4.b.vi**, including each contrast statement will create the following linear regression equation:

Cortical thickness = β0 + β1\*MVPA + β2\*Age + β3\*Sex + β4\*Education

* + 1. *dods*= “Different Onset Different Slope”. Each groupa has a different intercept and slope (similar to hierarchical linear model). You would want to do this if you have an interaction that you’re interested in looking at. *doss*= “different Onset Same Slope”. Each partici
    2. Open the *output\_directory.glmdir* which is a folder where all your output from the analysis will be placed. The following files will be in the folder for each contrast:

beta.mgh -- all parameter estimates (surface overlay)

dof.dat -- degrees of freedom (text)

fwhm.dat -- average FWHM of residual (text)

*Contrast\_directory* -- contrast subdirectory (multiple directories for multiple contrasts)

mask.mgh -- binary mask (surface overlay)

mri\_glmfit.log -- log file (text, send this with bug reports)

rstd.mgh -- residual standard deviation (surface overlay)

rvar.mgh -- residual variance (surface overlay)

sar1.mgh -- residual spatial AR1 (surface overlay)

surface -- the subject and hemisphere used for this analysis (text)

Xg.dat -- design matrix (text)

X.mat -- design matrix (MATLAB format)Contras

y.fsgd -- copy of input FSGD file (text)

* + 1. To view the results of the analysis, run the following command:

freeview -f $SUBJECTS\_DIR/fsaverage/surf/lh.inflated:overlay= *output\_directory.glmdir* /*Contrast\_directory*/sig.mgh:overlay\_threshold=2,5 -viewport 3d

* + 1. Threshold refers to the p-value threshold (p > log[-x]) which will be displayed. For example, a threshold of 2= p<0.01

1. **CLUSTERWISE CORRECTION FOR MULTIPLE COMPARISONS**
   1. The main advantage to using mri\_glmfit is that it allows the option to correct for multiple comparisons. This can be done one of two ways:
      1. False Positive Ratio correction – This assumes that a certain level of clusters which are significant are due to Type II error, and corrects for this by making the thresholds for significance higher.
      2. Monte-Carlo simulation (Clusterwise Correction) – The simulation is a way to get a measure of the distribution of the maximum cluster size under the null hypothesis. Once we have the distribution of the maximum cluster size, we correct for multiple comparisons by going back to the original data thresholding using same level and sign, and finding clusters in the threshold map. For each cluster, p = probability of seeing a maximum cluster that size or larger during simulation.
         1. To run a clusterwise simulation, use the following command:

mri\_glmfit-sim \

--glmdir *output\_directory.glmdir* \

--cache 4 *sign* \

--cwp 0.05\

--2spaces

* + - 1. *4* = Vertex-wise/cluster-forming threshold of 4 (p < .0001)
      2. *Sign =* Need to specify as either “abs”, “neg”, or “pos”. This determines the directionality of the correlation
      3. --cwp 0.05 : Keep clusters that have cluster-wise p-values < 0.05. To see all clusters
      4. --2spaces : adjust p-values for two hemispheres (this assumes you will eventually look at the right hemisphere too)
      5. In the contrast sub-directory, you will see new files by running:

ls *output\_directory.glmdir*/*Contrast\_directory*

* + - 1. Here are the files you will see in each contrast directory:

cache.th40.neg.sig.pdf.dat -- probability distribution function of clusterwise correction

cache.th40.neg.sig.cluster.mgh -- cluster-wise corrected map (overlay)

cache.th40.neg.sig.cluster.summary -- summary of clusters (text)

cache.th40.neg.sig.masked.mgh -- uncorrected sig values masked by the clusters that survive correction

cache.th40.neg.sig.ocn.annot -- output cluster number (annotation of clusters)

cache.th40.neg.sig.ocn.mgh -- output cluster number (segmentation showing where each numbered cluster is)

cache.th40.neg.sig.voxel.mgh -- voxel-wise map corrected for multiple comparisons at a voxel (rather than cluster) level

cache.th40.neg.sig.y.ocn.dat -- the average value of each subject in each cluster

* + - 1. In order to view the cluster summary, use the following command:

less *output\_directory.glmdir*/*Contrast\_directory* /cache.th40.neg.sig.cluster.summary

* + - 1. To exit, hit the **‘q’** button
      2. In order to view the cluster in freeview, use the following command:

freeview -f $SUBJECTS\_DIR/fsaverage/surf/lh.inflated:overlay= *output\_directory.glmdir*/*Contrast\_directory* /cache.th40.neg.sig.cluster.mgh:overlay\_threshold=2,5:annot= *output\_directory.glmdir*/*Contrast\_directory* /cache.th40.neg.sig.ocn.annot -viewport 3d

1. **GETTING SEGMENTATION STATS FROM CLUSTER OVERLAY**
   1. As detailed in **Part 6.a.ii.7** there will be several files located in each contrast directory

cache.th40.neg.sig.pdf.dat -- probability distribution function of clusterwise correction

cache.th40.neg.sig.cluster.mgh -- cluster-wise corrected map (overlay)

cache.th40.neg.sig.cluster.summary -- summary of clusters (text)

cache.th40.neg.sig.masked.mgh -- uncorrected sig values masked by the clusters that survive correction

cache.th40.neg.sig.ocn.annot -- output cluster number (annotation of clusters)

cache.th40.neg.sig.ocn.mgh -- output cluster number (segmentation showing where each numbered cluster is)

cache.th40.neg.sig.voxel.mgh -- voxel-wise map corrected for multiple comparisons at a voxel (rather than cluster) level

cache.th40.neg.sig.y.ocn.dat -- the average value of each subject in each cluster

The average cortical thickness from each cluster for a given participant will be located in the file highlighted above.