This document provides a GWAS Analysis Plan for use by groups with experience in running GWAS *and* for groups that have not imputed their data using the ENIGMA imputation protocol.

For groups with less experience and those who have run their imputation using the ENIGMA imputation protocol, please follow the protocols on github: <https://github.com/ENIGMA-git/ENIGMA/tree/master/Genetics/ENIGMA3>

Each group has a secure space on the ENIGMA upload server to upload the .info.gz (from imputation) and gzipped association result files. Please contact [enigma3helpdesk@gmail.com](mailto:enigma3helpdesk@gmail.com) to obtain upload information for your group’s data.

**Phenotype preparation:**

The preparation of the cortical surface area and cortical thickness phenotypes and the covariates will be completed using an R script that can be downloaded from the github. **Please do not rescale these variables in any way (do not use residuals, quantile normal transformation or any other transformations). The units of the measurements are mm for cortical thickness and mm2 for surface area.**

It is assumed that you have processed your images using the Extraction and QC protocols available from the ENIGMA webpage.

**Step 1 – Organise the files that you will need for phenotype pre-processing.**

* CorticalMeasuresENIGMA\_SurfAvg.csv - This file contains the mean surface area within each of the FreeSurfer ROI from the Desikan Atlas. In previous steps you marked individual ROIs as “NA” (without the quotes) if that ROI was poorly segmented for a given subject. There are no additional edits required, however, please make sure that the subject IDs in the first column (SubjID) match the ID format in the other files used in the analysis. Also, be sure to save this file as a comma separated (CSV) file after making any edits.
* CorticalMeasuresENIGMA\_ThickAvg.csv - This file contains the mean cortical thickness within each of the FreeSurfer ROI from the Desikan Atlas. In previous steps you marked individual ROIs as “NA” (without the quotes) if that ROI was poorly segmented for a given subject. There are no additional edits required, however, please make sure that the subject IDs in the first column (SubjID) match the ID format in the other files used in the analysis. Also, be sure to save this file as a comma separated (CSV) file after making any edits.
* HM3mds2R.mds.csv – In the genetic imputation protocol, you previously performed an MDS analysis to estimate the ancestry of each subject in your cohort (and to remove subjects with non-homogenous ancestry). Please make sure that the final version of the HM3mds2R.mds.csv file that you end up using in the association analysis contains only the subjects you want to keep in the analysis (i.e. that have a homogenous ancestry). You can plot the MDS values using the code provided in the [imputation protocols](https://github.com/ENIGMA-git/ENIGMA/blob/master/Genetics/ENIGMA2/Imputation). You can uncomment the line in the plotting code to plot the MDS values with the subject IDs overlaid and then remove those subjects from your HM3mds2R.mds.csv file in Excel (or any text editor).
* Covariates.csv – This file you will have to create, but it should be relatively similar (but not the same!) as the covariates files previously created for ENIGMA. Using Excel or your favorite spreadsheet program, create a file that contains the following columns: SubjID, Age, Sex. Additional columns for dummy covariates (i.e. a covariate to control for different MR acquisitions, if applicable) is optional. Save this spreadsheet as a comma delimited (.csv) text file called Covariates.csv.
  + **Note:** If your cohort has both patients and healthy controls, you should include a covariate called “AffectionStatus”, coded as a binary indicator variable where Controls = 0 and Patients = 1. The final file should have the following columns at a minimum: SubjID, Age, Sex, AffectionStatus. Additional columns for dummy covariates (i.e. a covariate to control for different MR acquisitions, if applicable) is optional.
  + **Note 2:** If your cohort has disease-only patients (i.e. no healthy controls) you need to include a covariate called "AffectionStatus", coded such that all patients are = 1. You will end up with a single column of 1's, but the code will handle this properly.
  + **Note 3:** Your Covariates.csv file should not contain any missing values or NA values. If any covariates are missing the whole subject will be removed from the analysis. Please remove any subject with missing values for covariates from the Covariates.csv file before continuing with the scripts below.

**Step 2 – Download the analysis scripts for the pre-processing.**

Using an ssh client download the analysis scripts and put them in the same folder as your .csv files:

svn checkout https://github.com/ENIGMA-git/ENIGMA/trunk/Genetics/ENIGMA3

svn checkout https://github.com/ENIGMA-git/ENIGMA/trunk/Genetics/enigma\_backend

Create a working directory (mkdir) and move (mv) all of the required files inside. Move the enigma\_backend/ folder to be inside of the ENIGMA3/ folder. Then mv the ENIGMA3 folder so that it is renamed to be SCRIPTS/. After this your working directory should look like this:

Your working directory should look something like this:

enigma@-> ls

CorticalMeasuresENIGMA\_SurfAvg.csv

CorticalMeasuresENIGMA\_ThickAvg.csv

HM3mds2R.mds.csv

Covariates.csv

SCRIPTS/

Make sure everything in your folder has executable permission:

chmod -R 777 \*

Change directories to move into the SCRIPTS/ folder.

**Step 3 – Edit run0\_E3\_GWAS\_format.sh .**

**(This script will be used to setup the phenotypes the same way across all groups to ensure homogeneity. Even though you will be using a different software for the GWAS please use this script to make the phenotypes.)**

open run0\_E3\_GWAS\_format.sh and set the following parameters (see descriptions):

#Set the directory where all the enigma association scripts are stored

*run\_directory*=/ENIGMA/CortexGWAS/SCRIPTS/enigma\_backend

#Give the **full path** to R binary, can be found by typing `which R` on the

#command line.

*Rbin*=/usr/local/bin/R

#Give the **full path** to the surf area csv file on your system

*csvFILE\_1*=/ENIGMA/CortexGWAS/CorticalMeasuresENIGMA\_SurfAvg.csv

#Give the **full path** to the thickness csv file on your system

*csvFILE\_2*=/ENIGMA/CortexGWAS/CorticalMeasuresENIGMA\_ThickAvg.csv

#Give the **full path** to a directory to write out the updated and filtered csv

#file (this folder will be created for you)

*csvFOLDER*=/ENIGMA/CortexGWAS/E3

#Please indicate the **full path** to the file where your covariate data is

#stored so that we can merge in relevant covariates to the ENIGMA phenotype files

*TableFile*=/ENIGMA/CortexGWAS/Covariates.csv

#What is the column name where the subject IDs are listed in your Covariates.csv

#file (needed to match subject-by-subject with the ENIGMA files)

*TableSubjectID\_column*="SubjID"

#How many covariates will you be using (note, at a minimum we would require 2

#or 3 -- age and sex and diagnosis (if dataset consists of patients and

#controls (or just patients, see Note 2)), and any additional site-specific #variables,

#please contact us with questions!)

*Ncov*=3

#**Note:** Remember to update this if you change the number of covariates in

#the line below. For example, if you have a healthy-only dataset (i.e. no

#AffectionStatus covariate) you might set this to 2.

#In your covariates file, what are the column headers for the covariates you

#would like to include? Make sure to separate them here with a semi-colon and no

#space!

*covariates*="Age;Sex;AffectionStatus"

#**Note:** If you have a healthy-only cohort you do not need to include an

#AffectionStatus covariate. If you add in additional covariates to control

#for site for example, you can add them here (the name just has to match the

#column name for that variable in the Covariates.csv file). Also, make sure

#that the number of covariates in this step matches the Ncov variable defined

#above.

**Step 4 – Run run0\_E3\_GWAS\_format.sh .**

./run0\_E3\_GWAS\_format.sh

**Step 5 – Edit run1\_GWAS\_flexible\_step1.sh**

**(This script will be used to produce a log file containing the summary statistics for each of the phenotypes that will be included in the tables for the supplementary methods section.)**

open run1\_GWAS\_flexible\_step1.sh and set the following parameters (see below):

#Give the **full path** to where all the enigma association scripts are stored

*run\_directory*=/ENIGMA/CortexGWAS/SCRIPTS/enigma\_backend

#Give the **full path** to R binary (can be found by typing `which R` on the

#command line)

*Rbin*=/usr/local/bin/R

#Give the **full path** to your HM3mds2Rmds.csv file -- has 4 MDS components to

#use as covariates

#(output from the MDS Analysis Protocol)

*csvFILE*=/ENIGMA/CortexGWAS/HM3mds2R.mds.csv

#Give the **full path** to the csv file where your phenotypes and covariates are

#stored after running ./run0\_E3\_GWAS\_format.sh

*combinedROItableFILE*=/ENIGMA/CortexGWAS/E3/combinedROItable\_eCORTEX4GWAS.csv

#Please give some information about the covariate coding you used:

*ageColumnHeader*='Age' # The column header for your age covariate

*sexColumnHeader*='Sex' # The column header for your sex covariate

*maleIndicator*=1 # Males in the sex column coded as (M? 1? 2? ... )

*patients*=1 # Does your dataset contain patients? (mark 0 for no,

# 1 for yes). If your sample has patients and

# controls (or just patients) make sure you have a

# column,(called 'AffectionStatus') where patients are

# marked with 1 and healthy controls with a 0.

#Give the **full path** of the output directory for the ped and dat file

outputs (folder will be created for you)

*peddatdir*=/ENIGMA/CortexGWAS/PedDat/

#**Does you sample have related or unrelated subjects?**

*related*=0 # LEAVE THIS AS 0

**mach2qtl\_DL**=0 # Leave as is.

**run\_machdir**=${run\_directory}/mach2qtl/ # If you downloaded the folder from git

# you do not need to change this

**# You don’t need to edit the rest of the script**

**localfamFILE**="None" # Leave as is.

else

**localfamFILE**=/ENIGMA/CortexGWAS/local.fam # Leave as is.

**merlin\_DL**=0 # Leave as is.

**merlin\_directory**=${run\_directory}/merlin/ # Leave as is.

Fi

**Step 6 – Run run1\_GWAS\_flexible\_step1.sh**

./run1\_GWAS\_flexible\_step1.sh

**Step 7 – Reformat the phenotypes for use with your preferred GWAS program.**

Running run0\_E3\_GWAS\_format.sh will have produced a csv file (CorticalMeasures\_ENIGMA\_ALL\_Avg.csv) that contains the 70 pre-processed phenotypes. This file will be in the directory that you specified in Step 3:

*csvFOLDER*=/ENIGMA/CortexGWAS/E3 The first line of this CSV will contain the names of your ID variable and the 70 phenotypes.

Running run1\_GWAS\_flexible\_step1.sh will have produced merlin format ped and dat files that contain the 70 pre-processed phenotypes and the covariates. This file will be in the directory that you specified in Step 3:

*peddatdir*=/ENIGMA/CortexGWAS/PedDat/ The ped files will contain the names of the 70 phenotypes and the covariates.

The phenotypes will need to be reformatted for use with your preferred GWAS program – the choice of reformatting the csv or the ped/dat files should be based on the required format for you preferred GWAS program.

**For all GWAS analyses please include the follow covariates:**

* Sex [ male = 1, female =0 ]
* Age
* Mean centred Age2 [calculated as (Age-MeanAge)2 ]
* Age\*Sex
* Sex \* Mean centered Age2
* MDS C1 /PC1
* MDS C2 /PC2
* MDS C3 /PC3
* MDS C4 /PC4
* If your cohort has both patients and healthy controls, you should also include a covariate called “AffectionStatus”, coded as a binary indicator variable where Controls = 0 and Patients = 1.
* Please include any additional columns you need to correct for dummy covariates (i.e. a covariate to control for different MR acquisitions, if applicable).

**GWAS – Healthy only Cohorts (including related samples):**

Use linear regression for all analyses. Please use imputed genotypes in dosage or genotype probability format (no hard calls or best guess genotypes). Please do not rescale these variables in any way (do not use residuals, quantile normal transformation or any other transformations). The units of the measurements are mm3 for cortical thickness and mm2 for surface area.

Please run a **GWAS for each phenotype** using the following model resulting in 70 genome wide association analysis runs (A list of the 70 phenotypes is provided at the end of this document):

Phenotype ~ = b0 + b1\*SNP + b2\*Sex + b3\*Age + b4\* Mean centred Age2 + b5\* Age\*Sex + b6\* Sex \* Mean centered Age2 + b7\* C1+ b8\* C2+ b9\* C3+ b10\* C4+ bk\* dummy covariates (if needed)

Please name the output for each trait using the following convention:

{Cohort}\_E3\_cortex\_ healthy\_wo\_{Phenotype}\_{CHR}\_{CHUNK}\_{yyyymmdd}.out.gz

where CHUNK refers to data that have been imputed in chunks (leave this out if you have imputed whole chromosomes or have concatenated the data). Year is given as YYYY, Month is given as MM and Day is given as DD

**PLUS a GWAS for each Surface Area phenotype *correcting for total surface area*** (which is called Mean\_Full\_SurfArea in the csv) using the following model resulting in 34 genome wide association analysis runs (A list of the 34 surface area phenotypes is provided at the end of this document)::

Phenotype ~ = b0 + b1\*SNP **+ b2\*Mean\_Full\_SurfArea** + b3\*Sex + b4\*Age + b5\* Mean centred Age2 + b6\* Age\*Sex + b7\* Sex \* Mean centered Age2 + b8\* C1+ b9\* C2+ b10\* C3+ b11\* C4+ bk\* dummy covariates

Please name the output for each trait using the following convention:

{Cohort}\_E3\_cortex\_ healthy\_**wSA**\_{Phenotype}\_{CHR}\_{CHUNK}\_{yyyymmdd}.out.gz

**PLUS a GWAS for each Thickness phenotype *correcting for mean thickness*** (which is called Mean\_Full\_Thickness) using the following model resulting in 34 genome wide association analysis runs (A list of the 34 Thickness phenotypes is provided at the end of this document):

Phenotype ~ = b0 + b1\*SNP **+ b2\*Mean\_Full\_Thickness** + b3\*Sex + b4\*Age + b5\* Mean centred Age2 + b6\* Age\*Sex + b7\* Sex \* Mean centered Age2 + b8\* C1+ b9\* C2+ b10\* C3+ b11\* C4+ bk\* dummy covariates

Please name the output for each trait using the following convention:

{Cohort}\_E3\_cortex\_ healthy\_**wTHICK**\_{Phenotype}\_{CHR}\_{CHUNK}\_{yyyymmdd}.out.gz

**For each GWAS please retain the following columns of data for each trait:**

* CHR:BP
* Effect\_allele *aka Coded or modelled allele (A/C/G/T/R/I/D)*

*For a A/G SNP in which AA=0, AG=1 and GG=2, the effect allele is G*

* NonEffect\_allele *The other allele (A/C/G/T/R/I/D)*
* Beta *Beta estimate from genotype-phenotype association, at least 5 decimal places –*

*‘NA’ if not available*

* SE *Standard error of beta estimate, to at least 5 decimal places – ‘NA’ if not available*
* Pval *p-value of test statistic – ‘NA’ if not available*
* N *If sample size varies across SNPs – if it does not vary please exclude this column*

***As the information about Allele frequency, Rsq/INFO, RS number and Imputed/Genotyped will be constant across phenotypes please do not include these data in the results files.*** All files should be gzipped. You do not need to concatenate the files across chromosomes – you can upload your results in 22 files for each trait. (*continued over page*)

***Meta data***

In addition to the results files from the 138 GWAS please upload a Meta information file for your cohort {Cohort}\_E3\_cortex\_{CHR}\_{Year}\_{Month}\_{Day}.meta.gz containing the following:

* CHR:BP
* RS\_number *if available*
* Effect\_allele *aka Coded or modelled allele (A/C/G/T/R/I/D)*

*For a A/G SNP in which AA=0, AG=1 and GG=2, the coded allele is G*

* NonEffect\_allele *The other allele (A/C/G/T/R/I/D)*
* Freq\_EA *Frequency of the Effect allele*
* MAF *Minor Allele Frequency*
* Rsq *Rsq or INFO imputation accuracy metric*
* Imputed *Imputed=1, Genotyped =0*

***Phenotypic Summary Statistics***

A file called RUN\_NOTES.txt containing the phenotypic summary statistics was produced when you ran run1\_GWAS\_flexible\_step1.sh. This file should be located in the directory you specified in peddatdir=/ENIGMA/CortexGWAS/PedDat/.

Please rename this file {Cohort}\_RUN\_NOTES\_{yyyymmdd}.txt and include it in your file upload.

**GWAS – Case/Control Cohorts:**

Use linear regression for all analyses. Please use imputed genotypes in dosage or genotype probability format (no hard calls or best guess genotypes). Please do not rescale these variables in any way (do not use residuals, quantile normal transformation or any other transformations). The units of the measurements are mm3 for cortical thickness and mm2 for surface area.

If your cohort includes cases and controls we ask that you run 2 GWAS for each phenotype (1 for controls only and 1 for cases and controls together) resulting in 276 GWAS runs. If you are unsure if your cohort is large enough to allow case only or control only analyses or you do not have sufficient resources to run this many GWAS please email us at [enigma3helpdesk@gmail.com](mailto:enigma3helpdesk@gmail.com).

If you are unsure as to whether the disease/trait that your cohort was selected on requires you to run as a case/control cohort (ie if there was soft selection or the selection was on a non-neurological trait such as BMI) please contact the helpdesk.

***Healthy Control only analyses***

Please run a **GWAS for each phenotype** using the following model resulting in 70 genome wide association analysis runs (A list of the 70 phenotypes is provided at the end of this document):

Phenotype ~ = b0 + b1\*SNP + b2\*Sex + b3\*Age + b4\* Mean centred Age2 + b5\* Age\*Sex + b6\* Sex \* Mean centered Age2 + b7\* C1+ b8\* C2+ b9\* C3+ b10\* C4+ bk\* dummy covariates (if needed)

Please name the output for each trait using the following convention:

{Cohort}\_E3\_cortex\_ healthy\_wo\_{Phenotype}\_{CHR}\_{CHUNK}\_{yyyymmdd}.out.gz

where CHUNK refers to data that have been imputed in chunks (leave this out if you have imputed whole chromosomes or have concatenated the data). Year is given as YYYY, Month is given as MM and Day is given as DD

**PLUS a GWAS for each Surface Area phenotype *correcting for total surface area*** (which is called Mean\_Full\_SurfArea in the csv) using the following model resulting in 34 genome wide association analysis runs (A list of the 34 surface area phenotypes is provided at the end of this document)::

Phenotype ~ = b0 + b1\*SNP **+ b2\*Mean\_Full\_SurfArea** + b3\*Sex + b4\*Age + b5\* Mean centred Age2 + b6\* Age\*Sex + b7\* Sex \* Mean centered Age2 + b8\* C1+ b9\* C2+ b10\* C3+ b11\* C4+ bk\* dummy covariates

Please name the output for each trait using the following convention:

{Cohort}\_E3\_cortex\_ healthy\_**wSA**\_{Phenotype}\_{CHR}\_{CHUNK}\_{yyyymmdd}.out.gz

**PLUS a GWAS for each Thickness phenotype *correcting for mean thickness*** (which is called Mean\_Full\_Thickness) using the following model resulting in 34 genome wide association analysis runs (A list of the 34 Thickness phenotypes is provided at the end of this document):

Phenotype ~ = b0 + b1\*SNP **+ b2\*Mean\_Full\_Thickness** + b3\*Sex + b4\*Age + b5\* Mean centred Age2 + b6\* Age\*Sex + b7\* Sex \* Mean centered Age2 + b8\* C1+ b9\* C2+ b10\* C3+ b11\* C4+ bk\* dummy covariates

Please name the output for each trait using the following convention:

{Cohort}\_E3\_cortex\_ healthy\_**wTHICK**\_{Phenotype}\_{CHR}\_{CHUNK}\_{yyyymmdd}.out.gz

***Case + Control (Full Sample) analyses***

Please run a **GWAS for each phenotype** using the following model resulting in 70 genome wide association analysis runs:

Phenotype ~ = b0 + b1\*SNP + b2\*Sex + b3\*Age + b4\* Mean centred Age2 + b5\* Age\*Sex + b6\* Sex \* Mean centered Age2 + b7\* C1+ b8\* C2+ b9\* C3+ b10\* C4+ b11\* AffectionStatus+ bk\* dummy covariates

Please name the output for each trait using the following convention:

{Cohort}\_E3\_cortex\_ **mixedHD**\_**wo**\_{Phenotype}\_{CHR}\_{CHUNK}\_{yyyymmdd}.out.gz

**PLUS a GWAS for each Surface Area phenotype *correcting for total surface area*** (which is called Mean\_Full\_SurfArea in the csv) using the following model resulting in 34 genome wide association analysis runs:

Phenotype ~ = b0 + b1\*SNP **+ b2\*Mean\_Full\_SurfArea** + b3\*Sex + b4\*Age + b5\* Mean centred Age2 + b6\* Age\*Sex + b7\* Sex \* Mean centered Age2 + b8\* C1+ b9\* C2+ b10\* C3+ b11\* C4 + b11\* AffectionStatus+ bk\* dummy covariates

Please name the output for each trait using the following convention:

{Cohort}\_E3\_cortex\_ **mixedHD**\_**wSA**\_{Phenotype}\_{CHR}\_{CHUNK}\_{yyyymmdd}.out.gz

**PLUS a GWAS for each Thickness phenotype *correcting for mean thickness*** (which is called Mean\_Full\_Thickness) using the following model resulting in 34 genome wide association analysis runs:

Phenotype ~ = b0 + b1\*SNP **+ b2\*Mean\_Full\_Thickness** + b3\*Sex + b4\*Age + b5\* Mean centred Age2 + b6\* Age\*Sex + b7\* Sex \* Mean centered Age2 + b8\* C1+ b9\* C2+ b10\* C3+ b11\* C4+ b11\* AffectionStatus + bk\* dummy covariates

Please name the output for each trait using the following convention:

{Cohort}\_E3\_cortex\_ mixedHD\_**wTHICK**\_{Phenotype}\_{CHR}\_{CHUNK}\_{yyyymmdd}.out.gz

**For each GWAS please retain the following columns of data for each trait:**

* CHR:BP
* Effect\_allele *aka Coded or modelled allele (A/C/G/T/R/I/D)*

*For a A/G SNP in which AA=0, AG=1 and GG=2, the effect allele is G*

* NonEffect\_allele *The other allele (A/C/G/T/R/I/D)*
* Beta *Beta estimate from genotype-phenotype association, at least 5 decimal places –*

*‘NA’ if not available*

* SE *Standard error of beta estimate, to at least 5 decimal places – ‘NA’ if not available*
* Pval *p-value of test statistic – ‘NA’ if not available*
* N *If sample size varies across SNPs – if it does not vary please exclude this column*

***As the information about Allele frequency, Rsq/INFO, RS number and Imputed/Genotyped will be constant across phenotypes please do not include these data in the results files.*** All files should be gzipped. You do not need to concatenate the files across chromosomes – you can upload your results in 22 files for each trait.

***Meta data***

In addition to the results files from the 138 GWAS please upload a Meta information file for your cohort {Cohort}\_E3\_cortex\_{CHR}\_{Year}\_{Month}\_{Day}.meta.gz containing the following:

* CHR:BP
* RS\_number *if available*
* Effect\_allele *aka Coded or modelled allele (A/C/G/T/R/I/D)*

*For a A/G SNP in which AA=0, AG=1 and GG=2, the coded allele is G*

* NonEffect\_allele *The other allele (A/C/G/T/R/I/D)*
* Freq\_EA *Frequency of the Effect allele*
* MAF *Minor Allele Frequency*
* Rsq *Rsq or INFO imputation accuracy metric*
* Imputed *Imputed=1, Genotyped =0* (*continued over page*)

***Phenotypic Summary Statistics***

A file called RUN\_NOTES.txt containing the phenotypic summary statistics was produced when you ran run1\_GWAS\_flexible\_step1.sh. This file should be located in the directory you specified in peddatdir=/ENIGMA/CortexGWAS/PedDat/.

Please rename this file {Cohort}\_RUN\_NOTES\_{yyyymmdd}.txt and include it in your file upload.

**List of the 70 phenotypes (names taken from the csv file made by run0\_E3\_GWAS\_format.sh):**

Mean\_bankssts\_surfavg

Mean\_bankssts\_thickavg

Mean\_caudalanteriorcingulate\_surfavg

Mean\_caudalanteriorcingulate\_thickavg

Mean\_caudalmiddlefrontal\_surfavg

Mean\_caudalmiddlefrontal\_thickavg

Mean\_cuneus\_surfavg

Mean\_cuneus\_thickavg

Mean\_entorhinal\_surfavg

Mean\_entorhinal\_thickavg

Mean\_fusiform\_surfavg

Mean\_fusiform\_thickavg

Mean\_inferiorparietal\_surfavg

Mean\_inferiorparietal\_thickavg

Mean\_inferiortemporal\_surfavg

Mean\_inferiortemporal\_thickavg

Mean\_isthmuscingulate\_surfavg

Mean\_isthmuscingulate\_thickavg

Mean\_lateraloccipital\_surfavg

Mean\_lateraloccipital\_thickavg

Mean\_lateralorbitofrontal\_surfavg

Mean\_lateralorbitofrontal\_thickavg

Mean\_lingual\_surfavg

Mean\_lingual\_thickavg

Mean\_medialorbitofrontal\_surfavg

Mean\_medialorbitofrontal\_thickavg

Mean\_middletemporal\_surfavg

Mean\_middletemporal\_thickavg

Mean\_parahippocampal\_surfavg

Mean\_parahippocampal\_thickavg

Mean\_paracentral\_surfavg

Mean\_paracentral\_thickavg

Mean\_parsopercularis\_surfavg

Mean\_parsopercularis\_thickavg

Mean\_parsorbitalis\_surfavg

Mean\_parsorbitalis\_thickavg

Mean\_parstriangularis\_surfavg

Mean\_parstriangularis\_thickavg

Mean\_pericalcarine\_surfavg

Mean\_pericalcarine\_thickavg

Mean\_postcentral\_surfavg

Mean\_postcentral\_thickavg

Mean\_posteriorcingulate\_surfavg

Mean\_posteriorcingulate\_thickavg

Mean\_precentral\_surfavg

Mean\_precentral\_thickavg

Mean\_precuneus\_surfavg

Mean\_precuneus\_thickavg

Mean\_rostralanteriorcingulate\_surfavg

Mean\_rostralanteriorcingulate\_thickavg

Mean\_rostralmiddlefrontal\_surfavg

Mean\_rostralmiddlefrontal\_thickavg

Mean\_superiorfrontal\_surfavg

Mean\_superiorfrontal\_thickavg

Mean\_superiorparietal\_surfavg

Mean\_superiorparietal\_thickavg

Mean\_superiortemporal\_surfavg

Mean\_superiortemporal\_thickavg

Mean\_supramarginal\_surfavg

Mean\_supramarginal\_thickavg

Mean\_frontalpole\_surfavg

Mean\_frontalpole\_thickavg

Mean\_temporalpole\_surfavg

Mean\_temporalpole\_thickavg

Mean\_transversetemporal\_surfavg

Mean\_transversetemporal\_thickavg

Mean\_insula\_surfavg

Mean\_insula\_thickavg

Mean\_Full\_SurfArea

Mean\_Full\_Thickness

**List of the 34 surface area phenotypes**

Mean\_bankssts\_surfavg

Mean\_caudalanteriorcingulate\_surfavg

Mean\_caudalmiddlefrontal\_surfavg

Mean\_cuneus\_surfavg

Mean\_entorhinal\_surfavg

Mean\_fusiform\_surfavg

Mean\_inferiorparietal\_surfavg

Mean\_inferiortemporal\_surfavg

Mean\_isthmuscingulate\_surfavg

Mean\_lateraloccipital\_surfavg

Mean\_lateralorbitofrontal\_surfavg

Mean\_lingual\_surfavg

Mean\_medialorbitofrontal\_surfavg

Mean\_middletemporal\_surfavg

Mean\_parahippocampal\_surfavg

Mean\_paracentral\_surfavg

Mean\_parsopercularis\_surfavg

Mean\_parsorbitalis\_surfavg

Mean\_parstriangularis\_surfavg

Mean\_pericalcarine\_surfavg

Mean\_postcentral\_surfavg

Mean\_posteriorcingulate\_surfavg

Mean\_precentral\_surfavg

Mean\_precuneus\_surfavg

Mean\_rostralanteriorcingulate\_surfavg

Mean\_rostralmiddlefrontal\_surfavg

Mean\_superiorfrontal\_surfavg

Mean\_superiorparietal\_surfavg

Mean\_superiortemporal\_surfavg

Mean\_supramarginal\_surfavg

Mean\_frontalpole\_surfavg

Mean\_temporalpole\_surfavg

Mean\_transversetemporal\_surfavg

Mean\_insula\_surfavg

**List of the 34 thickness phenotypes**

Mean\_bankssts\_thickavg

Mean\_caudalanteriorcingulate\_thickavg

Mean\_caudalmiddlefrontal\_thickavg

Mean\_cuneus\_thickavg

Mean\_entorhinal\_thickavg

Mean\_fusiform\_thickavg

Mean\_inferiorparietal\_thickavg

Mean\_inferiortemporal\_thickavg

Mean\_isthmuscingulate\_thickavg

Mean\_lateraloccipital\_thickavg

Mean\_lateralorbitofrontal\_thickavg

Mean\_lingual\_thickavg

Mean\_medialorbitofrontal\_thickavg

Mean\_middletemporal\_thickavg

Mean\_parahippocampal\_thickavg

Mean\_paracentral\_thickavg

Mean\_parsopercularis\_thickavg

Mean\_parsorbitalis\_thickavg

Mean\_parstriangularis\_thickavg

Mean\_pericalcarine\_thickavg

Mean\_postcentral\_thickavg

Mean\_posteriorcingulate\_thickavg

Mean\_precentral\_thickavg

Mean\_precuneus\_thickavg

Mean\_rostralanteriorcingulate\_thickavg

Mean\_rostralmiddlefrontal\_thickavg

Mean\_superiorfrontal\_thickavg

Mean\_superiorparietal\_thickavg

Mean\_superiortemporal\_thickavg

Mean\_supramarginal\_thickavg

Mean\_frontalpole\_thickavg

Mean\_temporalpole\_thickavg

Mean\_transversetemporal\_thickavg

Mean\_insula\_thickavg