

THE ENIGMA SUBCORTICAL SHAPE ANALYSIS PIPELINE

Contact:

Dr. Boris Gutman (bgutman@gmail.com)

Chris Ching (Chris.Ching@loni.usc.edu)

Zvart Abaryan (zvart06@gmail.com)

Adam Mezher (adammezher@gmail.com)

Sean Hatton (sean.hatton@sydney.edu.au)

TABLE OF CONTENTS

[1. Introduction 3](#_Toc430638965)

[2. Quality checking Freesurfer output using ENIGMA protocol 4](#_Toc430638966)

[3. Generating meshes and shape data (medial daemons) 4](#_Toc430638967)

[3.1. Introduction 4](#_Toc430638968)

[3.2. Citations 5](#_Toc430638969)

[3.3. Setting up the software packages 6](#_Toc430638970)

[3.3.1 Set environment variables in shape script 6](#_Toc430638971)

[3.3.2 Set environment variables in AutoQA script 7](#_Toc430638972)

[3.4. Running mesh analysis on individual subjects 7](#_Toc430638973)

[3.5. Medial demons output 8](#_Toc430638974)

[3.6. Troubleshooting mesh creation 9](#_Toc430638975)

[3.6.1. LogJacs\_[ROI\_id].raw files are empty 9](#_Toc430638976)

[4. Quality Assurance (AutoQA) 10](#_Toc430638977)

[4.1. Introduction 10](#_Toc430638978)

[4.2. AutoQA prerequisites 10](#_Toc430638979)

[4.3. Running AutoQA 11](#_Toc430638980)

[4.3.1. Create your group file 11](#_Toc430638981)

[4.3.2. Modify Matlab paths in scripts 11](#_Toc430638982)

[4.3.3. Arguments for running AutoQA.sh script 11](#_Toc430638983)

[4.3.4. Run the AutoQA script 12](#_Toc430638984)

[4.4. AutoQA outputs 12](#_Toc430638985)

[4.5. Reviewing the AutoQA output 13](#_Toc430638986)

[4.5.1. View snapshots of all subjects by viewing the snapshots 13](#_Toc430638987)

[4.6. Rerunning AutoQA scripts 14](#_Toc430638988)

[4.7. Troubleshooting AutoQA 14](#_Toc430638989)

[4.7.1. Matlab not running 14](#_Toc430638990)

[4.7.2. Matlab openGL Fix 14](#_Toc430638991)

[5. Group batch script: shape analysis and data extraction 16](#_Toc430638992)

[5.1. Introduction 16](#_Toc430638993)

[5.2. Configuring the group batch script 16](#_Toc430638994)

[5.3. Running the group batch script 16](#_Toc430638995)

[6. Concluding remarks 17](#_Toc430638996)

# Introduction

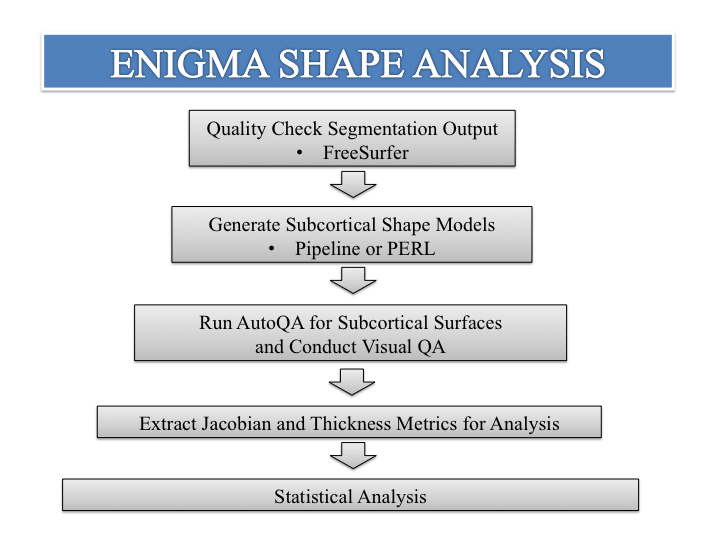
Neuroscientists and researchers in related fields are increasingly looking into novel methods of assessing changes in subcortical structures in the context of normal brain development as well as neurodegenerative diseases. The Shape Analysis tools discussed here offer one way to study subcortical structures in a semi-automated fashion. While currently designed to provide cross-sectional information, this method will support longitudinal analysis in the future.

The overall process for Shape Analysis involves three main steps.

1) Using FreeSurfer generated subcortical labels to generate meshes and shape metrics (medial deamon stage).

2) Quality checking data (AutoQA), and

3) Extracting shape measures (Jacobians and thickness values) for statistical analysis (extraction scripts).



These steps are described in detail on the following pages with additional instructions on how to modify the process should you find it necessary. You will find all necessary files in the Dropbox link that has been provided to you.

* Items in yellow highlight you may need to amend to your environment.
* Items in green highlight are important information.

# Quality checking Freesurfer output using ENIGMA protocol

The shape processing pipeline requires FreeSurfer or FSL generated subcortical labels. Prior to shape processing, it is ESSENTIAL that you quality inspect the subcortical segmentations you are interested in analyzing with the shape modeling pipeline. The ENIGMA FreeSurfer Quality Checking of Subcortical Structures protocol can found here:

<http://enigma.ini.usc.edu/wp-content/uploads/2015/06/Instructions_ENIGMA_subcortical_volumes.doc>

# Generating meshes and shape data (medial daemons)

## Introduction

Shape registration is based on existing shape templates and template medial models. The USC IGC team has provided a master template of all 14 subcortical structures of interest built from 200 young adults. The atlas was made by registering all subjects to a representative subject. The Euclidean average served as the template surface, on which the template medial curve was computed.

Two point-wise measures of shape morphometry are derived for each subject run through the pipeline. The first, termed radial distance, is derived by a medial model approach where for each point onthe surface, and given a medial curve, the radial distance is defined by

The second, based on surface Tensor Based Morphometry (TBM), generalizes TBM on Euclidean spaces to surfaces. The differential map between the tangent spaces of two surfaces replaces the Jacobian.

In our model, Mt is the average template, and M is the surface we wish to study. J is a linear mapping, and may be thought of as the restriction of the standard Jacobian to the tangent spaces of the template and study surfaces. While analysis of the full tensor using Log-Euclidean metrics on SPD matrices is possible, such analyses are difficult to interpret. Instead, our model considers the Jacobian determinant, representing the surface dilation ratio between the template and the study subject. An interpretation of this measure is that of a ratio of the area of a small surface patch around a particular point of the subject surface and the small patch of area around the corresponding point on the template. A higher Jacobian may indicate larger volume of a structure’s subfield corresponding to the region. Our final TBM measure is the logarithm of the Jacobian determinant, to obtain a distribution closer to Gaussian.

In this way, both radial size (termed *thickness* from now on) and the logarithm of the Jacobian determinant (termed simply *Jacobian* from now on) are calculated for up to 2,500 points across each subcortical structure, providing a sensitive index of regional shape differences across subjects.

## Citations

Please cite this work as follows:

1. Gutman, B.A., Madsen, S.K., Toga, A.W., Thompson, P.M.: A Family of Fast Spherical Registration Algorithms for Cortical Shapes. In: Shen, L., Liu, T., Yap, P.-T., Huang, H., Shen, D., Westin, C.-F. (eds.) Multimodal Brain Image Analysis, vol. 8159, pp. 246-257. Springer International Publishing (2013)

2. Gutman, B.A., Wang, Y., Rajagopalan, P., Toga, A.W., Thompson, P.M.: Shape matching with medial curves and 1-D group-wise registration. In: Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on, pp. 716-719. (2012)

For Alzheimer’s-related work, you may also wish to cite:

3. Gutman, B.A., Hua, X., Rajagopalan, P., Chou, Y.-Y., Wang, Y., Yanovsky, I., Toga, A.W., Jack Jr, C.R., Weiner, M.W., Thompson, P.M.: Maximizing power to track Alzheimer's disease and MCI progression by LDA-based weighting of longitudinal ventricular surface features. Neuroimage 70, 386-401 (2013)

## Setting up the software packages

1. Copy and unzip the ENIGMA\_Shape package to a Linux folder of your liking, for example:

/usr/local/ENIGMA\_Shape

Your project directory structure should look like the following:

ENIGMA\_Shape/

AutoQA\_prep/*AutoQA scripts and references*

MedialDemonsShared/

bin/*medial\_demons binary files*

atlas/*mesh reference atlases*

groupfile.csv: example group file

shape\_group\_run.sh: group batch script

2. Ensure files are read/write/executable by using *chmod* in terminal:

chmod –R 777 /usr/local/ENIGMA\_Shape

3. CD to the MedialDemonsShared/bin folder and test if the ccbbm, MeshfromAnalyze, and raw\_operations programs can be run*.* In all three cases, usage instructions should be printed if these programs are called with no arguments.

## Set environment variables in shape script

Open /usr/local/ENIGMA\_Shape*/*shape\_group\_run.sh in a text editor. Find ‘Path to Freesurfer.’

1. Change ‘FS’ variable (approximately line 22) to your local FreeSurfer installation. For example:

FS=/home/common/applications/freesurfer/freesurfer/ (yours will likely be different)

1. Change ‘runDirectory’ (approximately line 26) to you ENIGMA\_Shape directory. Following our example:

runDirectory=*/path/to/ENIGMA\_Shape/*

Save the shape\_group\_run.sh file.

## Set environment variables in AutoQA script

Open /usr/local/ENIGMA\_Shape*/AutoQA\_prep/AutoQA*.sh in a text editor. Find ‘MatlabPath.’

1. Change ‘MatlabPath’ variable to your local Matlab binary (approximately line 14). Note that your entry must contain the binary file ‘matlab’ at the end, so that if copied and pasted into your terminal, it opens the Matlab environment:

MatlabPath=/home/common/applications/MATLAB/matlab/bin/matlab (yours will likely be different)

1. Change the FREESURFER\_HOME variable, as done previously in shape\_group\_run.sh (approximately line 24):

FREESURFER\_HOME=/home/common/applications/freesurfer/freesurfer/ (this must be exactly the same as in step 3.3.1)

Save the *AutoQA*.sh file. You are now ready to use the ENIGMA shape package!

## Running mesh analysis on individual subjects

On Linux command line, type:

perl bin/Medial\_Demons\_shared.pl [FreeSurfer aseg file] [label id list (integer)] [output directory] [medial package directory] [freesurfer bin directory]

* ***[FreeSurfer aseg file]*** – This is the aseg.mgz file, a.k.a. the FreeSurfer segmentation
* ***[label id list (int)]*** – This is the list label indices, which indicates which regions of interest (ROIs) will be analyzed. In this package, atlases for the 14 ENIGMA ROIs are available, their index followed by the region:

|  |  |  |
| --- | --- | --- |
| 10 Left-Thalamus-Proper  11 Left-Caudate  12 Left-Putamen  13 Left-Pallidum  17 Left-Hippocampus | 18 Left-Amygdala  26 Left-Accumbens-area  49 Right-Thalamus-Proper  50 Right-Caudate | 51 Right-Putamen  52 Right-Pallidum  53 Right-Hippocampus  54 Right-Amygdala  58 Right-Accumbens-area |

* ***[output directory]*** – this is the path to the folder where per-vertex shape data is stored. It is created if it does not exist.
* ***[medial package directory]*** – This is the folder where the medial demons package is placed, in our example above it is ***/usr/local/ENIGMA\_Shape/MedialDemonsShared.*** It should contain two subfoders, ***bin*** and ***atlas.***
* ***[freesurfer bin directory]*** – this is the directory containing the freesurfer executables, in particular ***mri\_convert***. For example, ***/usr/local/freesurfer-5.3.0\_64bit/bin***

An example command:

perl bin/Medial\_Demons\_shared.pl

/ifshome/bgutman/tate2/aseg\_S\_2.mgz

10 11 12 13 17 18 26 49 50 51 52 53 54 58

/usr/local/ENIGMA\_Shape/ShapeAnalysis/

/usr/local/ENIGMA\_Shape/medial\_demons

/usr/local/freesurfer-5.3.0\_64bit/bin

NOTE: Running this program will create a temporary directory within the folder containing aseg.mgz - you must have write permissions in the FreeSurfer folder containing aseg.mgz. Upon completing execution, the script erases this folder. If you are running this for several ROI’s simultaneously, there will be several temporary folders created.

## Medial demons output

The script above creates a folder for each subject in which shape data is written. For each ROI specified in the ***[label id list (int)]*** argument, there are four outputs:

* ***resliced\_mesh\_[ROI\_id].m*** - this is a mesh with points corresponding to a pre-set atlas that can be viewed in shapeviewer for basic QC
* ***curve\_[ROI\_id].ucf*** - medial curve of the shape, can be loaded into shapeviewer simultaneously as the mesh for more basic QC
* ***thick\_[ROI\_id].raw*** - thickness, or radial distance in binary format, literally how thick the shape is
* ***LogJacs\_[ROI\_id].raw*** - Log of the Jacobian determinant, in binary format (shape TBM variant), see e.g. Wang, Y., et al.: Applying tensor-based morphometry to parametric surfaces can improve MRI-based disease diagnosis. Neuroimage 74, 209-230 (2013). This measure can be interpreted as the ratio of surface areas between corresponding patches on the average healthy atlas, and the subject. Really, it is just like volume TBM, suggesting relative size of the underlying regional volume.

The last two items are the per-vertex quantitative traits. Statistics resulting from these traits can be mapped as p, beta, or any other kind of maps on the subcortical surface atlas.

For additional information regarding loading mesh data in Matlab and visualization refer to the appendix.

## Troubleshooting mesh creation

### LogJacs\_[ROI\_id].raw files are empty

Ensure that all the files in the bin directory are executable. In terminal within the folder type:

chmod +x \*

# Quality Assurance (AutoQA)

## Introduction

It is important to check the quality of the shape models generated by the ENIGMA Shape Analysis pipeline. The AutoQA will generate a .png image for each shape model that includes the following for quality inspection:

1. 3-D surface model (generated by Shape Analysis Pipeline).
2. Outline of 3-D surface model overlaid on each subjects T1 image (from FreeSurfer).

## AutoQA prerequisites

* If you are running AutoQA without a display, you need to have VNC set up (matlab requires an X display for snapshots).
* Binaries are compiled for CentOS 6 x86\_64, if you have another architecture, let me (k-alpert@northwestern.edu) know and I can send you other binaries (or potentially src)
* Use ROI and SUBJID in the commands to stand-in for file names that include FS ROI numbers (like 17 or 53) and subject IDs. i.e. do not change them.

The AutoQA\_prep directory contains:

* ***AutoQA.sh*** - the main QA script that streamlines the snapshot generating processes
* ***erase\_byu.sh*** - a script that erases the byu from shape analysis subject directories.
* ***scripts*** (directory) -
  + ***matlab*** (directory) : contains the matlab scripts
  + ***bin*** (directory): compiled matlab binaries and a version of ccbbm
  + ***func.sh*** : auxiliary functions & log writing

## Running AutoQA

### Create your group file

In the root directory is an example groupfile.csv file. Most likely, you already have this file for your study. The format is the same as for the group file required in the medial demons step. The first row is ignored and lists the column headers “id” or “SubjId” for subject ID in the first column, and other data in other columns (also ignored). If you do not already have a file like this, you can use vi, gedit, or Excel to create your groupfile.csv. Note: The csv file MUST BE SAVE IN PC or UNIX format, also, the .csv file must have at least one additional column besides subject ids. If needed, this could be a dummy variable.

### Modify Matlab paths in scripts

Note that the AutoQA.sh script should be used for cross-sectional datasets with only a single timepoint but can be applied to datasets where there are multiple time points (running each time point separately). All paths that require modification are marked with “#----modify-this---- “.

In the AutoQA.sh script find the two instances of “#Edit the 2 lines below providing the path to your matlab location” and insert the path to your Matlab directory. The location can be found by typing in the terminal window “which matlab”. (This is already covered in the set-up section)

### Arguments for running AutoQA.sh script

The following arguments must be set when executing the AutoQA.sh script:

* Mandatory flags:

**-s** [ Mesh file directory ending with "SUBJID/resliced\_mesh\_ROI.m" ]

**-f** [ FreeSurfer Output Directory ]

**-g**  [ Path to the a .csv file containing subject ids]

**-o** [ QA (this script’s) Output directory ]

**-n** [ List of ROI region comma separated; e.g. '17,49'. You can also enter 'All'

to selected all of the ROIs ]

**-r** [ Redo option, enter either '0' or '1'; redoing will only delete and regenerate

the ".byu" files, the QC images will always get updated. ]

* Optional flags:

**-m** [ When using matlab version 2014b and above, you must turn this flag on for the script to function. ]

**-e** [ Do not overwrite previously generated CSV files. ]

**-h** [ This will print out the script guide. ]

EXAMPLE:

./AutoQA\_prep.sh

-s /Project\_directory/ShapeAnalysis/SUBJID/resliced\_mesh\_ROI.m

-f /Project\_directory/FreeSurfer

-g /Project\_directory/AutoQA\_prep/subjectfile/subjectlist.txt

-o /Project\_directory/ShapeAnalysis/QA

-n 17,53

-r 0

NOTE:

Please provide full paths to the required inputs. (i.e. /Project\_directory/ShapeAnalysis/SUBJID/resliced\_mesh\_ROI.m)

### Run the AutoQA script

In terminal, change directories into your AutoQA directory and run the AutoQA.sh script. The processing time varies depending on number or ROIs and subjects. It is roughly 2-3 minutes per subject if running with –n All flag. IMPORTANT!!! Because this script does some low-level computing on your local machine, you should:

1. Keep it logged in while the script is running
2. Log in with –XY flag, for example:

$ ssh –XY [your\_username@your.server.institution.edu](mailto:your_username@your.server.institution.edu)

This script seems to run more reliably when logged in from a MAC than a PC, particularly if your server is running a pre-2014b version of MATLAB

## AutoQA outputs

Your output will be located in the QA folder that you specified when modifying the AutoQA.sh script.

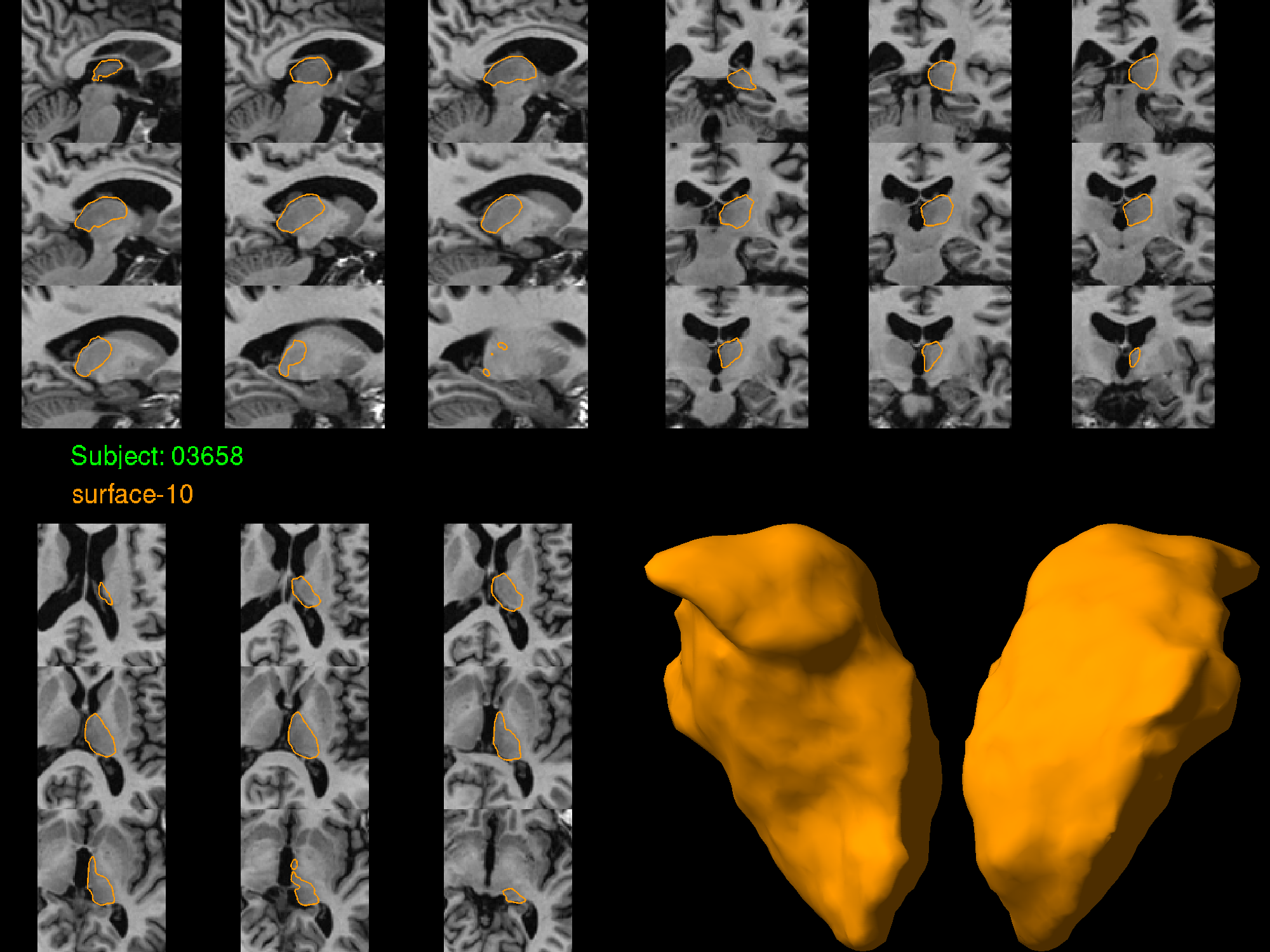
* A “quickcheck” directory with a subdirectory for each group. Each group directory in turn has an ROI subdirectory, each of which contains a “{subject}\_surface\_{roi}.png” file. This PNG file is important as it contains surface images in 3D and cross section. This is the primary means by which you should be looking through your AutoQA results.

## Reviewing the AutoQA output

### View snapshots of all subjects by viewing the snapshots

Following the example command from 4.3.3, the .png files to perform a visual QA check on will be in the following directory (your actual path will be specific to your command):

/Project\_directory/ShapeAnalysis/QA/quickcheck/ROI\_$roi



* Notice that you are given a set of images in in the 3 different planes, the subject number, ROI or “surface” number, and rendered surfaces (two views) of the ROI in question.
* Using these images, you can decide whether the surface model captures the boundary of the ROI reasonably well. The script will generate 2 .csv spreadsheets in the directory /Project\_directory/ShapeAnalysis/QA,

following our example above (the directory set with –o flag).

These files are: QA\_Failed\_Subjects.csv and QA\_Status.csv. These files are initialized with a list of missing meshes:

- QA\_Failed\_Subjects.csv – lists the subject id’s under the corresponding ROI

- QA\_Status.csv – contains an ROI by Subject matrix, with 0’s for missing mesh and 1’s for existing mesh. Note that the QA rating defines 1 as fail, 2 as moderate pass and 3 as pass. For using this file in the subsequent statistical analysis, you will have to change entries for shapes which passed visual QA to 2’s and 3’s.

* For examples of good and bad meshes refer to ENIGMA\_SHAPE\_QA\_MANUAL.pdf guide located in the AutoQA\_prep directory.

## Rerunning AutoQA scripts

The AutoQA scripts create a byu surface file located where you designated the mesh outputs from the Medial Demons scripts. If you wish to rerun the AutoQA because you have re-run the Medial Demons scripts, you will need to regenerate the .byu files. This can be done by setting –r 1 in the command. If there is no need to overwrite these, for example if the script could not execute mri\_convert in the previous attempt, setting –r 0 will save a little bit of time. If you’re not sure, use –r 1.

## Troubleshooting AutoQA

### Matlab not running

Ensure that you are using Matlab 2014a or earlier, or use the –m flag if using a later Matlab version.

### Matlab openGL Fix

Prior to running any of the scripts, be sure to type the following into terminal:

export LIBGL\_ALWAYS\_INDIRECT=1

For assistance, please contact Kate Alpert (k-alpert@northwester.edu), Boris Gutman (bgutman@gmail.com) Chris Ching (Chris.Ching@loni.usc.edu); Adam Mezher (adammezher@gmail.com).

# Group batch script: shape analysis and data extraction

## Introduction

With an understanding of how to create mesh files (Section 3) for an individual subject, this section describes how to run the process for a whole dataset in a single bash command.

* One batch file with all ROIs pointing to one groupfile.csv (slower).
* Multiple batch files with all ROIs pointing to different .csv files (faster).

## Configuring the group batch script

Assuming a standard FreeSurfer directory structure, the set-up done in section 3.3.1 will be sufficient. However, if you have a different structure (for example, you’ve downloaded FreeSurfer results on a publicly available dataset), you have a different file extension or name for your aseg files, for example aseg+aparc.nii.gz, or you have used FSL, you may still use the group script. In this case, you will have to modify shape\_analysis\_group.sh around line 99 so that the line containing ${runFS}/${subject}/mri/aseg.mgz is modified appropriately to point to your parcellation data.

## Running the group batch script

The script may be run as below (your path to ENIGMA\_Shape, and other inputs/outputs will likely be different):

$ /usr/local/ENIGMA\_Shape/shape\_group\_run.sh /path/to/Subject/Info/File/groupfile.csv /path/to/FreeSurfer/Output /path/to/MedialDemons/Output

Ensure that you have write permissions in the FreeSurfer directory for the script to create temporary directories, for example with the command:

$ chmod –R 777 /path/to/FreeSurfer/Output

The script will produce the following output:

* /path/to/MedialDemons/Output/${subject\_id} directories corresponding to the subject identifiers in your groupfile.csv. For contents of these directories, refer to section 3.4
* Two .csv files containing the entire data matrix for your dataset in groupfile.csv:

/path/to/MedialDemons/Output/${groupfile}\_LogJacs.csv

/path/to/MedialDemons/Output/${groupfile}\_thick.csv

These files contain all the vertex-wise data for the two measures, with each subject contained in a single row. For subjects and ROI’s which have missing data or mesh files, the appropriate row section is filled with ‘NaN.’ Because the name of the file is based on the name of your groupfile, it is crucial to have differently named groupfiles if they point to different subjects (for example, if splitting the list to speed up computation)

On completion, run and review AutoQA (Section 4.5) and create the visual QA table as described in section 4. This is the most labor-intensive aspect. Once QA is done, you are ready to begin your statisticsThis can be done either by forwarding your /${groupfile}\_LogJacs.csv and /${groupfile}\_thick.csv files to the IGC if individual-level shape data can be shared, or running R-scripts specially designed for shape data for any number of association tests one may wish to run. In the latter case, the dataset will presumably participate only in a meta-analysis. Please consult the IGC team regarding existing code to run pointwise statistics on Jacobian and thickness shape measures.

# Concluding remarks

We appreciate your efforts and motivation to help us develop and improve the ENIGMA subcortical shape analysis pipeline protocol and tools. Please contact us (see contact information on title page) with any questions, comments, or suggestions.