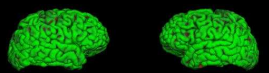
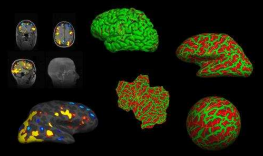
Surface-based Group Analysis in FreeSurfer



FreeSurfer



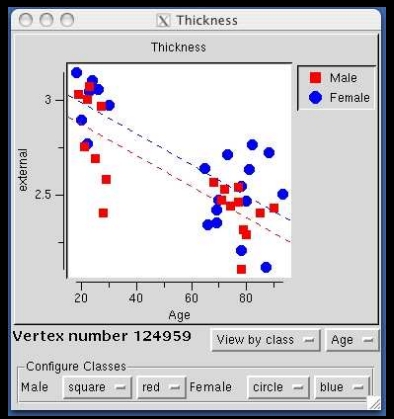
MASSACHUSETTS GENERAL HOSPITAL

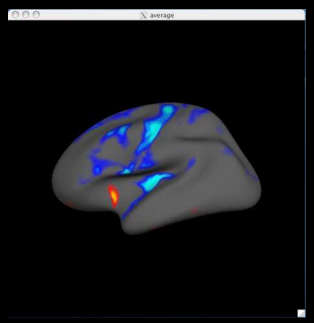
surfer.nmr.mgh.harvard.edu/docs/ftp/pub/docs/freesurfer.groupanalysis.pdf

Outline

- Processing Stages
- Manual Stream
  - Assemble Data
  - Design/Contrast (GLM Theory)
  - Analyze
  - Visualize
- Interactive/Automated GUI (QDEC)
- Correction for multiple comparisons

Surface-based Study (Thickness)

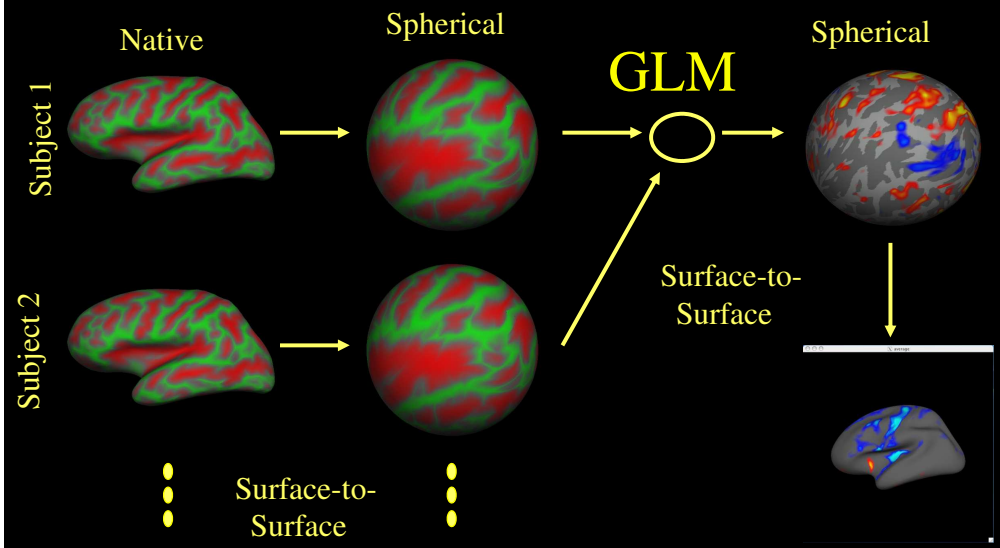




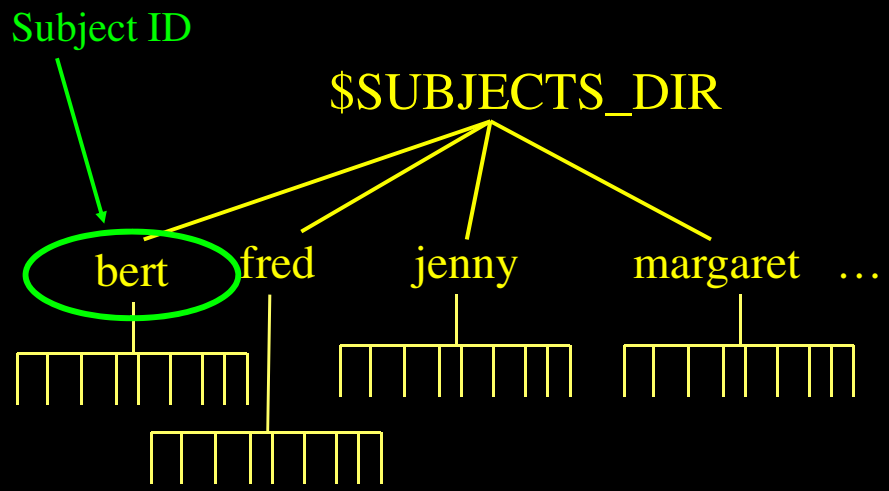
Processing Stages

- Specify Subjects and Surface measures
- Assemble Data:
  - Resample into Common Space
  - Smooth
  - Concatenate into one file
- Model and Contrasts (GLM)
- Fit Model (Estimate)
- Visualize

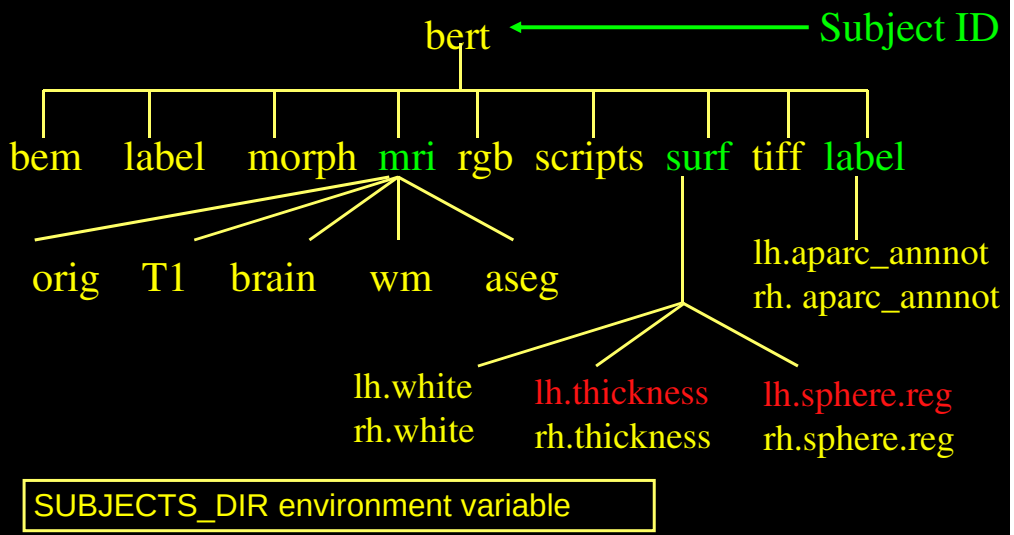
Inter-Subject Averaging



Specifying Subjects

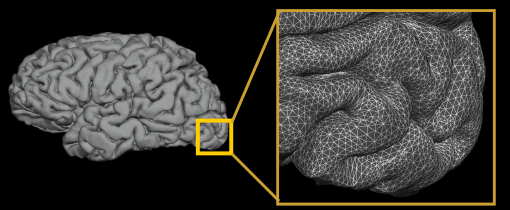


FreeSurfer Directory Tree



Surface-based Measures

- Morphometric (eg, thickness)
- Functional
- PET
- MEG/EEG
- Diffusion (?) sampled just under the surface



"Manual" Processing Stages

- Assemble Data (mris_preproc)
  - Resample into Common Space
  - Smooth
  - Concatenate into one file
  - Cache In/Out with –cache-xxx options
- Model and Contrasts (GLM) (FSGD)
- Fit Model (Estimate) (mri_glmfit)
- Visualize (tksurfer)

Assemble Data: mris_preproc

mris_preproc --help

--fsgd FSGDFile : Specify subjects thru FSGD File
--hemi lh : Process left hemisphere
--meas thickness : $SUBJECTS_DIR/subjectid/surf/hemi.thickness
--target fsaverage : common space is subject fsaverage
--o lh.thickness.mgh : output "volume-encoded surface file"

Lots of other options!

lh.thickness.mgh – file with thickness maps for all subjects à Input to Smoother or GLM

Surface Smoothing

- mri_surf2surf --help
- Loads lh.thickness.mgh
- 2D surface-based smoothing (eg, fwhm = 10 mm)
- Saves lh.thickness.sm10.mgh

Why should you smooth?
- Improve CNR
- Improve inter-subject registration

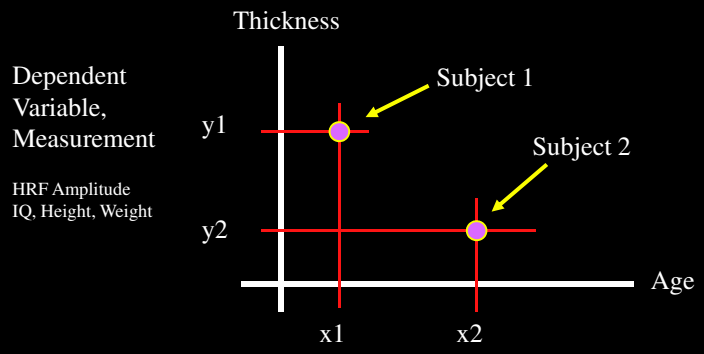
How much smoothing?
- Blob-size
- Typically 10-20 mm FWHM
- More forgiving than volume-based

General Linear Model (GLM)

- t-test
- Paired t-test
- F-test
- ANOVA
- ANCOVA
- MANOVA
- MANCOVA
- Multiple Regression

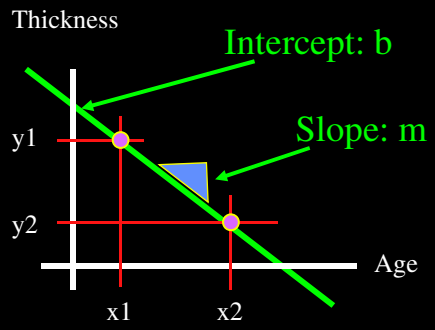
GLM Theory

Is Thickness correlated with Age?



Of course, you'd need more then two subjects …

Linear Model



Matrix Formulation

$$\begin{bmatrix} y1 \\ y2 \end{bmatrix} = \begin{bmatrix} 1 & x1 \\ 1 & x2 \end{bmatrix} * \begin{bmatrix} b \\ m \end{bmatrix}$$

$$Y = X*\beta \qquad \beta = \begin{bmatrix} b \\ m \end{bmatrix}$$

System of Linear Equations

$y1 = b + x1*m$

$y2 = b + x2*m$

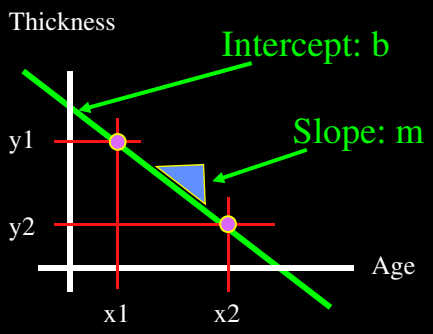
X = Design Matrix

β = Regression Coefficients
= Intercepts and Slopes

Intercept = Offset

Inference

Is Thickness correlated with Age?
Does m = 0?
Null Hypothesis: H0: m=0

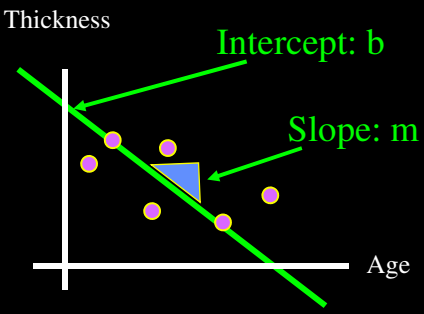


$$m = [0\ 1]*\begin{bmatrix} b \\ m \end{bmatrix} \qquad \beta = \begin{bmatrix} b \\ m \end{bmatrix}$$

$$\gamma = C*\beta \stackrel{?}{=} 0$$

C=[0 1]: Contrast Matrix

More than Two Data Points



$$\begin{bmatrix} y1 \\ y2 \\ y3 \\ y4 \end{bmatrix} = \begin{bmatrix} 1 & x1 \\ 1 & x2 \\ 1 & x3 \\ 1 & x4 \end{bmatrix} * \begin{bmatrix} b \\ m \end{bmatrix}$$

$$Y = X*\beta$$

$y1 = b + x1*m$

$y2 = b + x2*m$

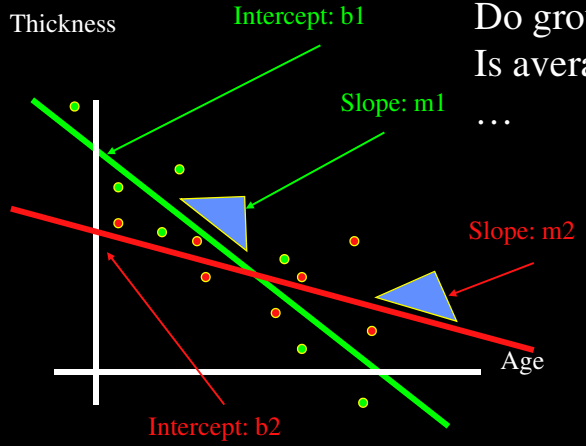
$y3 = b + x3*m$

$y4 = b + x4*m$

Contrast Matrix Stays Same:

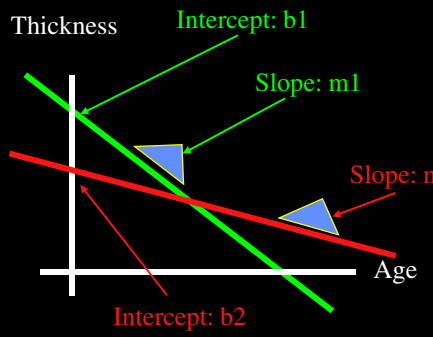
C = [0 1]

Two Groups



Do groups differ in Intercept?
Do groups differ in Slope?
Is average slope different than 0?
…

Two Groups



$$\begin{bmatrix} y11 \\ y12 \\ y21 \\ y22 \end{bmatrix} = \begin{bmatrix} 1 & 0 & x11 & 0 \\ 1 & 0 & x12 & 0 \\ 0 & 1 & 0 & x21 \\ 0 & 1 & 0 & x22 \end{bmatrix} * \begin{bmatrix} b1 \\ b2 \\ m1 \\ m2 \end{bmatrix}$$

$$Y = X*\beta$$

$y11 = b1 + x11*m1$
$y12 = b1 + x12*m1$
$y21 = b2 + x21*m2$
$y22 = b2 + x22*m2$

Two Groups

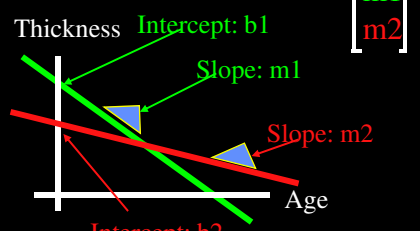
Do groups differ in Intercept?
Does b1=b2?
Does b1-b2 = 0?
C = [+1 -1 0 0], $\gamma = C*\beta$

Do groups differ in Slope?
Does m1=m2?
Does m1-m2=0?
C = [0 0 +1 -1], $\gamma = C*\beta$

Is average slope different than 0?
Does (m1+m2)/2 = 0?
C = [0 0 0.5 0.5], $\gamma = C*\beta$

$$\begin{bmatrix} y11 \\ y12 \\ y21 \\ y22 \end{bmatrix} = \begin{bmatrix} 1 & 0 & x11 & 0 \\ 1 & 0 & x12 & 0 \\ 0 & 1 & 0 & x21 \\ 0 & 1 & 0 & x22 \end{bmatrix} * \begin{bmatrix} b1 \\ b2 \\ m1 \\ m2 \end{bmatrix}$$

$$Y = X*\beta \qquad \beta = \begin{bmatrix} b1 \\ b2 \\ m1 \\ m2 \end{bmatrix}$$



Factors, Levels, Groups, Classes

Continuous Variables/Factors: Age, IQ, Volume, etc

Discrete Variables/Factors: Gender, Handedness, Diagnosis
Levels of Discrete :
- Handedness: Left and Right
- Gender: Male and Female
- Diagnosis: Normal, MCI, AD

Group or Class: Specification of All Discrete Factors:
- Left-handed Male MCI
- Right-handed Female Normal

Factors, Levels, Groups

Each Group/Class:

- Has its own Intercept
- Has its own Slope (for each continuous variable)

NRegressors = NClasses*(NVariables+1)

Example: Thickness Study

1. $SUBJECTS_DIR/bert/surf/lh.thickness
2. $SUBJECTS_DIR/fred/surf/lh.thickness
3. $SUBJECTS_DIR/jenny/surf/lh.thickness
4. $SUBJECTS_DIR/margaret/surf/lh.thickness
5. …

FSGD File

FSGDF = FreeSurfer Group Descriptor File

- One Discrete Factor (Gender) with Two Levels (M&F)
- Three Continuous Variables: Age, Weight, IQ

```
GroupDescriptorFile 1
Class Male
Class Female
Variables                      Age  Weight  IQ
Input bert       Male          10   100     1000
Input fred       Male          15   150     1500
Input jenny      Female        20   200     2000
Input margaret   Female        25   250     2500
```

Class = Group

Note: Can specify design matrix explicitly with --design

FSGDF à X (Automatic)

Female Group
Male Group
Male Age
Female Age

$$X = \left[\begin{array}{cc|cc|cc|cc} 1 & 0 & 10 & 0 & 100 & 0 & 1000 & 0 \\ 1 & 0 & 15 & 0 & 150 & 0 & 1500 & 0 \\ 0 & 1 & 0 & 20 & 0 & 200 & 0 & 2000 \\ 0 & 1 & 0 & 25 & 0 & 250 & 0 & 2500 \end{array}\right]$$

Age Weight IQ

$$C = [-1\ 1\ 0\ 0\ 0\ 0\ 0\ 0]$$

Tests for the difference in intercept/offset between groups

$$C = [0\ 0\ -1\ 1\ 0\ 0\ 0\ 0]$$

Tests for the difference in age slope between groups

DODS – Different Offset, Different Slope

mri_glmfit

- Reads in FSGD File and constructs X
- Reads in your contrasts (C1, C2, etc)
- Loads lh.thickness.sm10.mgh
- Fits GLM (ie, computes β)
- Computes contrasts (γ=C*β)
- t or F ratios, significances
- Significance -log10(p) (.01 à 2, .001 à 3)

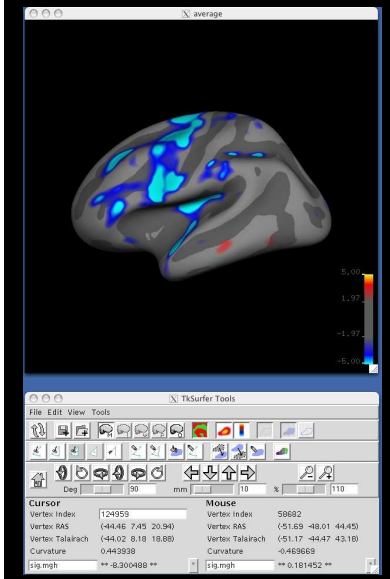
mri_glmfit

```
mri_glmfit
 --y lh.thickness.sm10.mgh
 --fsgd gender_age.txt doss
 --surf fsaverage lh
 --glmdir lh.gender_age.glmdir
 --C age.mat –C gender.mat

mri_glmfit --help
```

Creates: lh.gender_age.glmdir/
- beta.mgh – parameter estimates
- eres.mgh – residual error
- rvar.mgh – residual error variance
- etc …
- age/
  - sig.mgh – -log10(p)
  - gamma.mgh, F.mgh
- gender/
  - sig.mgh – -log10(p)
  - gamma.mgh, F.mgh

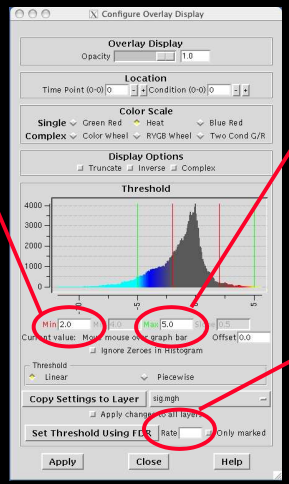
Visualization with tksurfer



File->LoadOverlay

Threshold:
-log10(p),
Eg, 2=.01

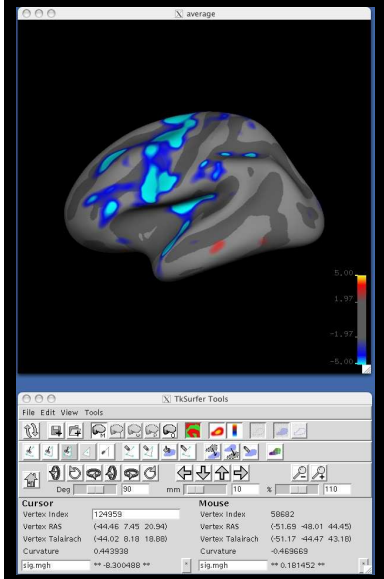
Saturation:
-log10(p),
Eg, 5=.00001

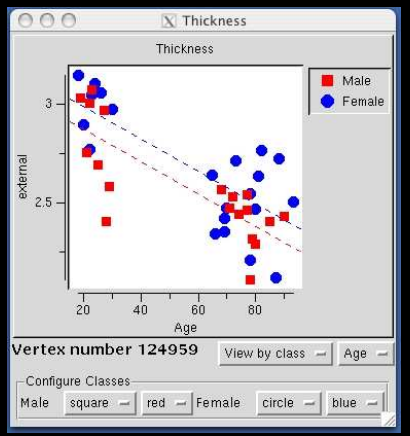


False
Dicovery
Rate
Eg, .01

View->Configure->Overlay

Visualization with tksurfer





File->
Load Group Descriptor File …

Another FSGD Example

- Two Discrete Factors
  - Gender: Two Levels (M&F)
  - Handedness: Two Levels (L&R)
- One Continuous Variable: Age

```
GroupDescriptorFile 1
Class MaleRight
Class MaleLeft
Class FemaleRight
Class FemaleLeft
Variables                        Age
Input bert        MaleLeft       10
Input fred        MaleRight      15
Input jenny       FemaleRight    20
Input margaret    FemaleLeft     25
```

Class = Group

QDEC – An Interactive Statistical Engine GUI

Query – Select subjects based on Match Criteria
Design – Specify discrete and continuous factors
Estimate – Fit Model
Contrast – Automatically Generate Contrast Matrices

Interactive – data caching (recon-all –qcache)

…a work in progress
- No Query yet
- Two Discrete Factors (Two Levels)
- Two Continuous Factors
- Surface only

QDEC – Spreadsheet

qdec.table.dat – spreadsheet with subject information – spreadsheet can be huge!

fsid	gender	age	diagnosis	Left-Cerebral-White-Matter-Vol
011121_vc8048	Female	70	Demented	202291
021121_62313-2	Female	71	Demented	210188
010607_vc7017	Female	73	Nondemented	170653
021121_vc10557	Male	75	Demented	142029
020718_62545	Male	76	Demented	186087
020322_vc8817	Male	77	Nondemented	149810

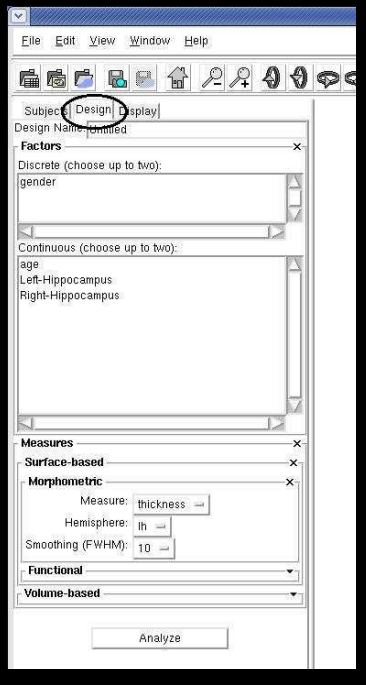
gender.levels

```
Female
Male
```

diagnosis.levels

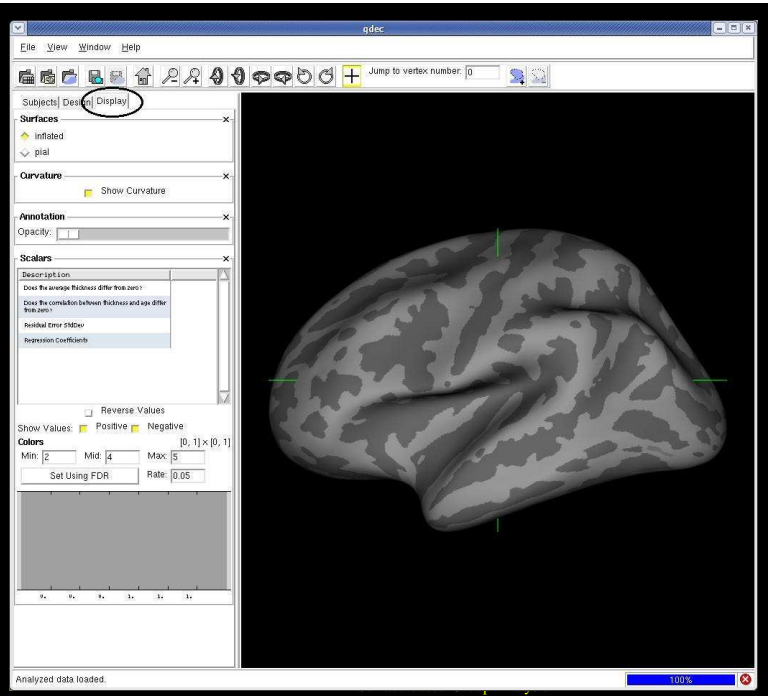
```
Demented
Nondemented
```

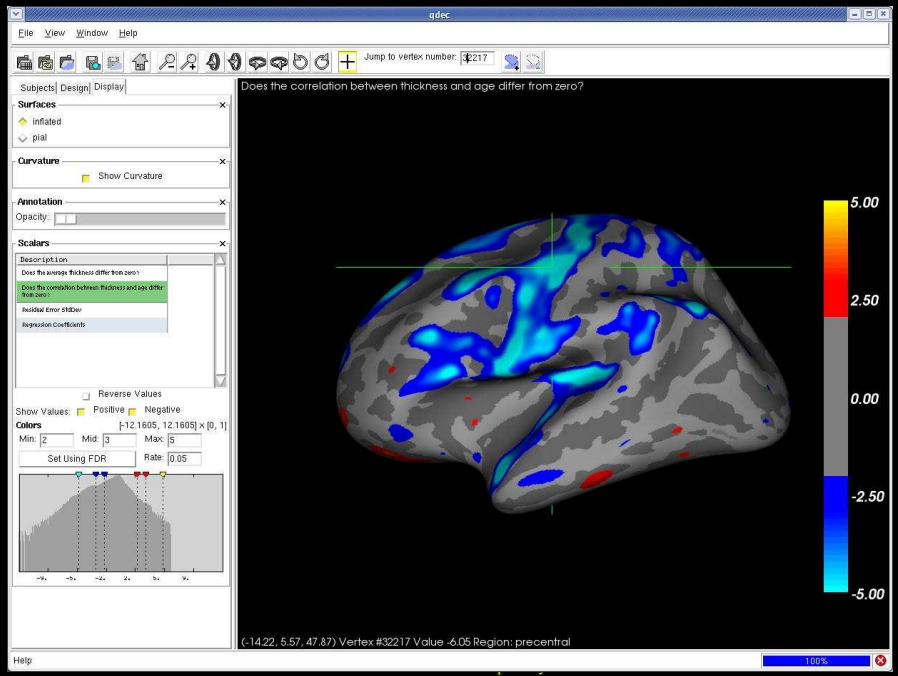
Discrete Factors need a factorname.level file



QDEC GUI

- Load QDEC Table File
  - List of Subjects
  - List of Factors (Discrete and Cont)
- Choose Factors
- Choose Input (cached):
  - Hemisphere
  - Measure (eg, thickness)
  - Smoothing Level
- "Analyze"
  - Builds Design Matrix
  - Builds Contrast Matrices
  - Constructs Human-Readable Questions
  - Analyzes
  - Displays Results





Correction for Multiple Comparisons

- False Discovery Rate (FDR) – built into tksurfer. (Genovese, et al, NI 2002)
- Cluster-based
  - Monte Carlo simulation.
  - Permutation Tests
  - Surface Gaussian Random Fields (GRF) – coming "soon"

Cluster-based Correction for Multiple Comparisons

1. Simulate data under Null Hypothesis:
   - Synthesize Gaussian noise and then smooth (Monte Carlo)
   - Permute rows of design matrix (Permutation, orthog)
2. Analyze, threshold, cluster, max cluster size
3. Repeat 10000 times
4. Analyze real data, get cluster sizes
5. P(cluster) = #MaxClusterSize > ClusterSize/10000

mri_glmfit …. –sim nulltype nsim thresh csdbasename
  nulltype: perm, mc-full, mc-z
  nsim: 10000
  thresh: vertex-wise threshold (-log10(p)), eg, p=.01, thresh=2
  csdbasename: file to save cluster simulation data (CSD)
mri_surfcluster

Practical

- Thickness study of aging
  - 40 subjects
  - Age and Gender
- Create QDEC table and level file
- Play with data

FSGDF à X

Input:
- y
- X
- C

DOSS – Different Offset, Same Slope

Female Class

Male Class

Age for Males and Females

$$X = \begin{bmatrix} 1 & 0 & 10 & 100 & 1000 \\ 1 & 0 & 15 & 150 & 1500 \\ 0 & 1 & 20 & 200 & 2000 \\ 0 & 1 & 25 & 250 & 2500 \end{bmatrix}$$

$$C = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 \end{bmatrix}$$ à Same test, different vector

#Regressors = Nv+Nc = 3+2=5 à Fewer regressors than DODS
DOF = #Rows - #Regressors