

$$Y \sim \beta_0 + \beta$$

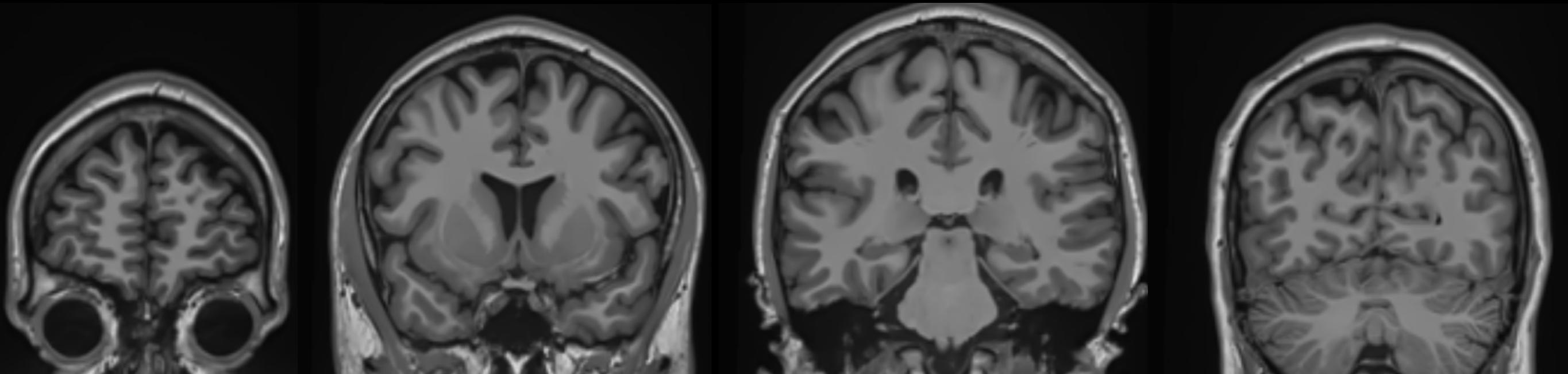
## WEEK 2: THE GENERAL LINEAR MODEL

Boris Bernhardt, PhD

Bratislav Misic, PhD

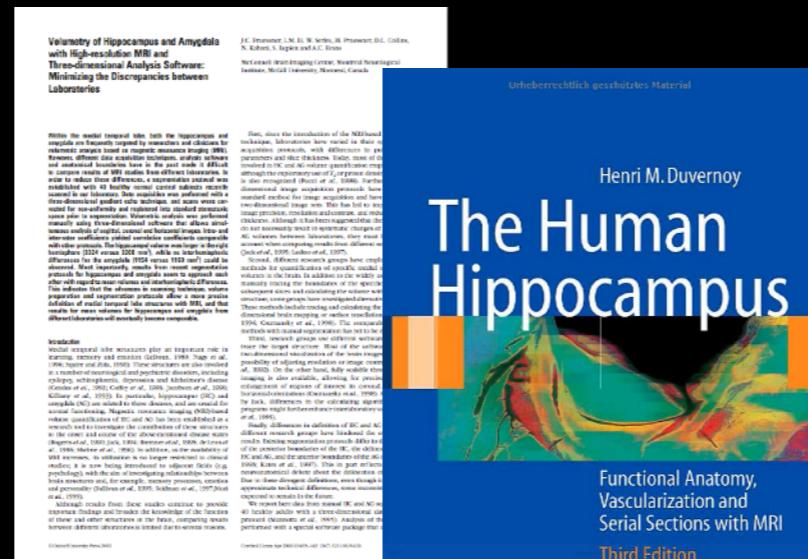
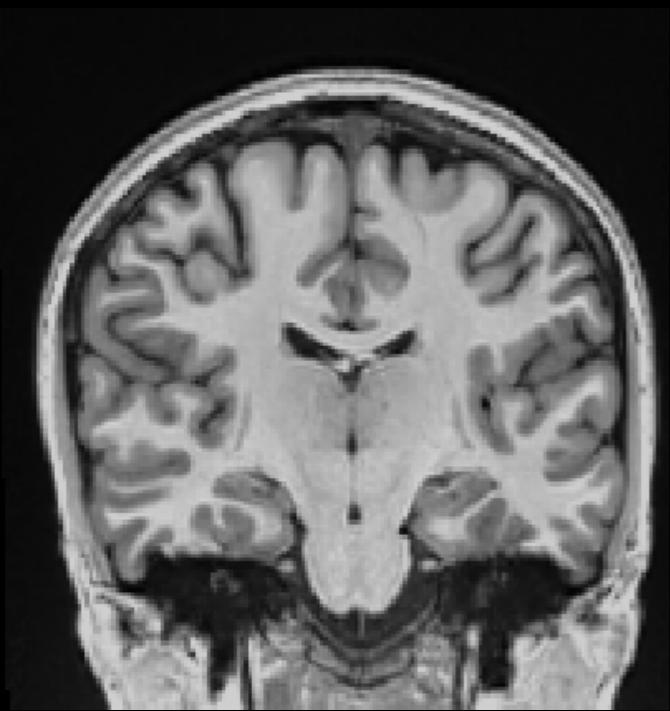


# structural MRI



T1-weighted MRI

# VOLUMETRY PIPELINE

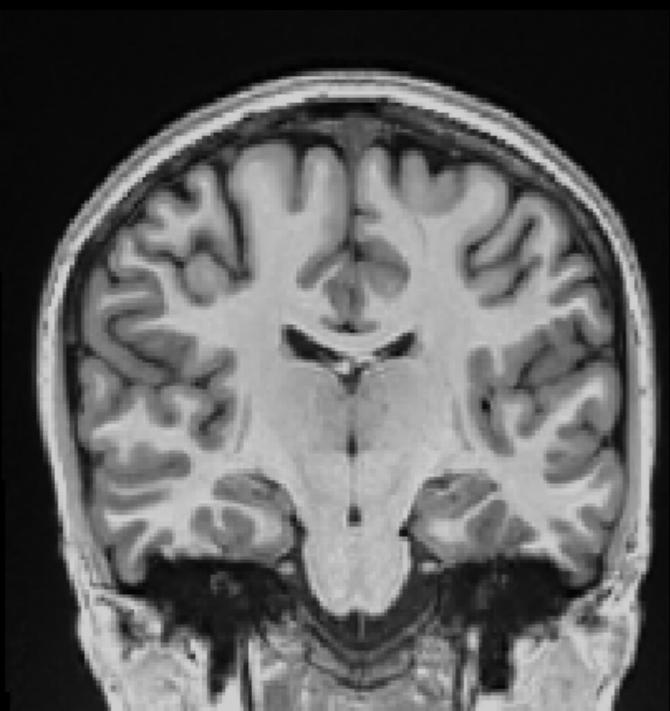


input

read and become expert

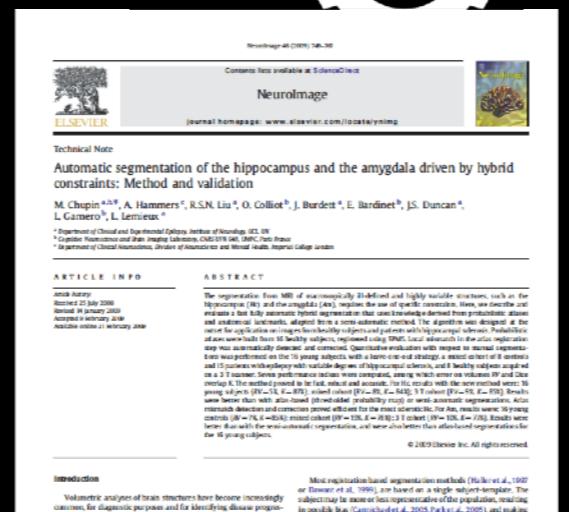
segment

# AUTOMATATION



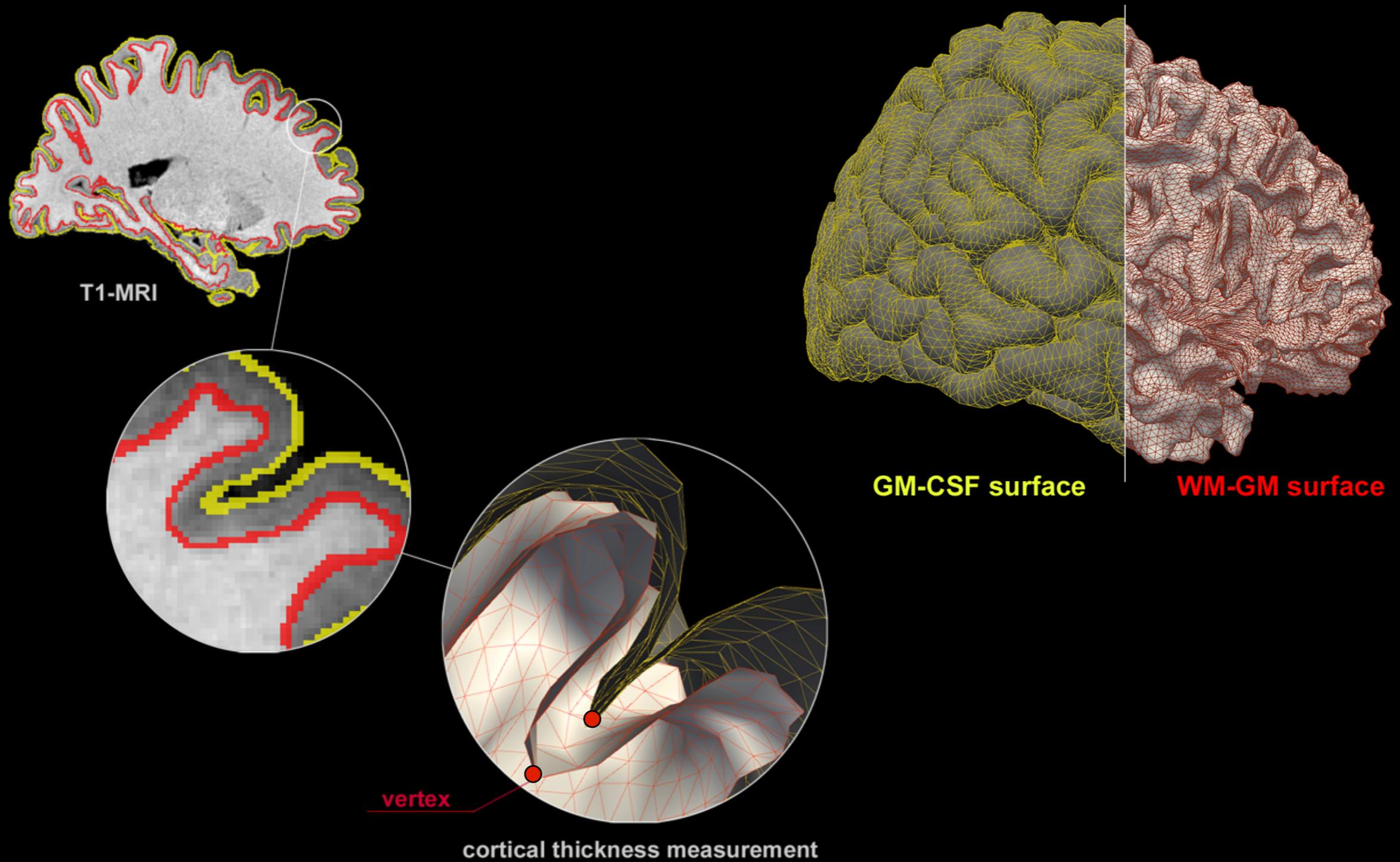
input

automatic segmentation approaches

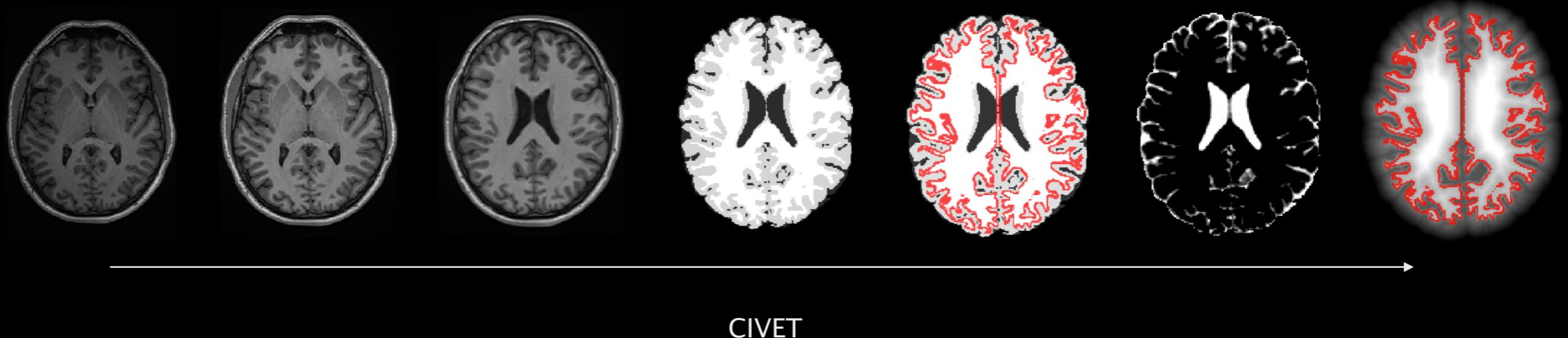


segment

# CORTICAL THICKNESS MEASUREMENT



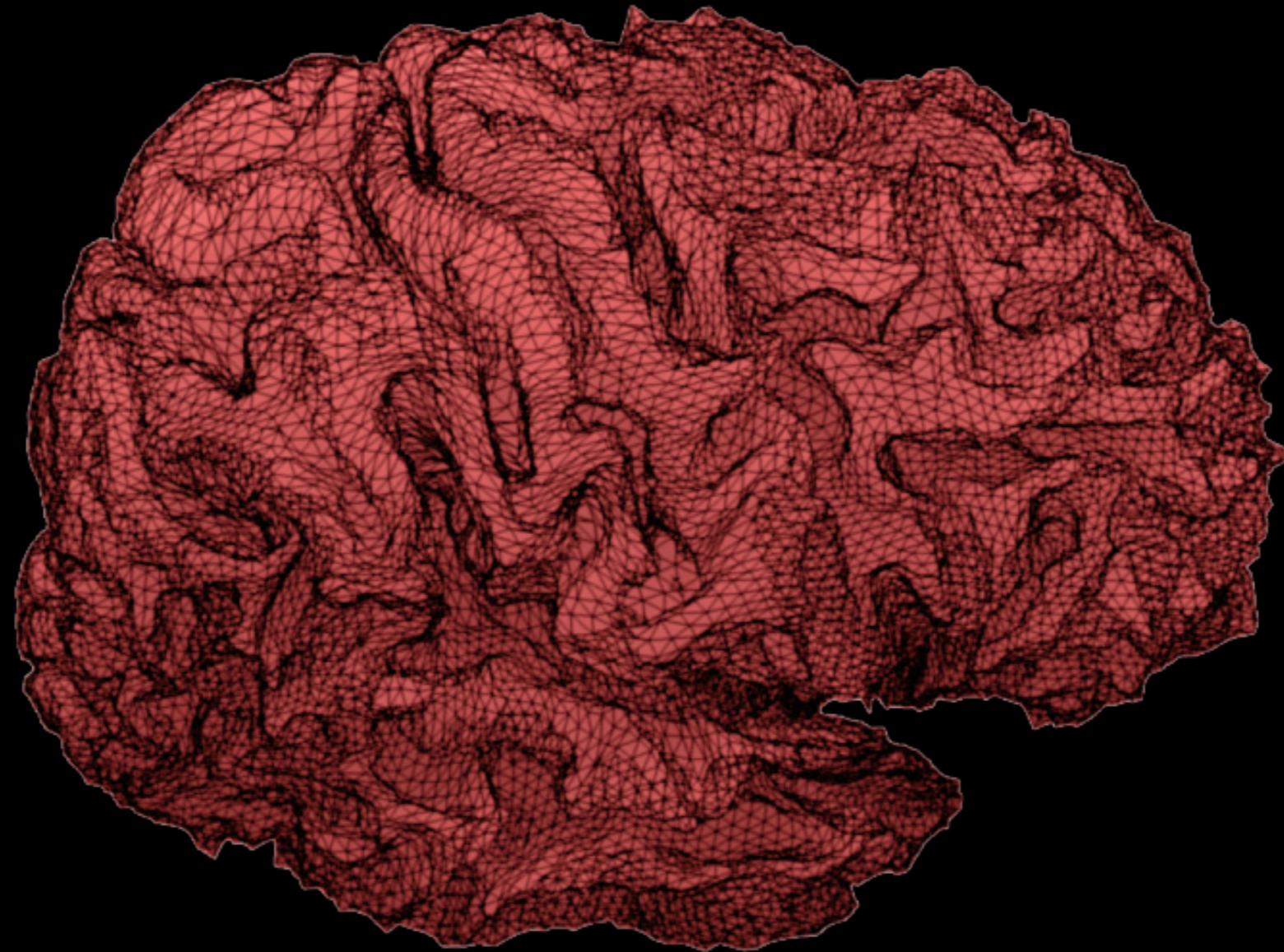
# PROCESSING PIPELINES



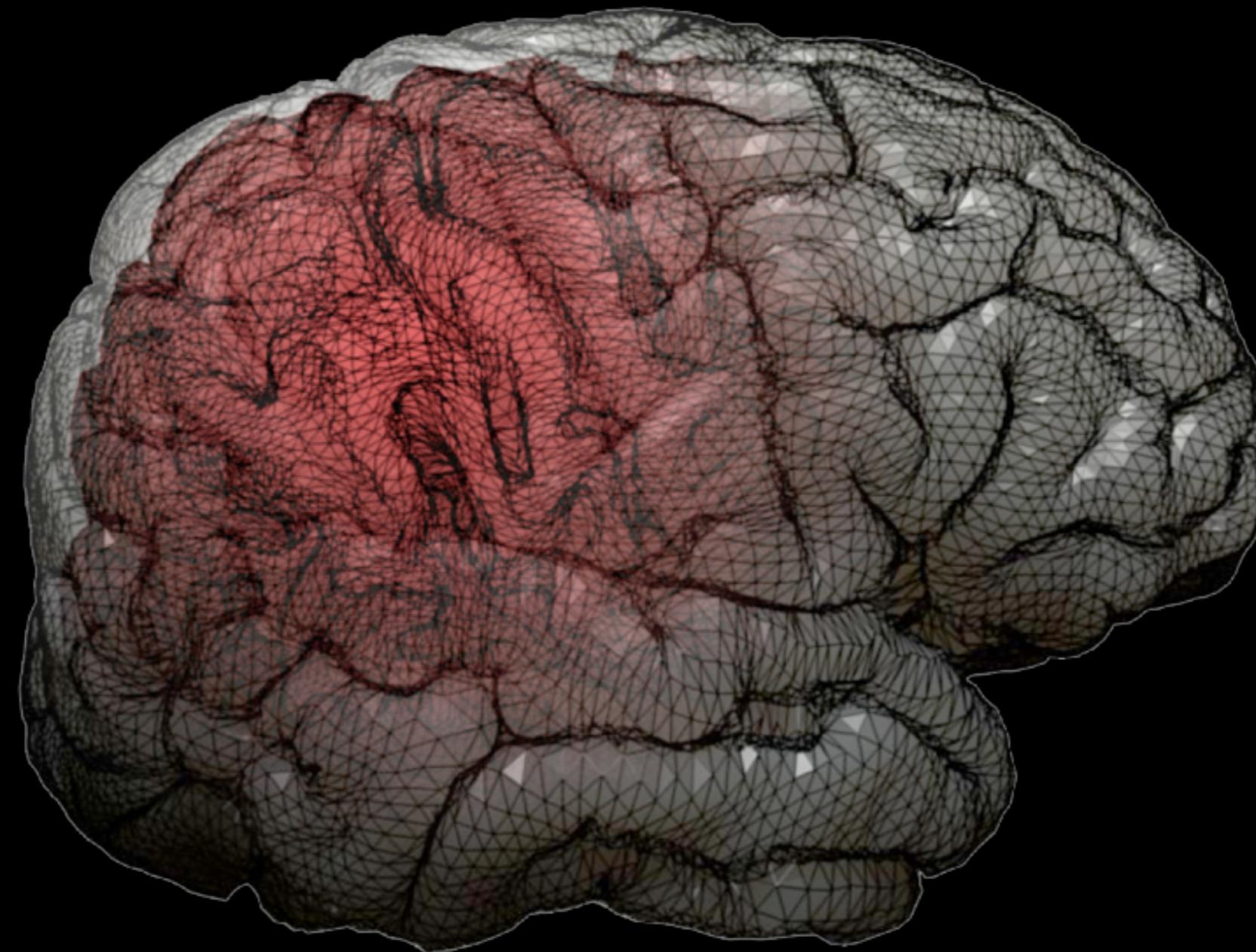
MacDonald et al. (2000) *NeuroImage*

Kim et al. (2005) *NeuroImage*

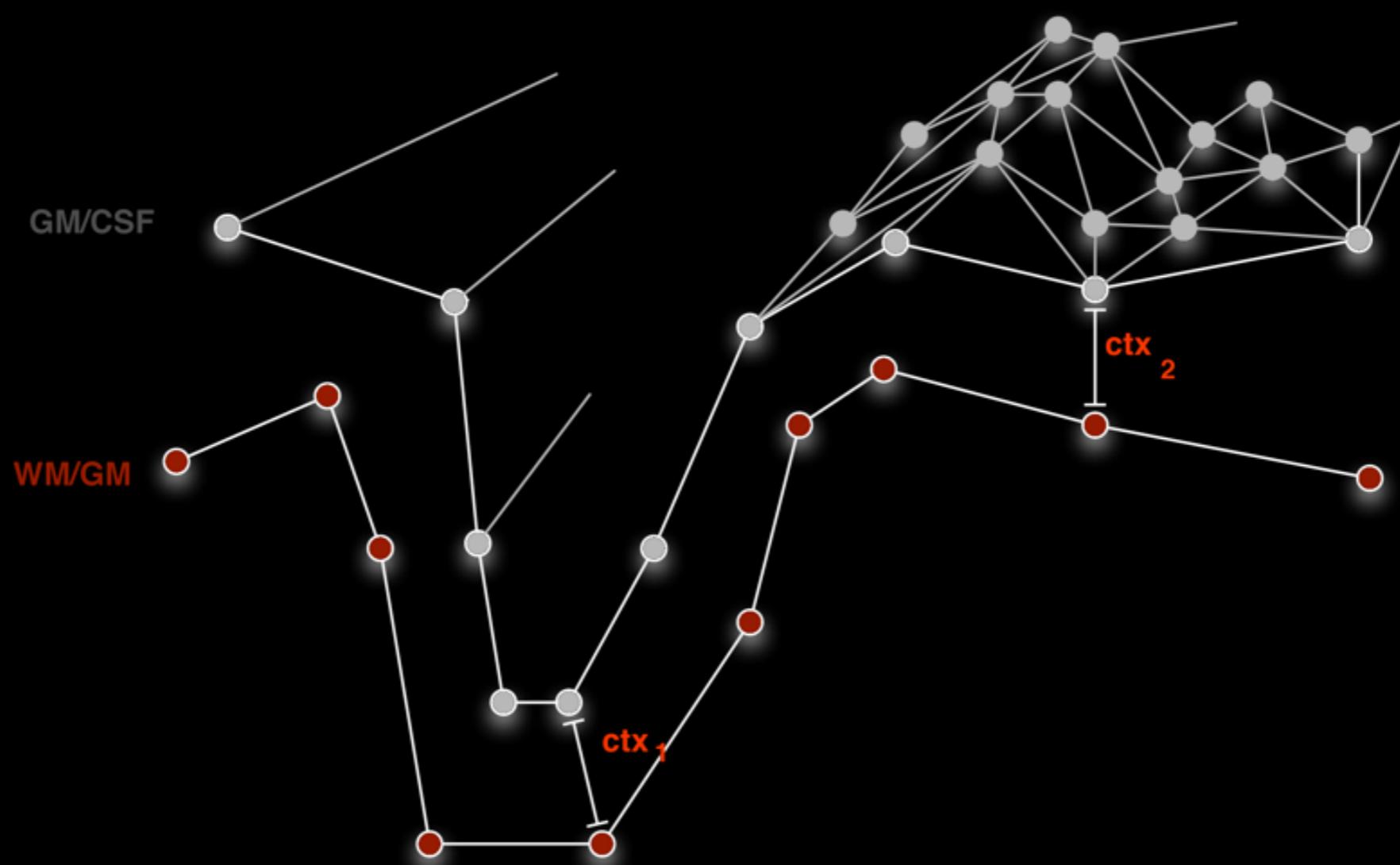
WM



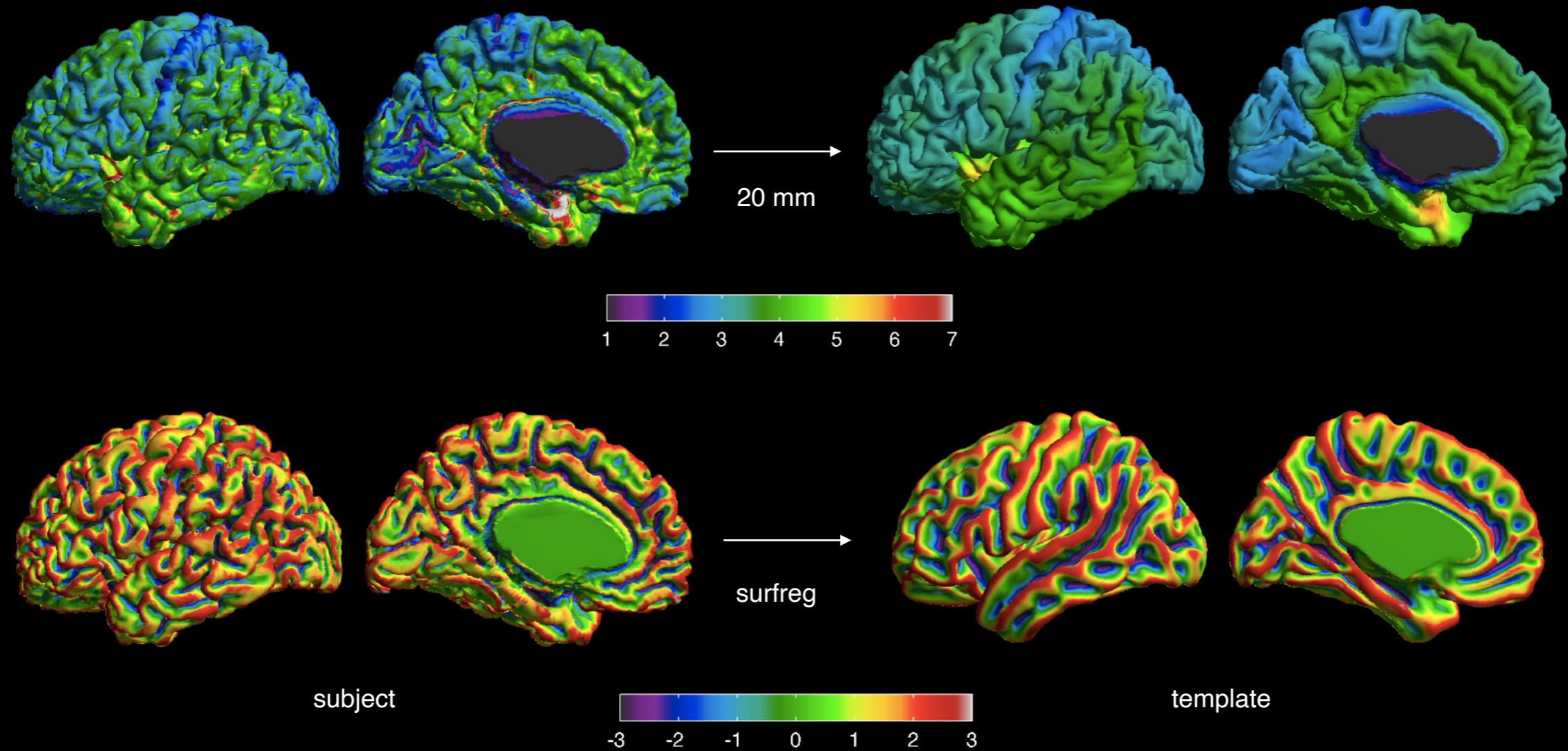
GM SURFACE



# MEASUREMENT OF THICKNESS



# SURFACE-BASED PROCESSING



Chung et al. (2003) NeuroImage  
Robbins et al. (2004) MedImaAnalysis

# NOW WE CAN FINALLY BUILD THE FIRST LINEAR MODELS

t-tests, correlations, partial correlations, ANOVAs, MANOVAs,...

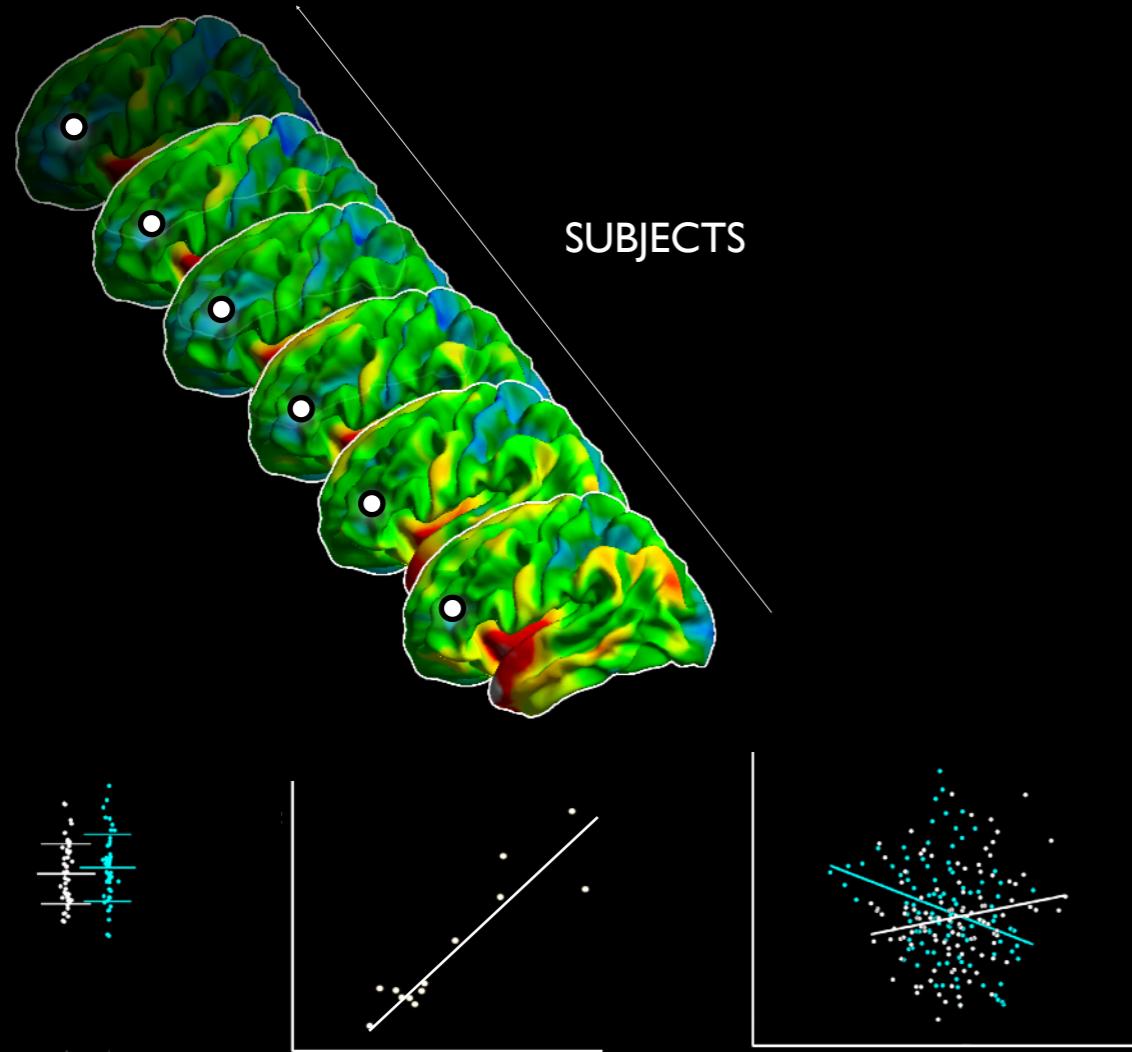
are just specific instances of the linear model of the form

$$Y \sim \beta_0 + \beta_1 * x_1 + \beta_2 * x_2 + \beta_3 * x_1 * x_2 \dots + \varepsilon$$

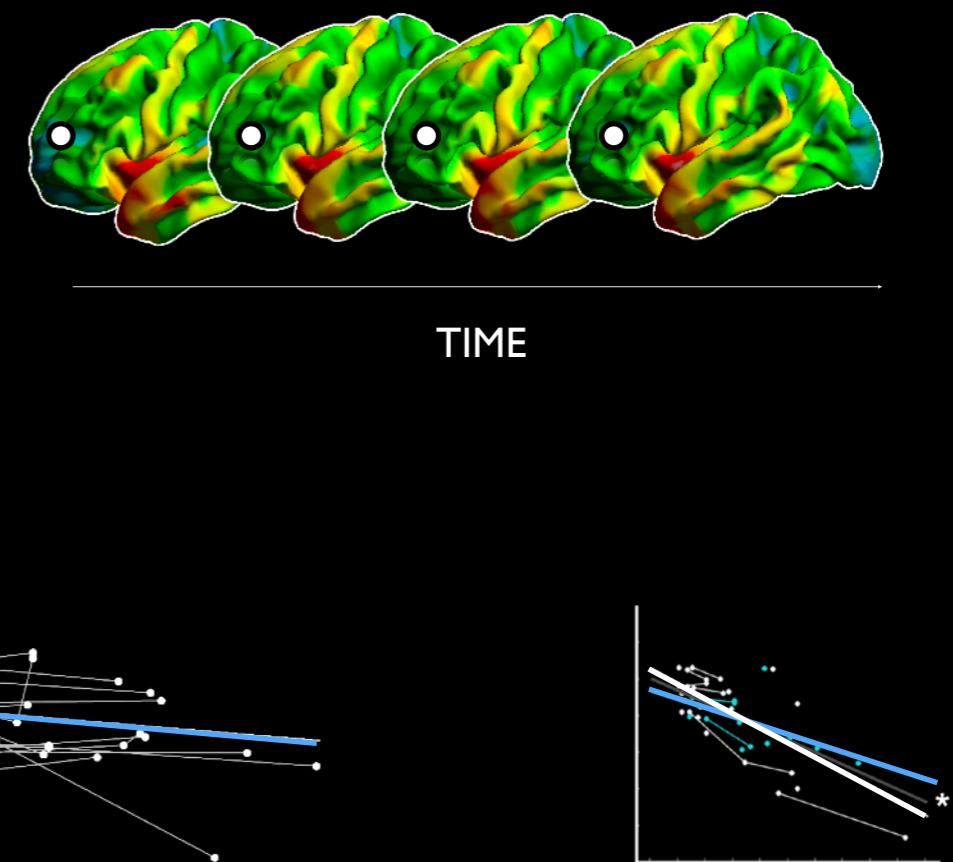
DATA                    INTERCEPT                    SIMPLE EFFECTS                    INTERACTIONS

# NOW WE CAN FINALLY BUILD THE FIRST LINEAR MODELS

CROSS-SECTIONAL ANALYSES

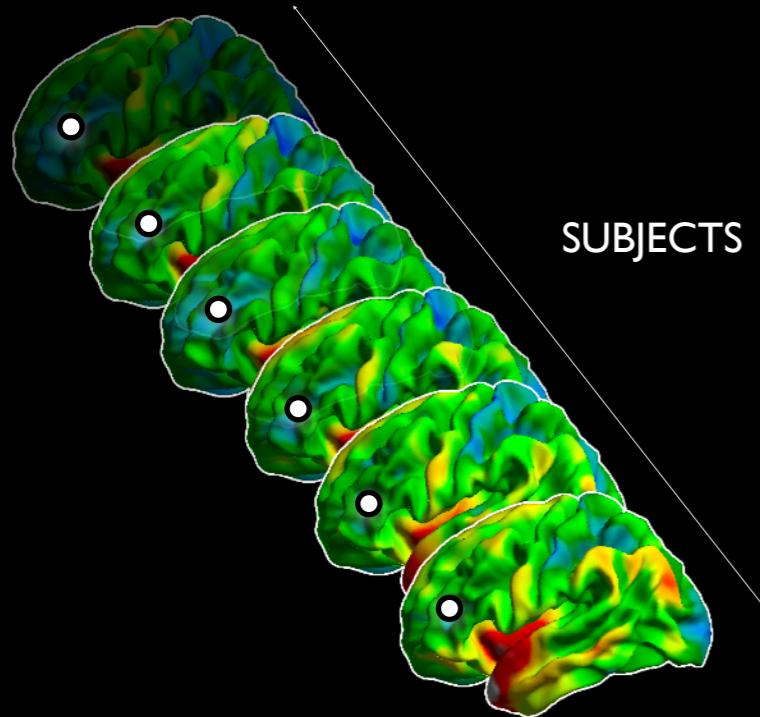


LONGITUDINAL ASSESSMENTS

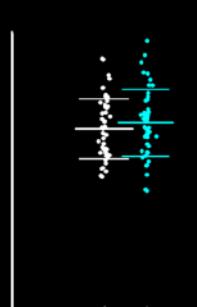


# NOW WE CAN FINALLY BUILD THE FIRST LINEAR MODELS

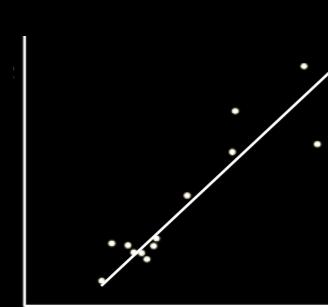
CROSS-SECTIONAL ANALYSES



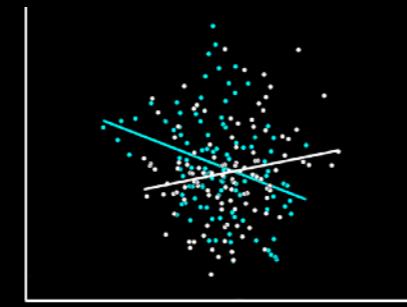
SUBJECTS



$$Y = I + G$$

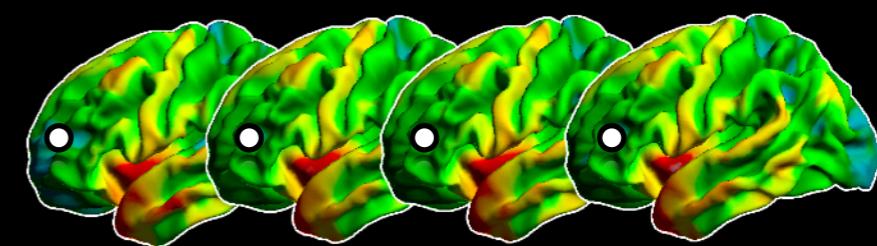


$$Y = I + A$$

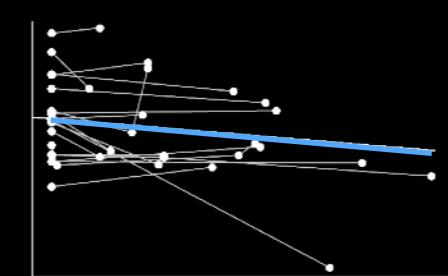


$$Y = I + G + A + G \times A$$

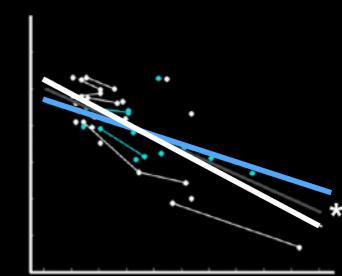
LONGITUDINAL ASSESSMENTS



TIME



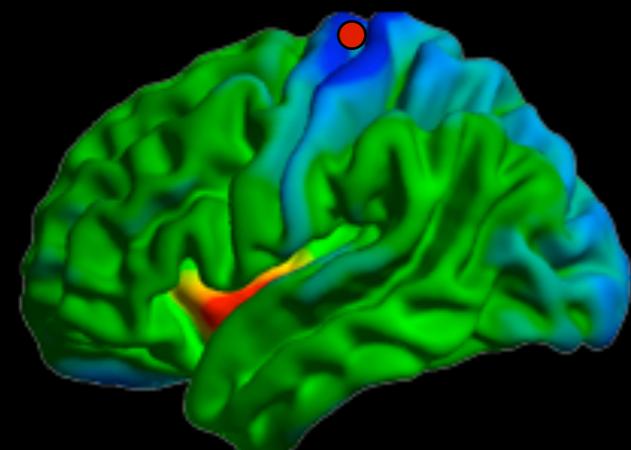
$$Y = I + r(S) + ISI$$



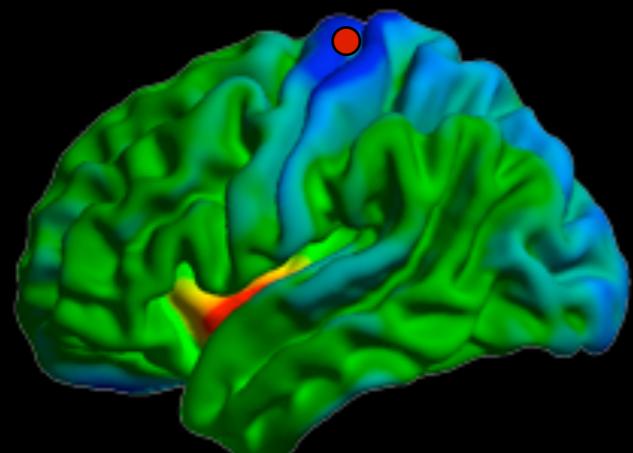
$$Y = I + r(S) + ISI + G + ISI \times G$$

# SURFSTAT

Controls

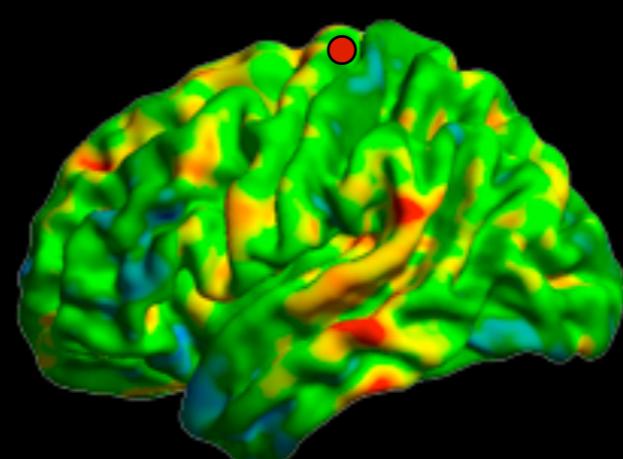


Patients

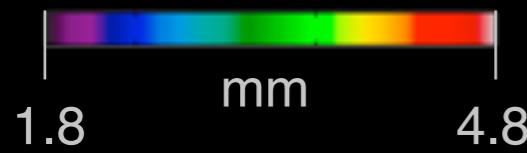
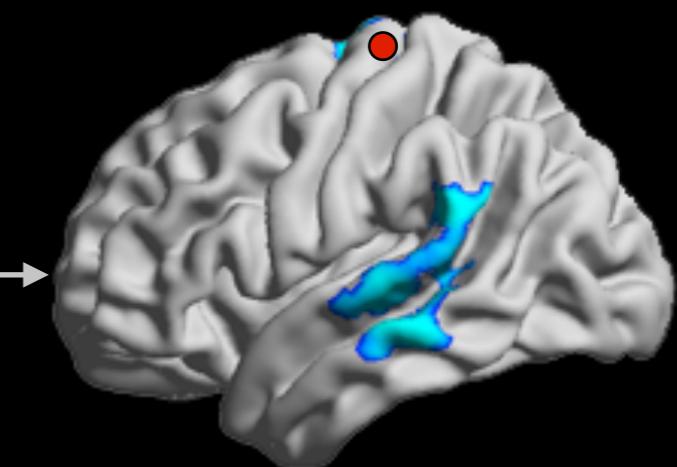


GLM

t-map



p-values



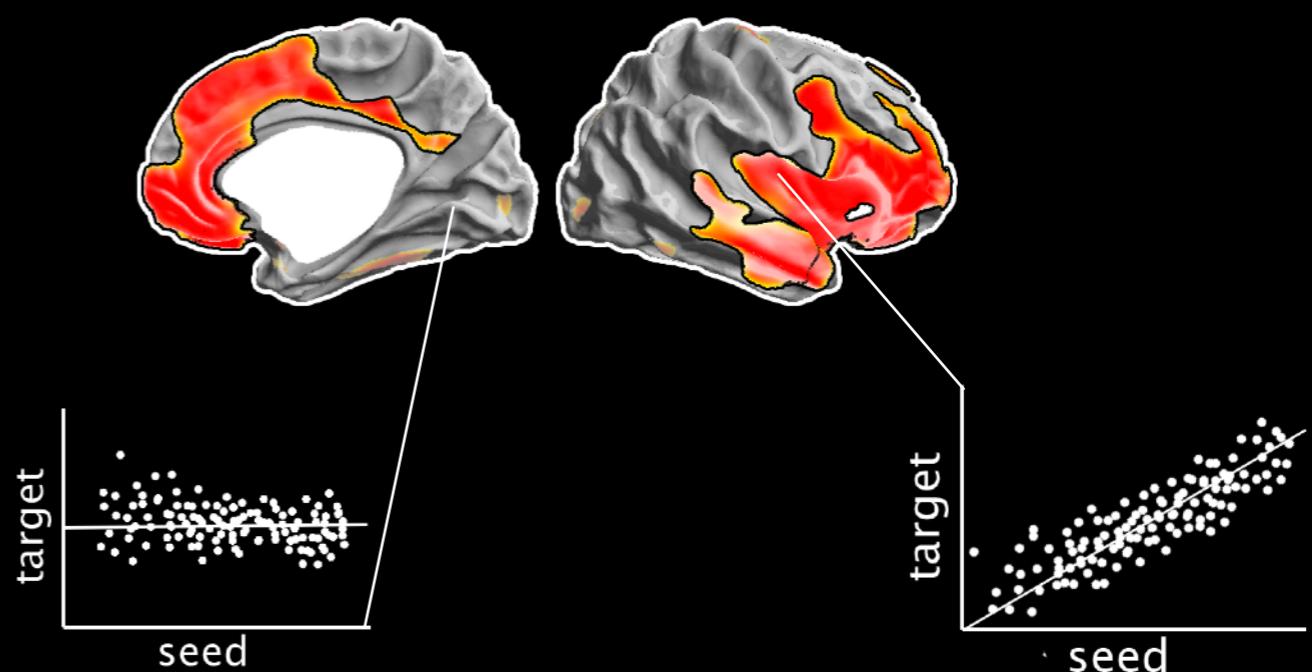
# MRI COVARIANCE ANALYSIS

A SPECIAL CASE OF A GENERALIZED LINEAR MODEL APPLICATION

THE IDEA:  
VARIABLES OF INTEREST ARE MORPHOLOGICAL MEASURES

SINGLE SEED CORRELATIONS POSSIBLE  
 $MODEL = 1 + SEED$

GROUP COMPARISONS VIA INTERACTIONS  
 $MODEL = 1 + SEED + GROUP + GROUP*SEED$



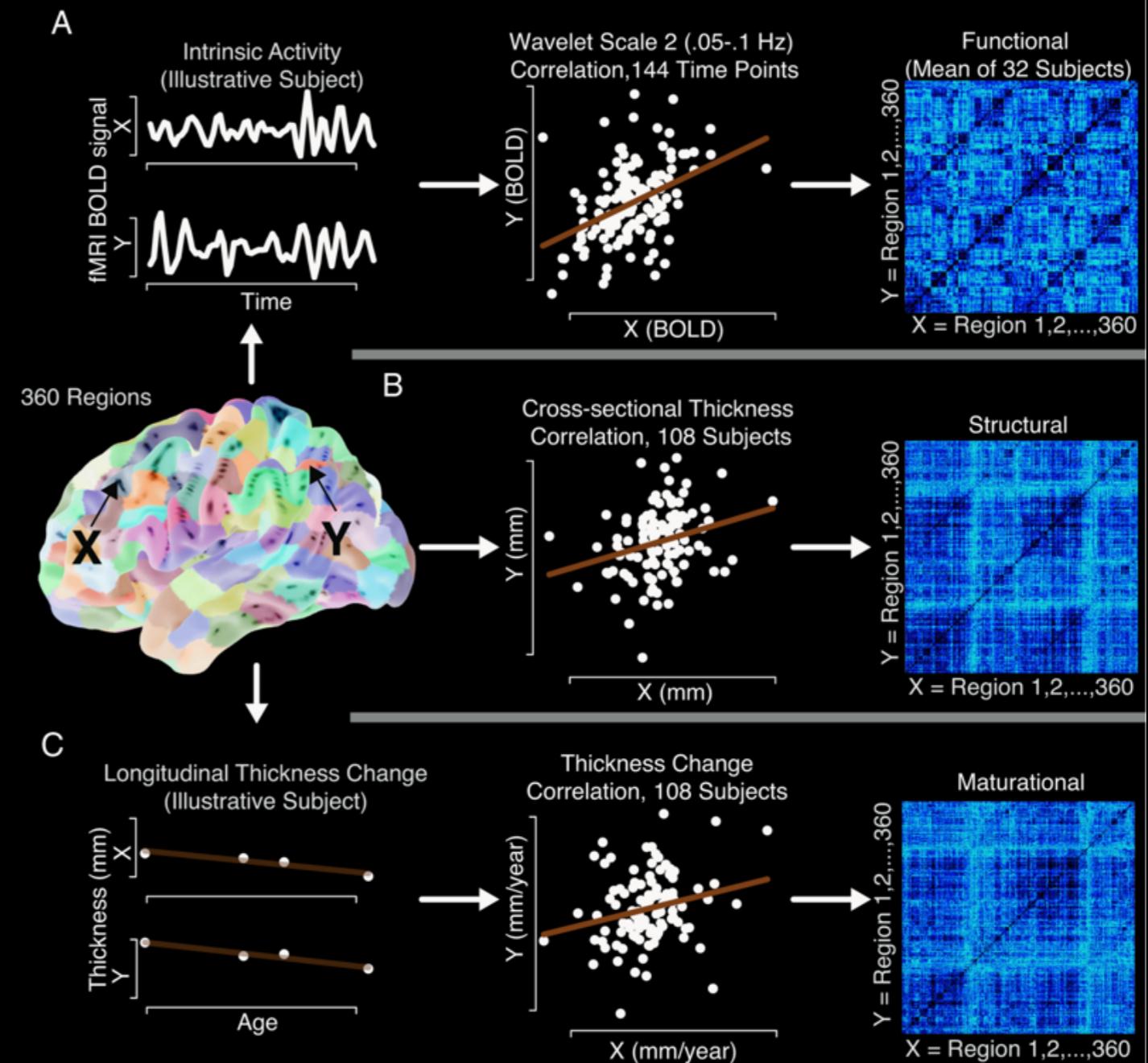
# MRI COVARIANCE ANALYSIS

AN EXTENSION OF THE COVARIANCE FRAMEWORK MAY BE A ‘MATURATIONAL COUPLING ANALYSIS’

IDEA:  
INSTEAD OF USING CROSS-SECTIONAL MORPHOLOGICAL MEASURES,  
LONGITUDINAL TRAJECTORIES ARE CORRELATED

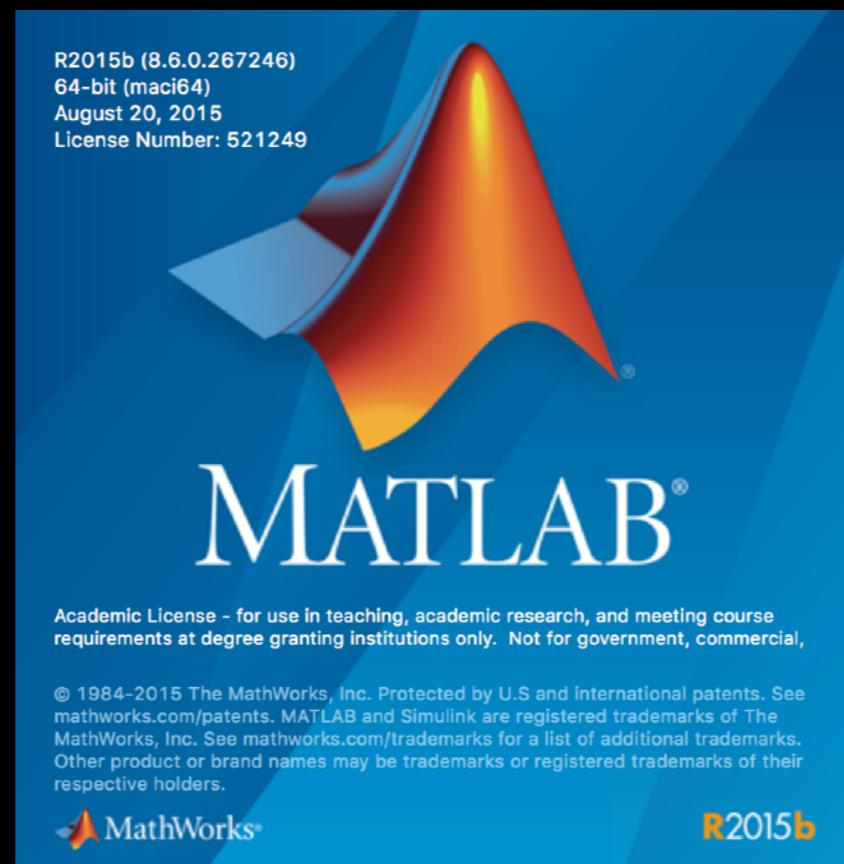
BOTH APPROACHES GIVE SIMILAR FINDINGS.  
CROSS-SECTIONAL COVARIANCE MAYBE MORE GENERAL, BUT DEPENDS ON AGE

LONGITUDINAL MATURATION ALSO DEPENDS ON AGE AND AGE-WINDOW, BUT MAY MORE SPECIFICALLY PROBE A CERTAIN PROCESS



# TO GET STARTED

LAUNCH MATLAB



# TO GET STARTED

Download surfstat tutorial from

<https://github.com/MICA-MNI/micaopen>

And do this:

```
P = '/Users/boris/Documents/1_github/micaopen/surfstat/surfstat_tutorial/'  
addpath(genpath(P));  
cd(P);
```

## LOAD SURFACES

```
%% 1. Load the surface data
SP = SurfStatAvSurf({[P 'fsaverage5/lh.pial'],[P 'fsaverage5/rh.pial']})
SW = SurfStatAvSurf({[P 'fsaverage5/lh.white'],[P 'fsaverage5/rh.white']})
```

## WHAT ARE THESE SURFACES: SW AND SP

```
>> SW
```

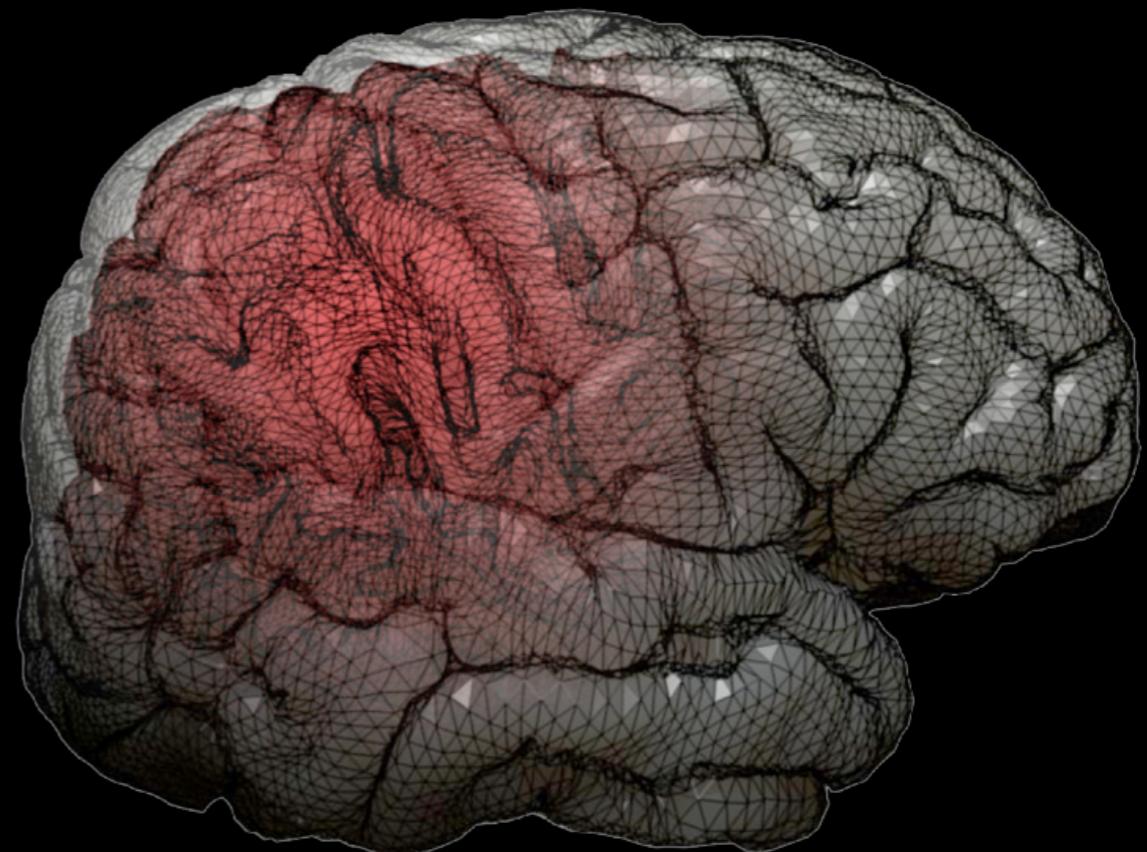
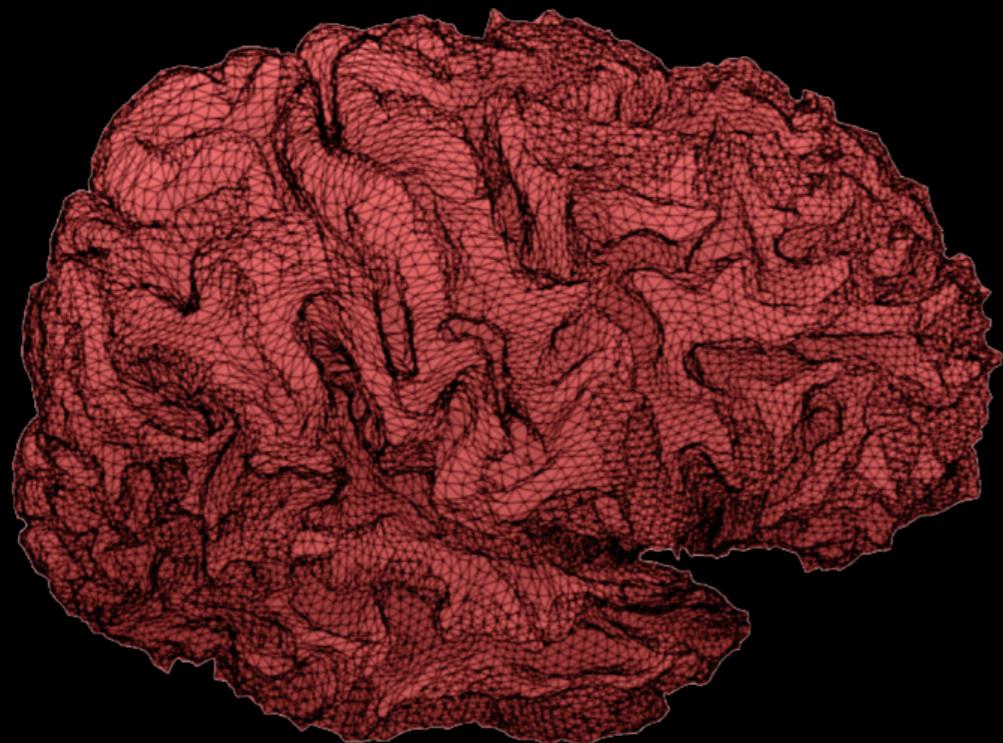
```
SW =
```

```
tri: [40960x3 int32]  
coord: [3x20484 double]
```

```
>> SP
```

```
SP =
```

```
tri: [40960x3 int32]  
coord: [3x20484 double]
```



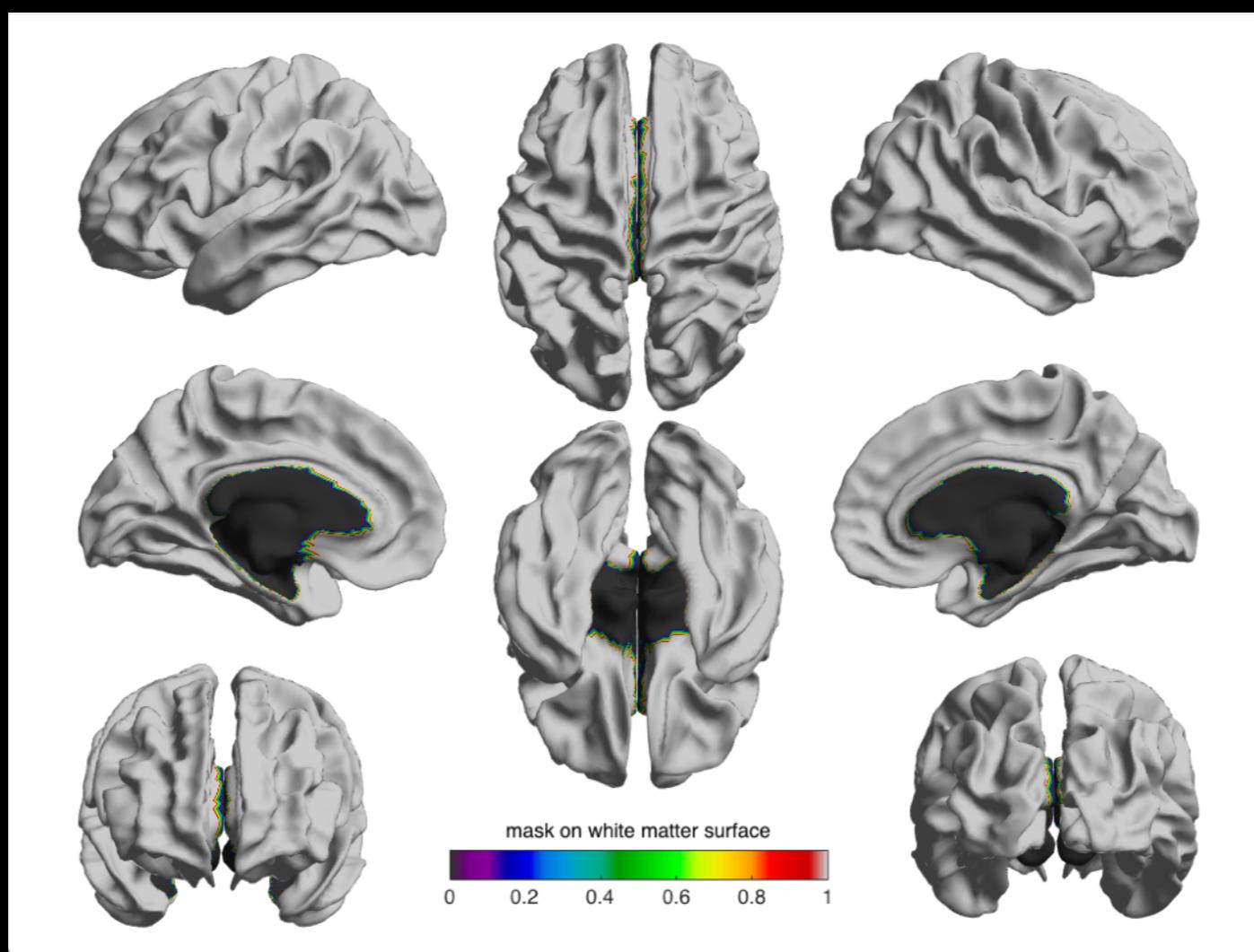
## LOAD SOME USEFUL FEATURES

```
%% 2. Load brain mask and some useful surface features
load([P 'fsaverage5/mask.mat']),
load([P 'fsaverage5/curv.mat'])
```

## FIRST DISPLAY

```
%% 3. Display what we have been loading  
% first the brain mask  
f=figure,  
SurfStatViewData(double(mask),SW,'mask on white matter surface')
```

# BRAIN MASK



## ON DIFFERENT SURFACES

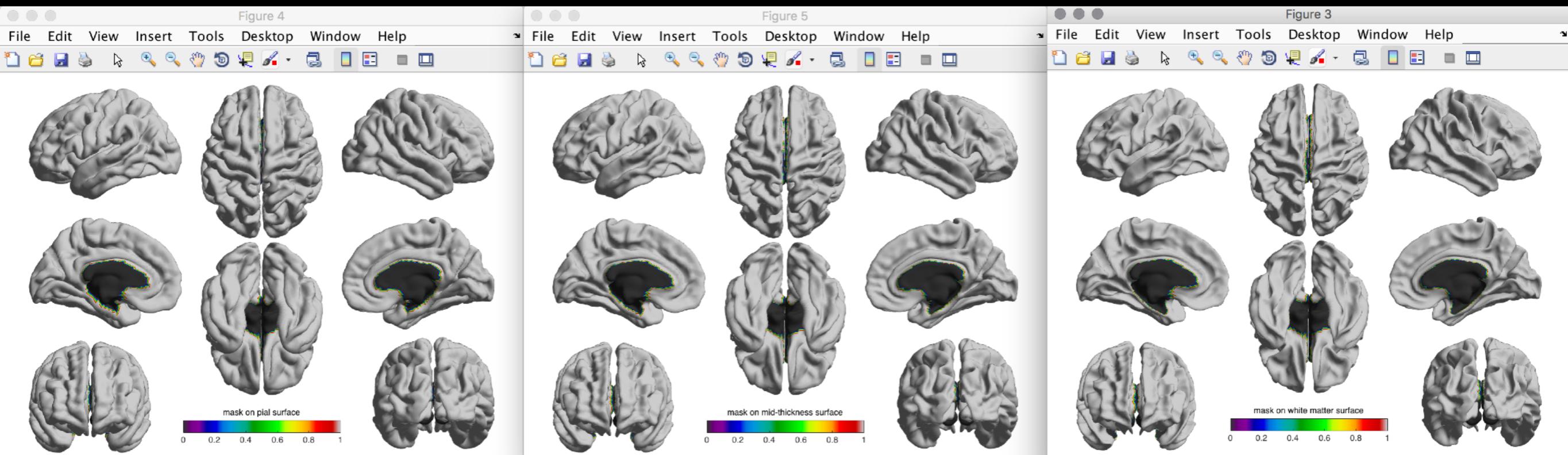
```
%% 3. Display what we have been loading

% first the brain mask
f=figure,
SurfStatViewData(double(mask),SW, 'mask on white matter surface')

f=figure,
SurfStatViewData(double(mask),SP, 'mask on pial surface')

f=figure,
SurfStatViewData(double(mask),SM, 'mask on mid-thickness surface')
```

# ON DIFFERENT SURFACES

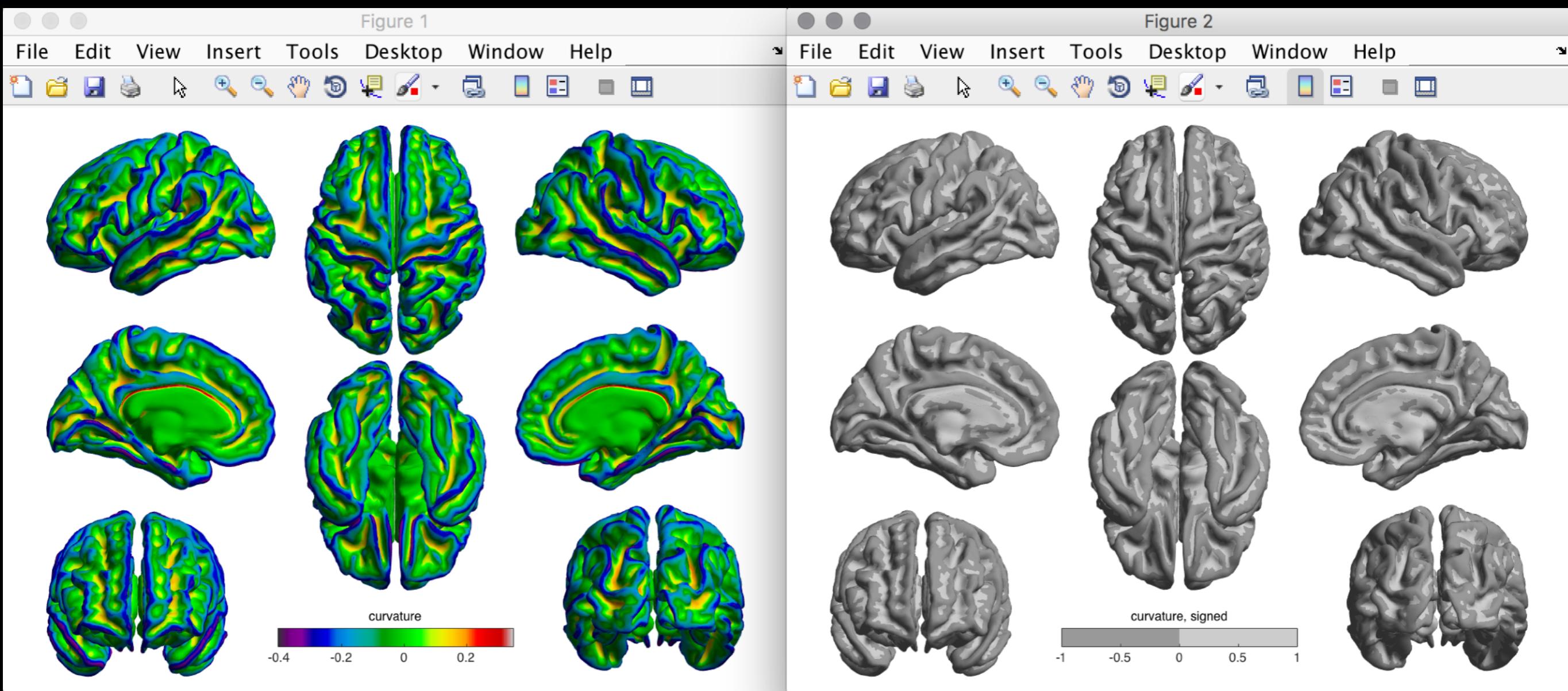


## WE CAN ALSO DISPLAY THE OTHER DATA

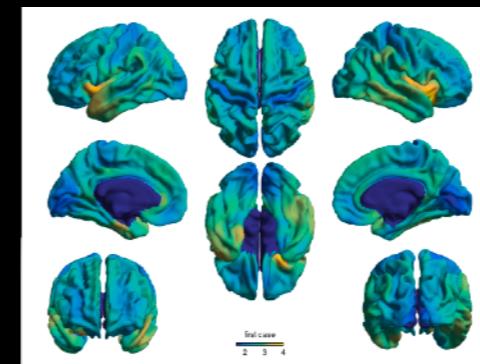
```
%% then the curvature data
f=figure,
SurfStatViewData(curv, SM, 'curvature')

% and now a binarized colormap
f=figure,
SurfStatViewData(sign(curv), SM, 'curvature, signed')
colormap([0.6 .6 .6; .8 .8 .8])
```

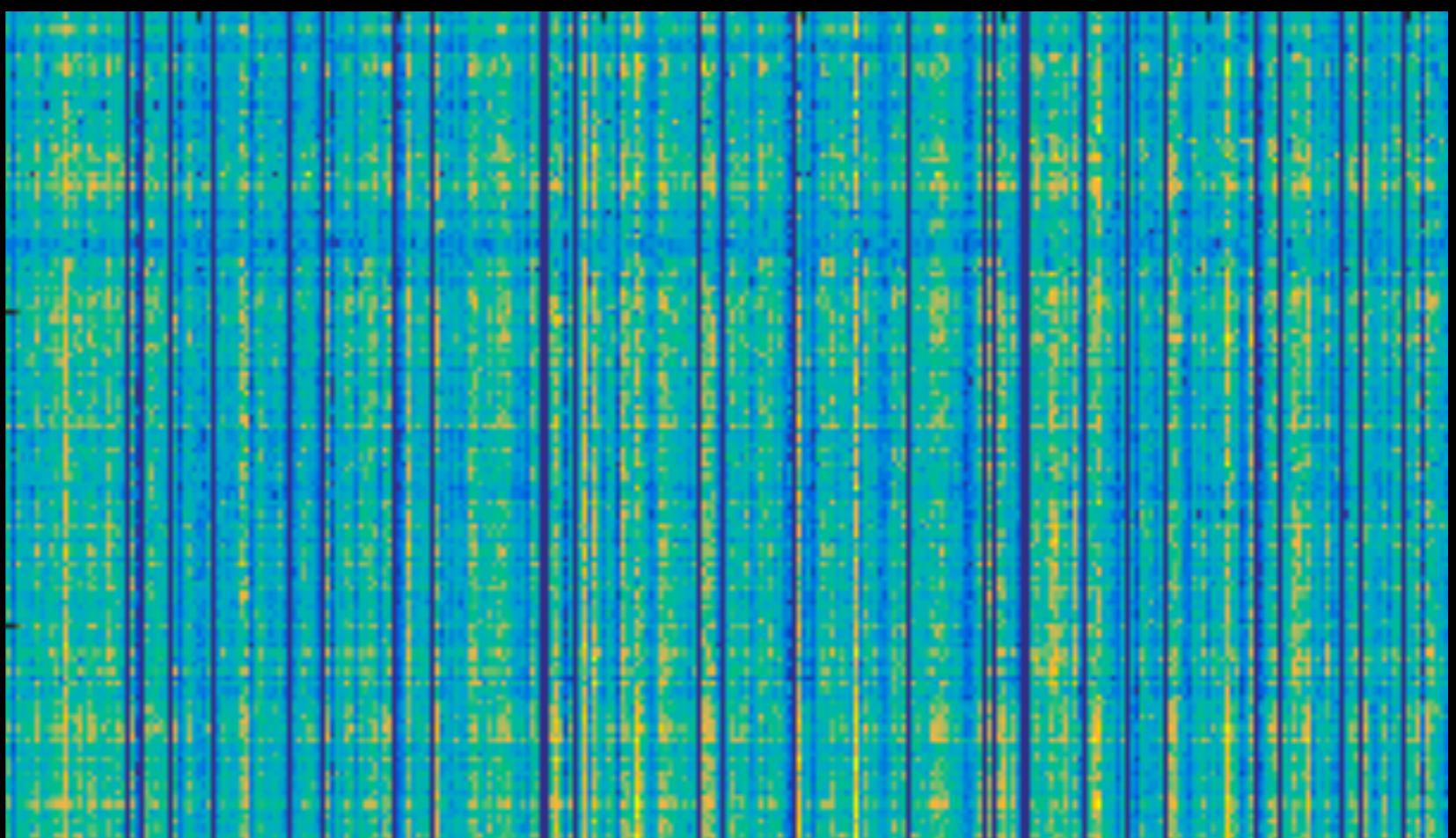
## WE CAN ALSO DISPLAY OTHER DATA



# NOW ITS TIME TO LOAD THE DATA OF THE INDIVIDUAL SUBJECTS



ID2	GROUP	AGE	HAND	IQ
33345	Group1	26	R	110
33346	Group1	19	R	93
33347	Group1	33	R	118
33348	Group1	31	R	110
33350	Group1	29	L	122
33351	Group1	30	R	118
33352	Group1	52	R	133
33354	Group1	42	R	114
33355	Group1	22	R	114
33356	Group1	18	R	95
33359	Group2	23	R	125
33362	Group2	23	R	122
33363	Group2	26	R	114
33364	Group2	30	R	110
33365	Group2	29	R	110
33366	Group2	29	R	110
33367	Group2	29	R	110
33368	Group2	23	R	107
33369	Group2	30	R	118
33370	Group2	22	R	118
33371	Group2	35	R	107
33372	Group2	25	R	118
33373	Group2	24	R	129
33375	Group1	11	R	105
33376	Group1	11	R	79



# NOW ITS TIME TO LOAD THE DATA OF THE INDIVIDUAL SUBJECTS

```
%% 4. ready for some analysis: load the spreadsheet
% load csv file that contains our participant ids, groups and IVs
fid = fopen([P 'myStudy.csv']); % final group
C = textscan(fid, '%s%s\n%s\n', 'Delimiter', ',', ...
    'headerLines', 1, 'CollectOutput', 1);
fclose(fid);

% we have to do a little bit of recoding
ID = C{1}(:,1);
GR = C{1}(:,2);
AGE = C{2};
HAND = C{3};
IQ = C{4};

%% 4b. Load the thickness data
% generate the file names
left = strcat(P, 'thickness/', ID, '_lh2fsaverage5_20.mgh');
right = strcat(P, 'thickness/', ID, '_rh2fsaverage5_20.mgh');

% load data into a matrix
T = SurfStatReadData([left, right]);
```



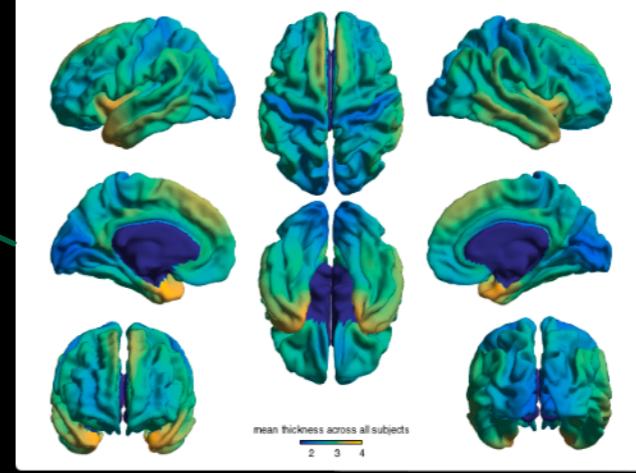
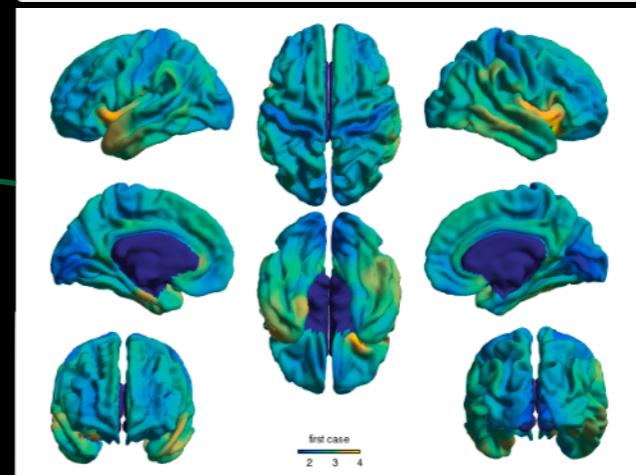
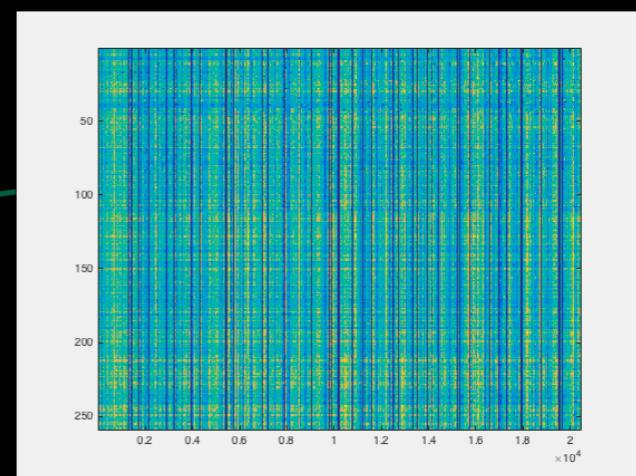
fsaverage5	▶	32074_lh2fsaverage5_20.mgh
myStudy.csv	▶	32074_rh2fsaverage5_20.mgh
surfstat	▶	32076_lh2fsaverage5_20.mgh
surfstat_example.m	▶	32076_rh2fsaverage5_20.mgh
surfstat_example.m~	▶	32077_lh2fsaverage5_20.mgh
thickness	▶	32077_rh2fsaverage5_20.mgh
		32081_lh2fsaverage5_20.mgh
		32081_rh2fsaverage5_20.mgh
		32082_lh2fsaverage5_20.mgh
		32082_rh2fsaverage5_20.mgh
		32083_lh2fsaverage5_20.mgh
		32083_rh2fsaverage5_20.mgh
		32085_lh2fsaverage5_20.mgh
		32085_rh2fsaverage5_20.mgh
		32090_lh2fsaverage5_20.mgh
		32090_rh2fsaverage5_20.mgh
		32091_lh2fsaverage5_20.mgh
		32091_rh2fsaverage5_20.mgh
		32093_lh2fsaverage5_20.mgh
		32093_rh2fsaverage5_20.mgh
		32094_lh2fsaverage5_20.mgh
		32094_rh2fsaverage5_20.mgh
		32099_lh2fsaverage5_20.mgh
		32099_rh2fsaverage5_20.mgh
		32102_lh2fsaverage5_20.mgh
		32102_rh2fsaverage5_20.mgh
		32103_lh2fsaverage5_20.mgh

# LETS VERIFY WHAT WE JUST LOADED

```
%% lets verify
f=figure,
imagesc(T,[1.5 4])
colormap(parula)

f=figure,
SurfStatViewData(T(1,:),SM, 'first case')
SurfStatColLim([1.5 4])
colormap(parula)

f=figure,
SurfStatViewData(mean(T,1),SM, 'mean thickness across all subjects')
SurfStatColLim([1.5 4])
colormap(parula)
```



# NOW WE CAN FINALLY BUILD THE FIRST LINEAR MODELS

t-tests, correlations, partial correlations, ANOVAs, MANOVAs,...

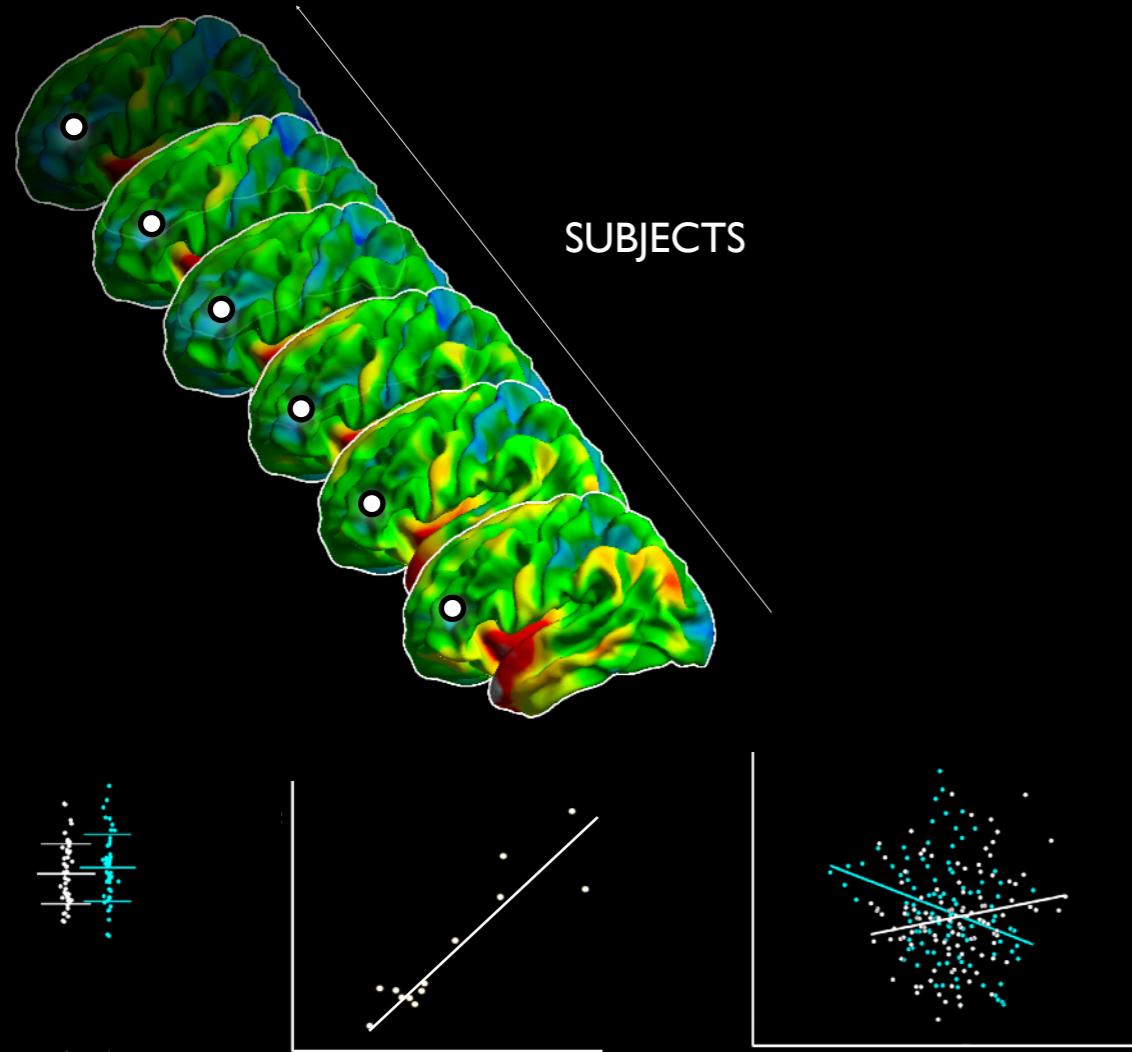
are just specific instances of the linear model of the form

$$Y \sim \beta_0 + \beta_1 * x_1 + \beta_2 * x_2 + \beta_3 * x_1 * x_2 \dots + \varepsilon$$

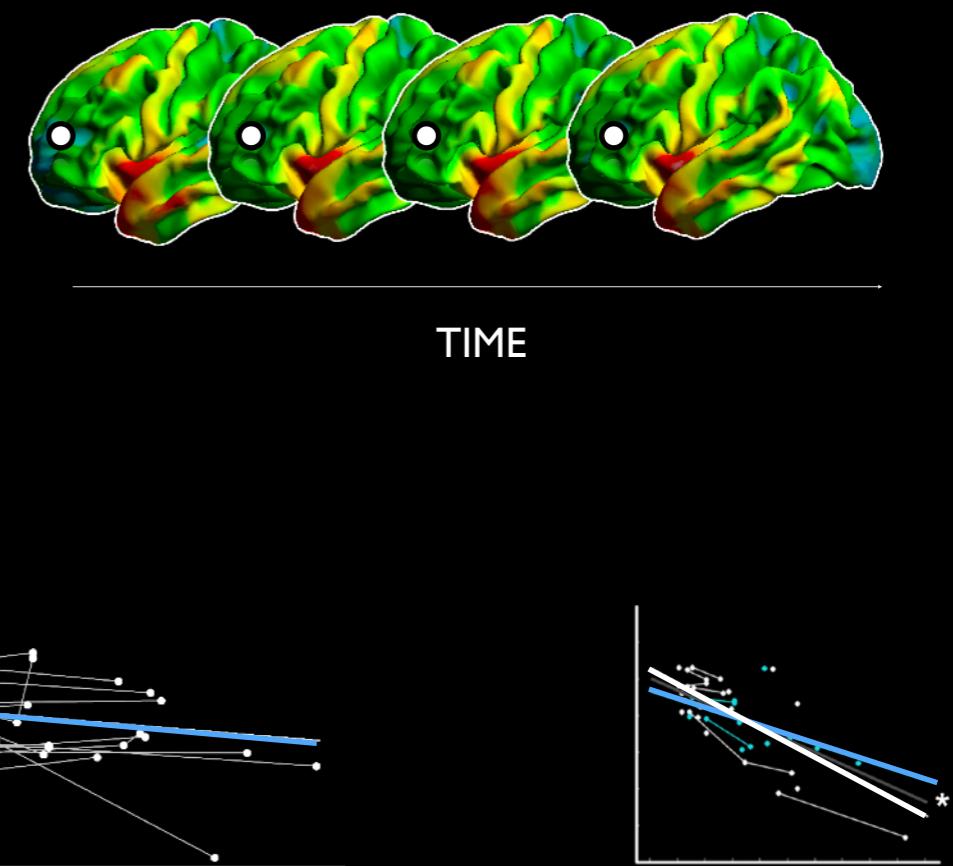
DATA                    INTERCEPT                    SIMPLE EFFECTS                    INTERACTIONS

# NOW WE CAN FINALLY BUILD THE FIRST LINEAR MODELS

CROSS-SECTIONAL ANALYSES

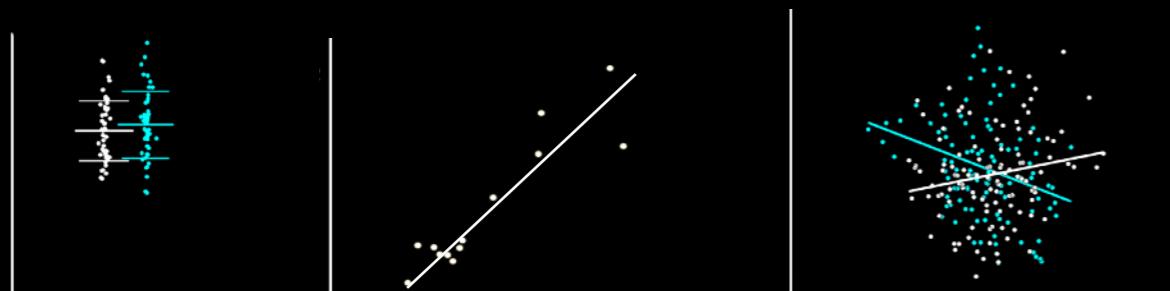
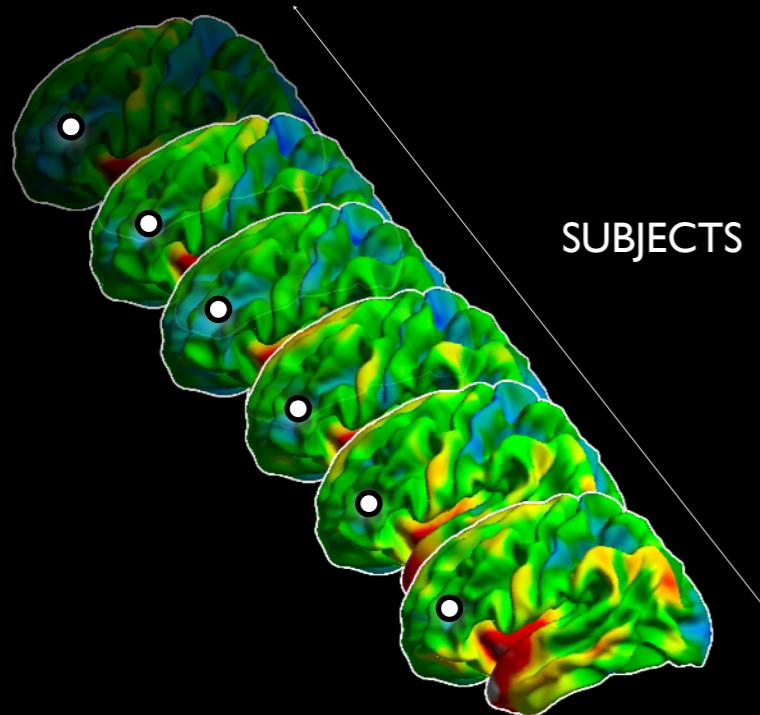


LONGITUDINAL ASSESSMENTS



# NOW WE CAN FINALLY BUILD THE FIRST LINEAR MODELS

CROSS-SECTIONAL ANALYSES

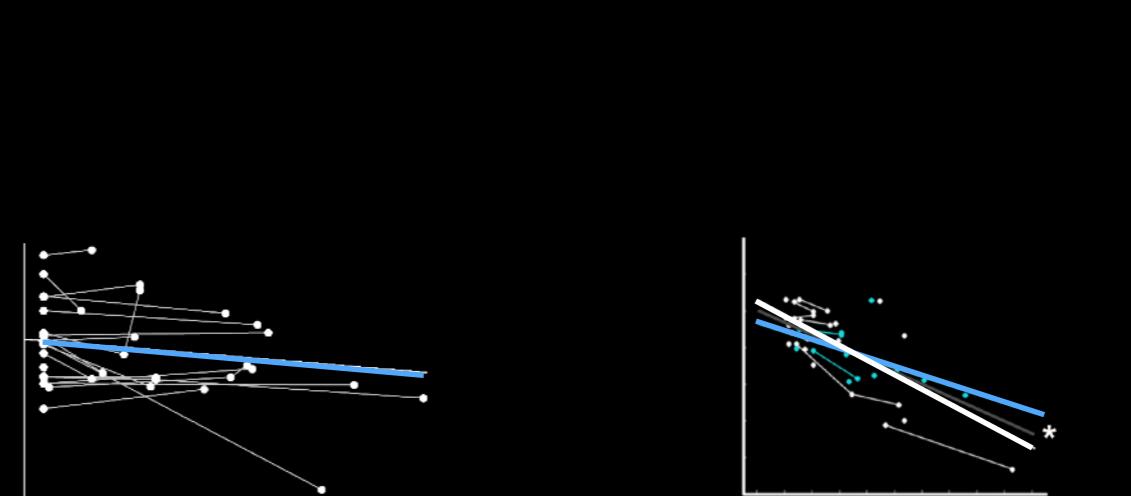
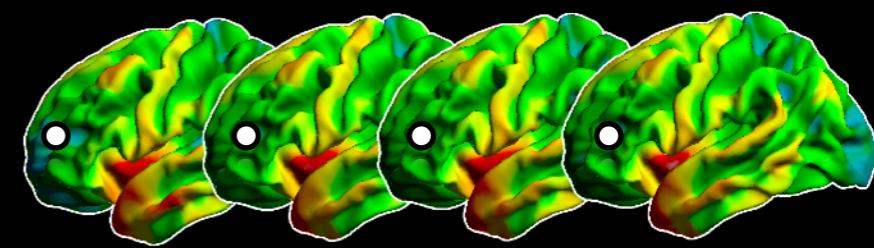


$$Y = I + G$$

$$Y = I + A$$

$$Y = I + G + A + G \times A$$

LONGITUDINAL ASSESSMENTS



$$Y = I + G$$

$$Y = I + r(S) + ISI + G + ISI \times G$$

# NOW WE CAN FINALLY BUILD THE FIRST LINEAR MODELS

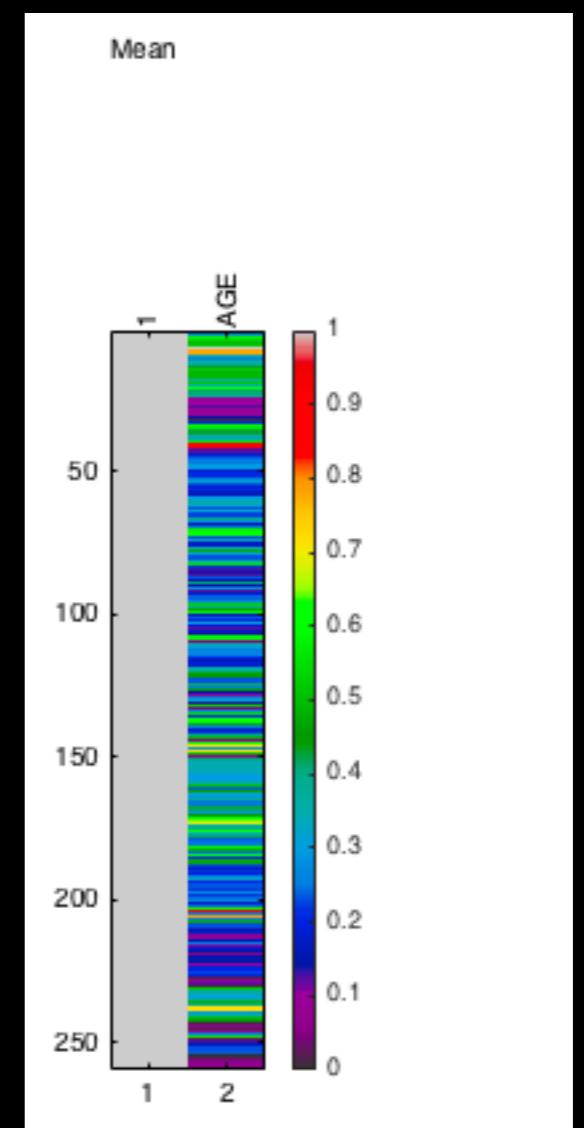
```
%% now we can finally do a firs models: lets look at effects of age

% first code some variables of interest
AGE_term = term(AGE);

% then build a model
M = 1 + AGE_term;
f=figure, image(M)

% estimaste the model parameter s
slm = SurfStatLinMod(T, M, SW);

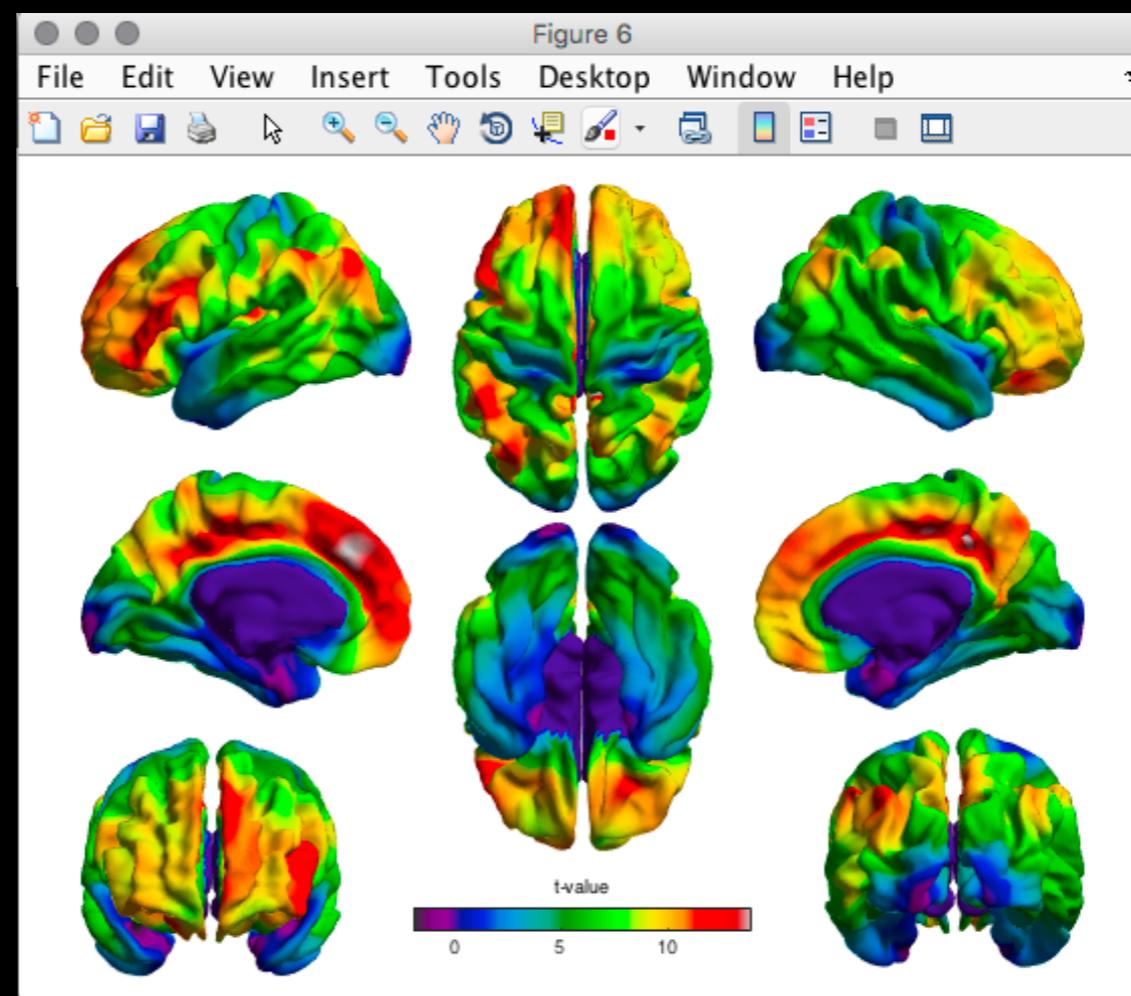
% specifiy contrast
slm = SurfStatT(slm, -AGE)
```



# NOW WE CAN FINALLY BUILD THE FIRST LINEAR MODELS

```
slm =  
  
    X: [259x2 double]  
    df: 257  
    coef: [2x20484 double]  
    SSE: [1x20484 double]  
    tri: [40960x3 int32]  
    resl: [61440x1 double]  
    c: [-2.9671e-14 -1.0000]  
    k: 1  
    ef: [1x20484 double]  
    sd: [1x20484 double]  
    t: [1x20484 double]
```

```
% display t-value  
f=figure  
SurfStatViewData(slm.t, SM, 't-value')
```



## AND CORRECT FOR MULTIPLE COMPARISONS

```
% multiple comparison correction: none
p = 1-tcdf(slm.t,slm.df);
f=figure
    SurfStatViewData(p, SM, 'p-value')
    SurfStatColLim([0 0.05])
    colormap([parula; .8 .8 .8])

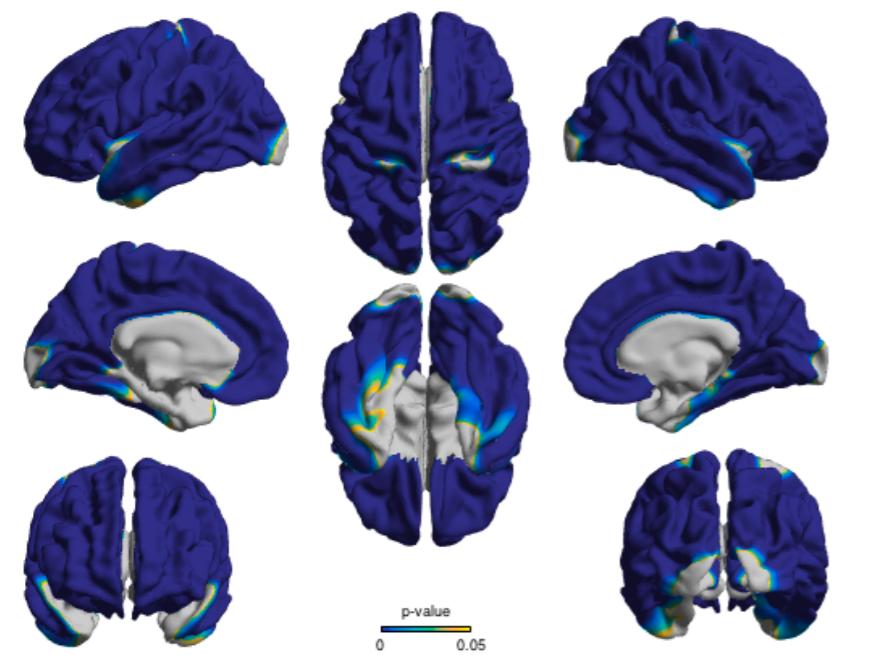
% multiple comparison correction: Bonferroni
p = 1-tcdf(slm.t,slm.df);
p = p*size(p,2);
f=figure
    SurfStatViewData(p, SM, 'Bonferroni p-value')
    SurfStatColLim([0 0.05])
    colormap([parula; .8 .8 .8])

% multiple comparions using fdr
qval = SurfStatQ(slm,mask);
f=figure
    SurfStatView(qval, SM, 'fdr')

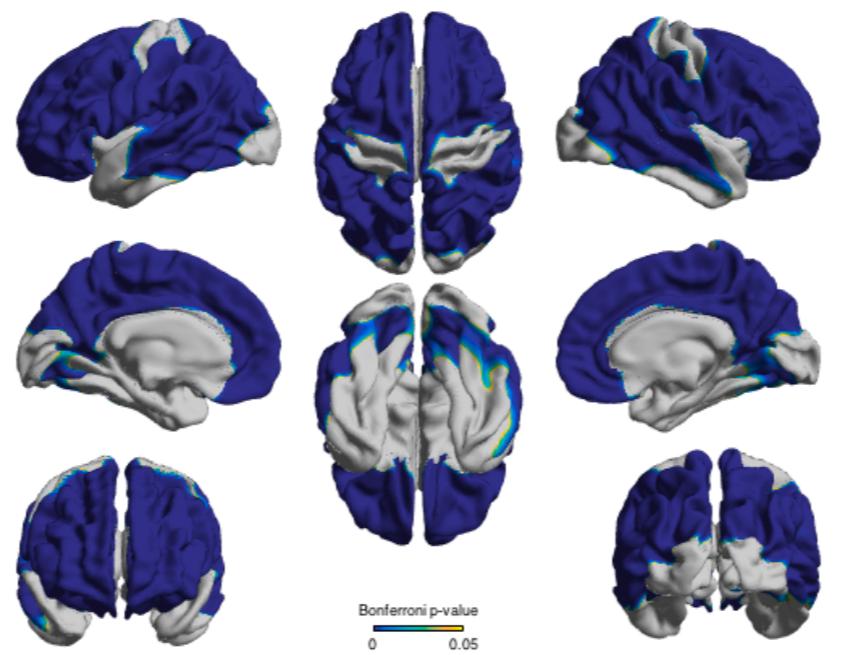
% multiple comparions using random field theory
pval = SurfStatP(slm,mask);
f=figure
    SurfStatView(pval, SM, 'rft')
```

# AND CORRECT FOR MULTIPLE COMPARISONS

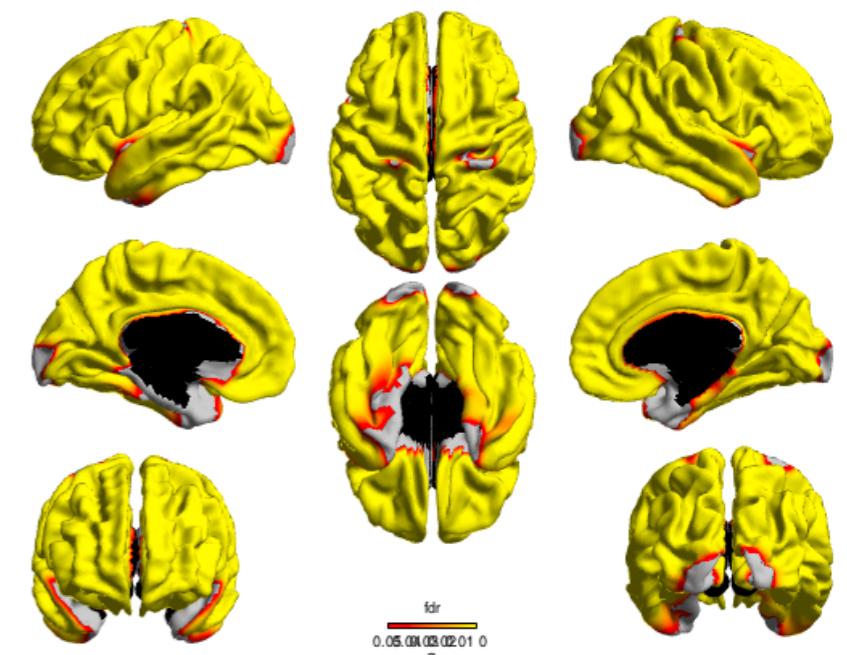
uncorrected



Bonferroni



FDR



## SIMPLE LINEAR MODELS CONTINUED

Different models and contrasts are possible

```
A = term(AGE); G = term(GR);
```

```
M = 1 + G
```

```
slm = SurfStatLinMod(T, M, SW)
```

```
slm = SurfStatT(slm, G.Group1-G.Group2)
```

is a model that assesses group differences

$$y = \beta_0 + \beta_1 * G + \varepsilon$$

## SIMPLE LINEAR MODELS CONTINUED

Different models and contrasts are possible

```
A = term(AGE); G = term(GR);
```

```
M = 1 + A + G
```

```
slm = SurfStatLinMod(T, M, SW)
```

```
slm = SurfStatT(slm, G.Group1-G.Group2)
```

is a model that assesses group differences, controlling for age

$$y = \beta_0 + \beta_1 * A + \beta_2 * G + \varepsilon$$

## SIMPLE LINEAR MODELS CONTINUED

Different models and contrasts are possible

```
A = term(AGE); G = term(GR);
```

```
M = 1 + A + G
```

```
slm = SurfStatLinMod(T, M, SW)
```

```
slm = SurfStatT(slm, -AGE)
```

is the same model but assesses age effects, controlling for group

$$y = \beta_0 + \beta_1 * A + \beta_2 * G + \varepsilon$$

## SIMPLE LINEAR MODELS CONTINUED

Different models and contrasts are possible

```
A = term(AGE); G = term(GR);
```

```
M = 1 + A + G + A*G
```

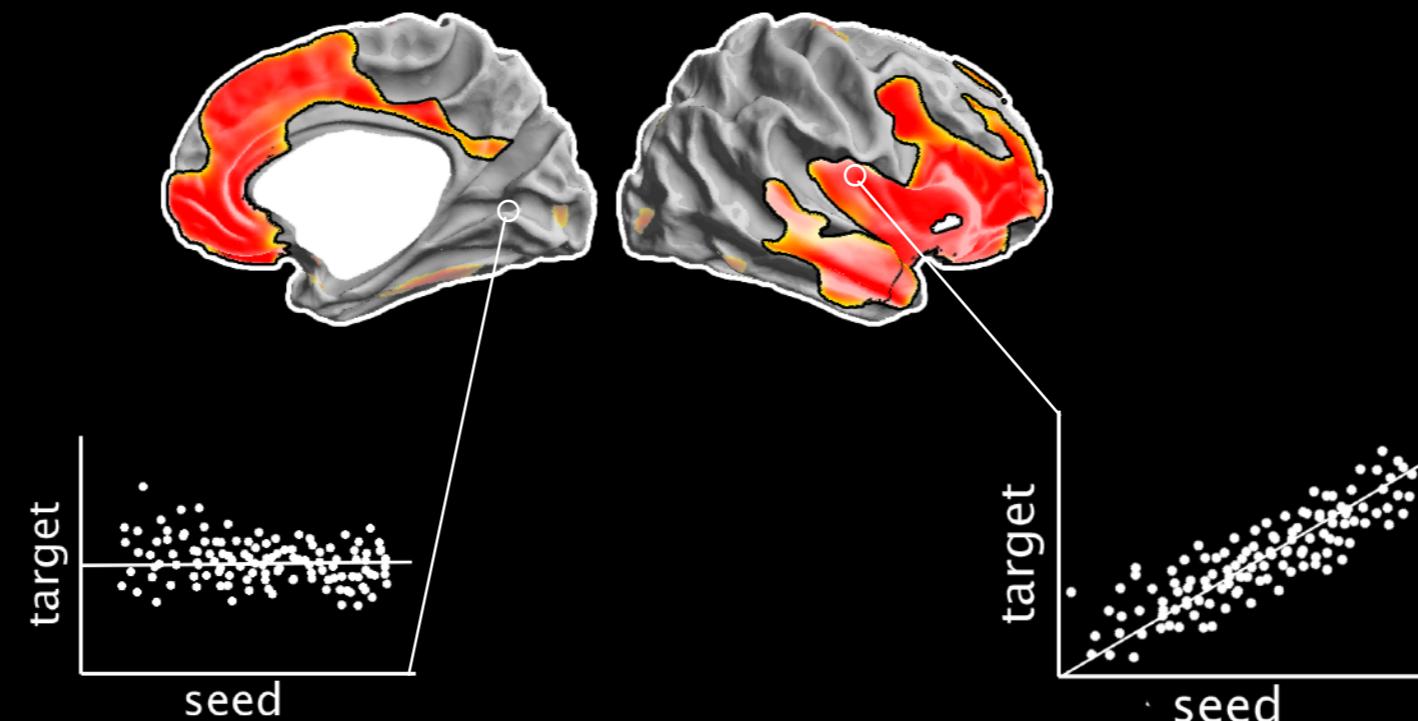
```
slm = SurfStatLinMod(T, M, SW)
```

```
slm = SurfStatT(slm, (-AGE.*G.Group2) - (-AGE.*G.Group1))
```

interaction model, assumes different age effect across groups

$$y = \beta_0 + \beta_1 * A + \beta_2 * G + \beta_3 * G * A + \varepsilon$$

# STRUCTURAL COVARIANCE ANALYSIS



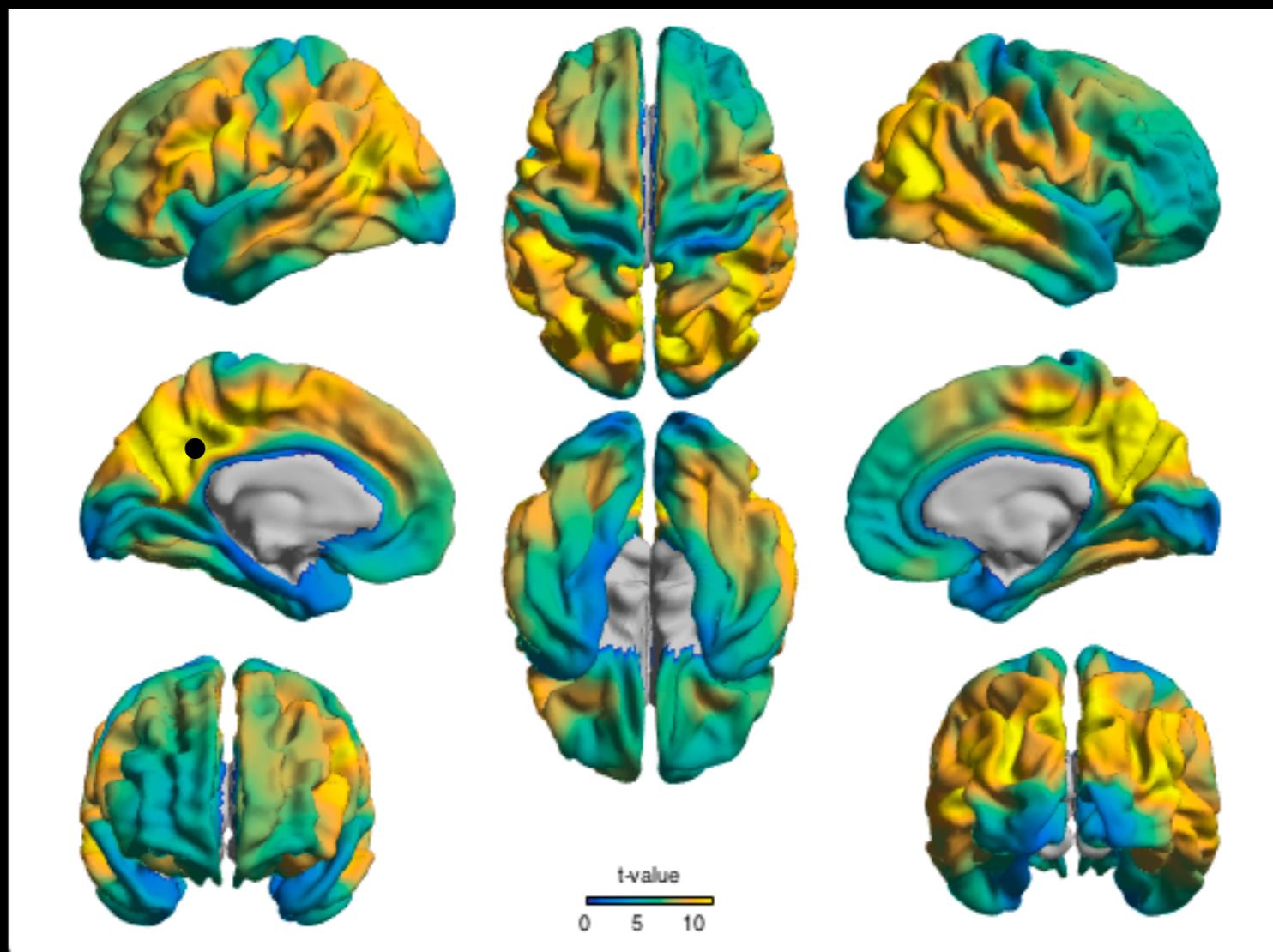
Lerch et al. (2006) NeuroImage  
Alexander-Bloch et al. (2013) Nat Rev Neurosci

## COVARIANCE MODELS

```
Seed = T(:,14446);  
  
S = term(Seed);  
  
M = 1 + S  
  
slm = SurfStatLinMod(T,M, S)  
  
slm = SurfStatT(slm, Seed )
```

is a model that assesses the correlation between a seed and cortical thickness at each surface point

## COVARIANCE MAP



## COVARIANCE DIFFERENCES

```
Seed = T(:,14446);  
  
S = term(Seed); G = term(Group);  
  
Model = 1 + S + G + G*S  
  
slm = SurfStatLinMod(T,Model, S)  
  
slm = SurfStatT(slm,(G.Group1.*Seed)-(Group2.*Seed))
```

is a model that assesses the interaction between seed and group, assessing a stronger correlation with seed thickness in controls than patients

# SUMMARY

SurfStat is a swiss army knife to flexibly analyze MRI data

- ▶ reading and writing data
- ▶ perform surface-based / volume-based statistical analysis
- ▶ correct for multiple comparison
- ▶ display results
- ▶ Not limited to thickness data: resting-state fMRI, intensity, diffusion

## OTHER COOL STUFF

Mixed effects models (for longitudinal designs)

```
M = 1 + A + S + random(SUBJID)
```

Non-surface based analysis in e.g. thickness in ROI

```
SurfStatLinMod(roi, Model)
```

Analysing volume data (e.g., VBM, DBM, rs-fMRI)

```
SurfStatReadVol1 ...
```

Smoothing on surfaces, mapping between volume and surface space

```
SurfStatSmooth ... SurfStatVol2Surf ...
```

- ▶ visit: <http://www.math.mcgill.ca/keith/surfstat/>

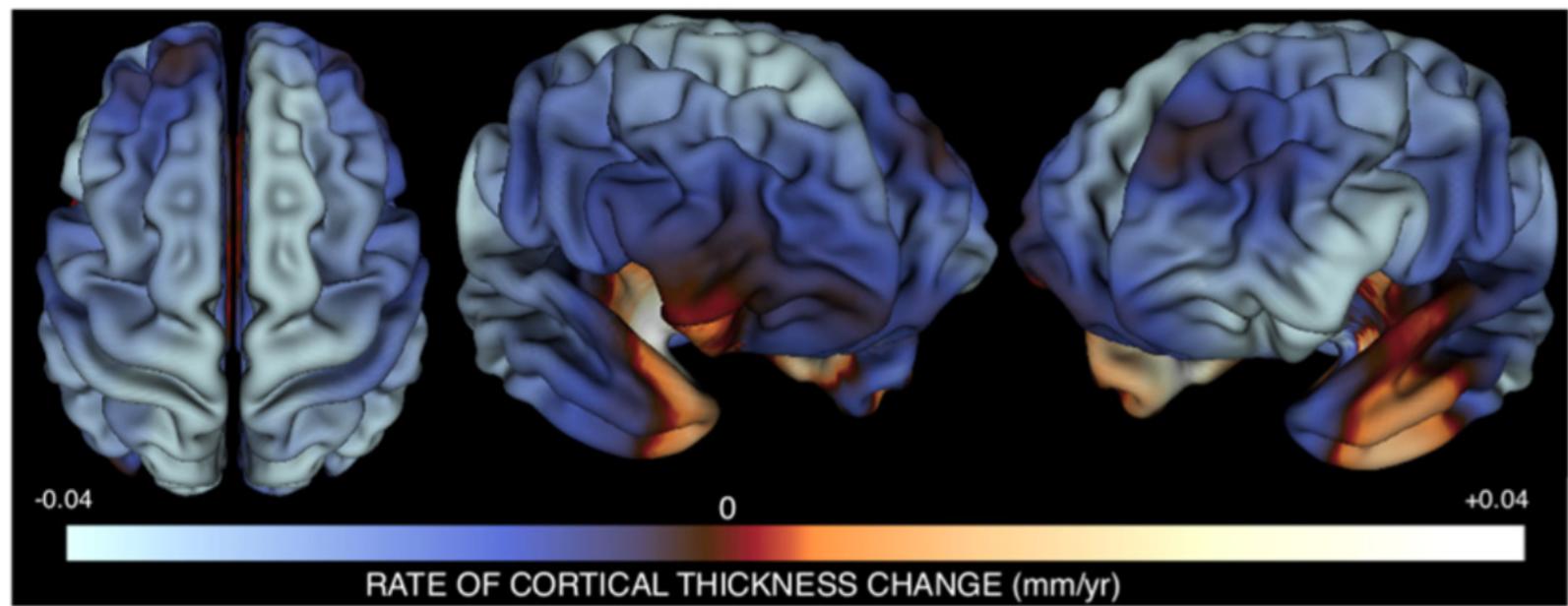


<https://github.com/MICA-MNI/micaopen>

RAZNAHAN

**Table 1. Participant Characteristics**

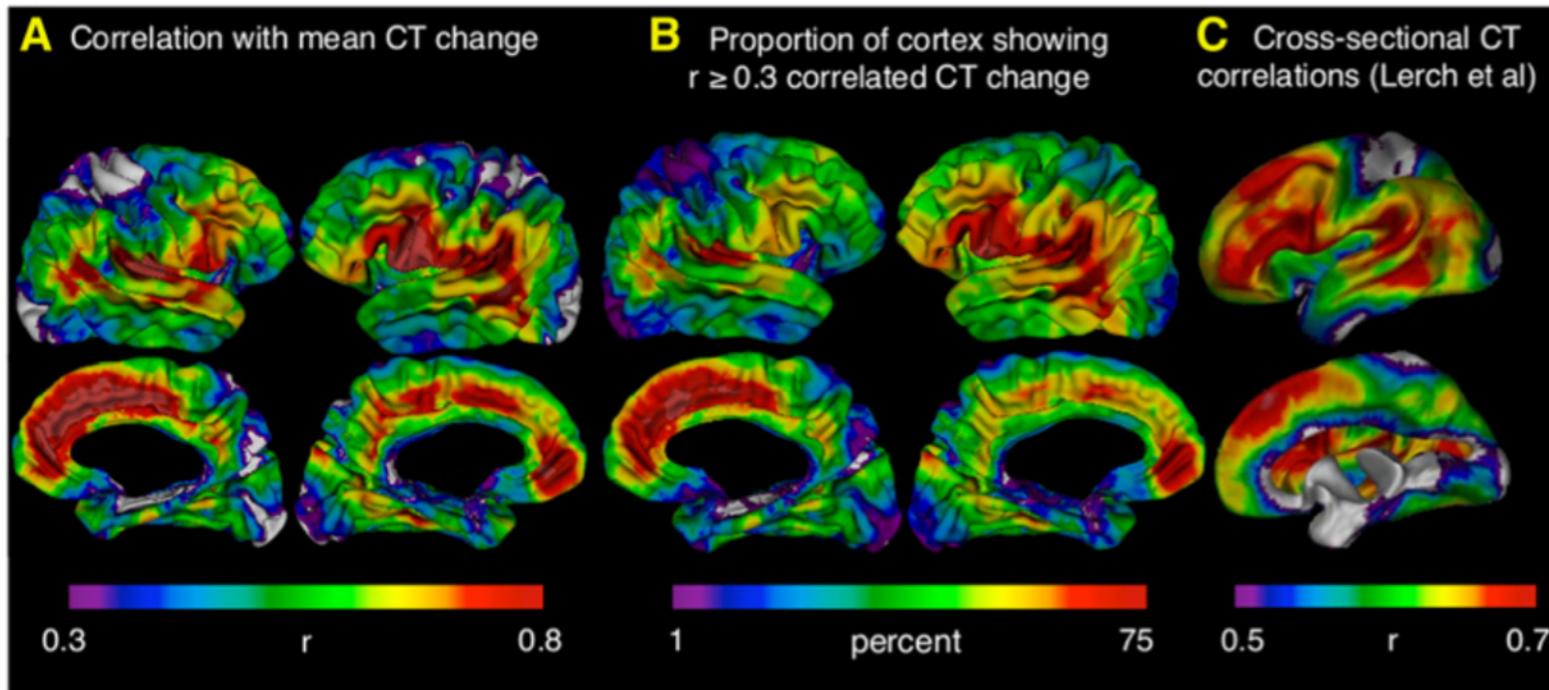
Characteristic	Group			Sex Difference
	All	Male	Female	
Number of Individuals	108	67	41	n.s.
Singleton	63	35	28	
Member of twin pair	45	32	13	
Handedness, No.				n.s.
Right	99	62	37	
Mixed	4	3	1	
Left	5	2	3	
Race, No.				n.s.
Caucasian	98	62	36	
African-American	3	1	2	
Asian	2	1	1	
Hispanic	3	2	1	
Other	2	1	1	
IQ				n.s.
Mean (SD)	115 (11.8)	116 (11.5)	114 (12.4)	
SES				n.s.
Mean (SD)	40 (17.6)	39 (18.5)	41 (16.2)	
Number of scans, No.				n.s.
3 scans	67	39	28	
4 scans	31	22	9	
5 scans	9	5	4	
6 scans	1	1	0	
Total	376	236	140	
Age Distribution of Scans (years)				
Mean (SD)	15.2 (3.5)	15.3 (3.5)	14.9 (3.5)	
Range	9.1–22.8	9.2–22.7	9.1–22.8	
n.s., not statistically significant at p < 0.05; SES, socioeconomic status.				



**Figure 1. Mapping the Mean Rate of CT Change per Year between Ages 9 and 22 Years using Person-Specific Estimates of CT Change**

Three views of the cortical sheet are shown. Colors represent the magnitude of mean annual cortical thickness (CT) change within our sample at each vertex. Mean change values were derived by averaging estimates of weighted annual CT change across all participants. Over the age range studied, most cortical regions are becoming thinner with advancing age, with the exception of bilateral anterior-medial temporal and right orbitofrontal cortices where CT is still increasing with age. This approach to mapping annual CT change closely replicates results derived using traditional mixed-model

approaches for analyzing longitudinal data (Figure S1), and converges with other larger mixed-model studies of CT change (Shaw et al., 2008), but has the added advantage of permitting correlational analysis of interindividual differences in CT change at different vertices.



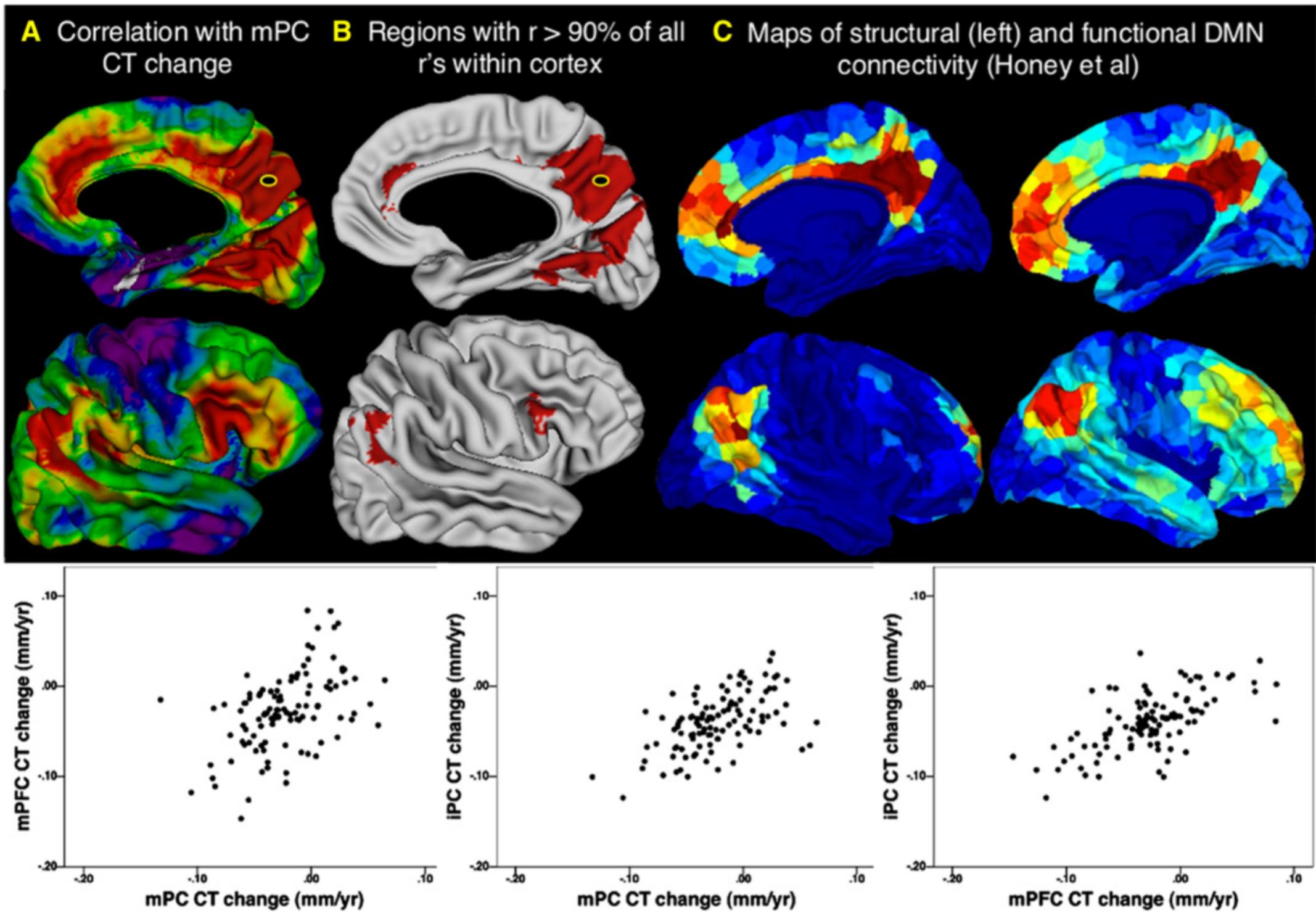
**Figure 2. Regional Differences in Correlation with Rates of CT Change throughout the Cortical Sheet**

(A) Map of correlation strength between CT change at each vertex and mean CT change across all vertices. This map has been arbitrarily thresholded at  $r \geq 0.3$  to highlight its similarity with a previously published thresholded map of cross-sectional CT correlations throughout the cortical sheet (Lerch et al., 2006). An unthresholded version of this map is provided in Figure S2A. Note that the strongest correlations with mean CT change are seen in fronto-temporal association cortices, whereas weakest correlations with mean CT change are seen in primary sensory cortices.

(B) An alternative representation of regional differences in maturational coupling. The color at a given cortical region represents the proportion of the cortical surface showing

correlated CT change with the region in question at  $r \geq 0.3$ . “Warmer” colors refer to higher proportions. Fronto-temporal regions show the most spatially extensive maturational coupling whereas primary sensory cortices show the least.

(C) A reproduction of earlier published (Lerch et al., 2006) maps showing the correlation between cross-sectional variation in CT at each vertex and mean CT across the whole vertex. Note the convergence between these maps and those for correlated CT change shown in (A) and (B).



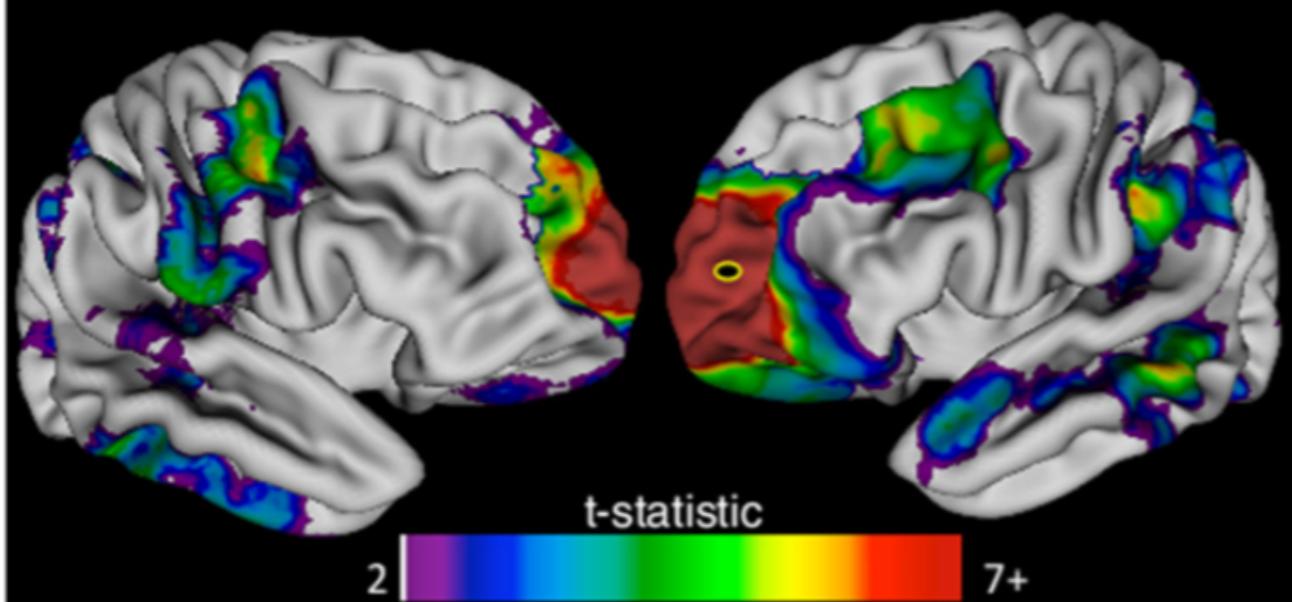
**Figure 3. Maturational Coupling within the Default Mode Network**

(A) Right hemisphere map of maturational coupling with the medial posterior cortex (mPFC) default mode network (DMN) node. Color gradations represent correlation centile position in the distribution of all possible correlations between cortical vertices (blue → red: 1st → 100th centile).

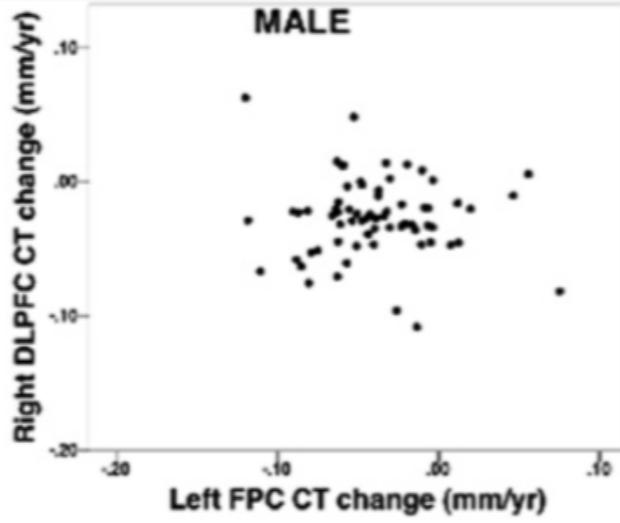
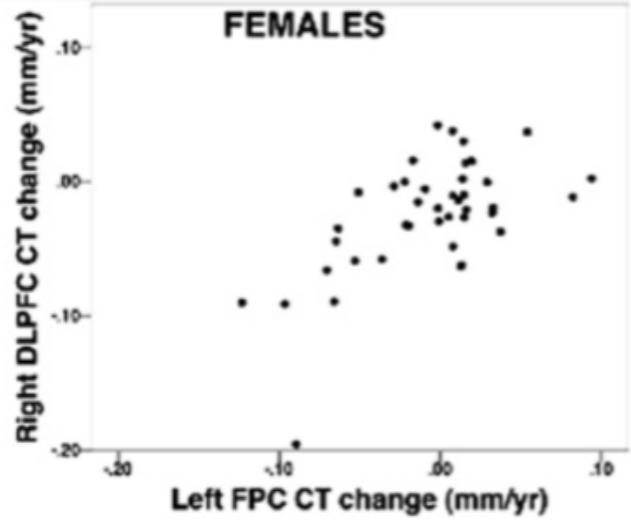
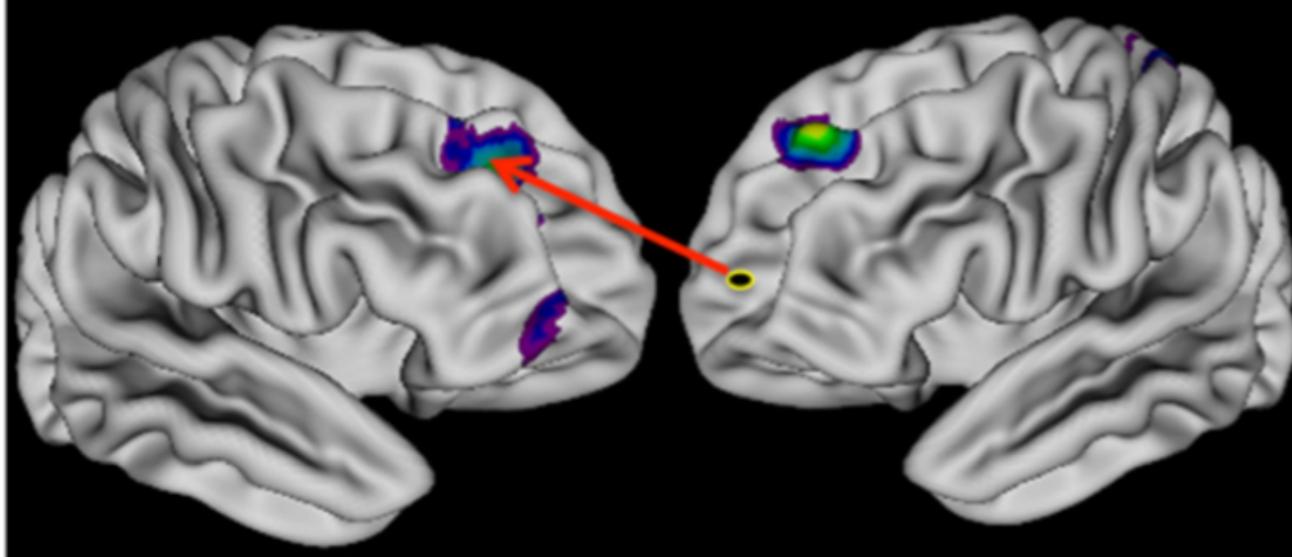
(B) Regions where correlations with mPFC change are in the top 90% of all possible correlations. Note mPFG and iPC overlaps between the distribution of regions showing highly coordinated maturation with the mPFC DMN seed, and the distribution of regions that show high functional and structural connectivity within the DMN.

(C and D) Figures from Honey et al. (2009) depicting the DMN by analysis of diffusion tensor imaging and functional magnetic resonance data, respectively.

**A** Regions showing significant coupling with left frontopolar maturation in both sexes



**B** Regions showing significant sexually-dimorphic coupling with left frontopolar maturation



**Figure 4. Maturational Coupling with the Left Frontopolar Cortex and Its Variation by Sex**

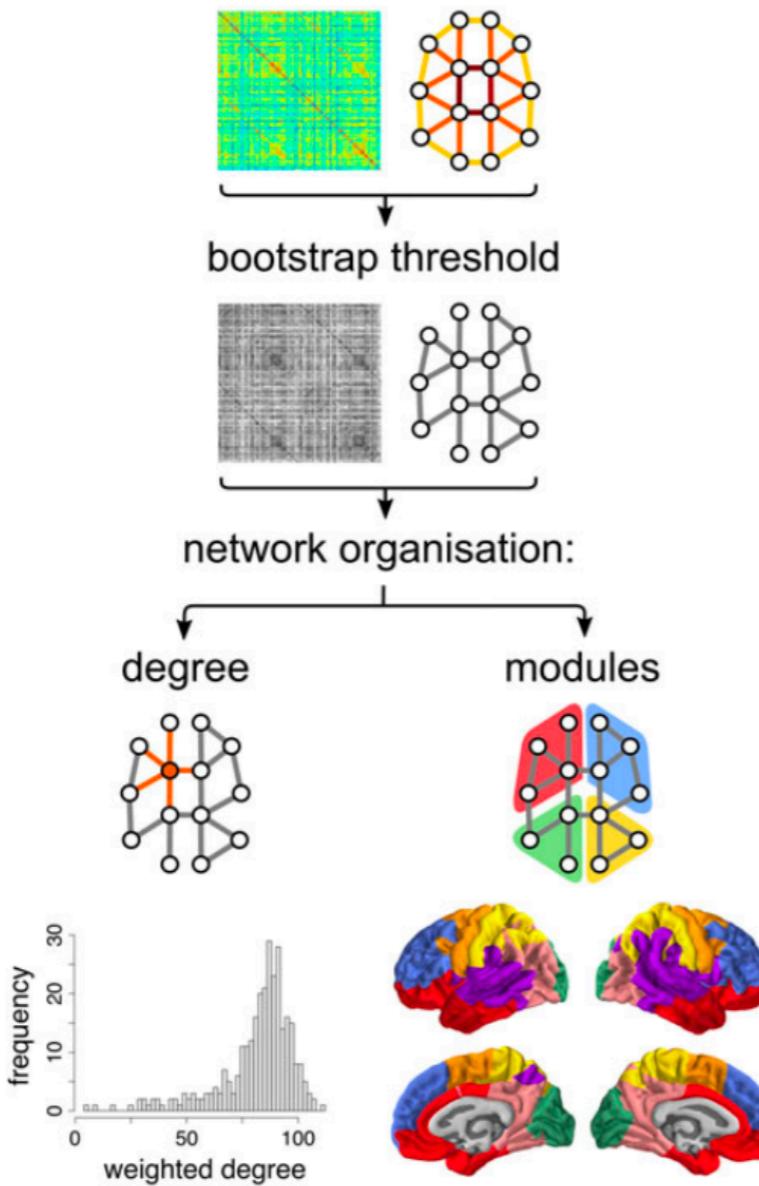
The left frontopolar cortex (FPC) was used as a seed to explore sex differences in maturational coupling because it is where rate of cortical thickness (CT) change shows statistically significant sex differences over the age range studied—in both prior work (Raznahan et al., 2010; Figure S4A) and our current study (Figure S4B).

(A) Map of regions showing significant maturational coupling with left FPC that is not significantly different in magnitude between males and females. Note the very strong relationship between left FPC change and change at its contralateral homolog. Several regions show bilateral coupling with IFPC change (e.g., inferior temporal, planum temporale, angular gyrus and orbitofrontal cortex).

(B) Regions where coupling with IFPC CT change differs significantly between males and females. These consist of areas where coupling is specific to females, as shown for the right dorsolateral prefrontal cortex (rDLPFC) in the inset scatter plot. Furthermore, sex differences in FPC-DLPFC coupling also remained statistically significant after removal of nine outliers (defined using a conservative Cooks distance threshold of  $4/n$ ).

VASA

### A age-invariant structural network



### B age-resolved structural network development

