

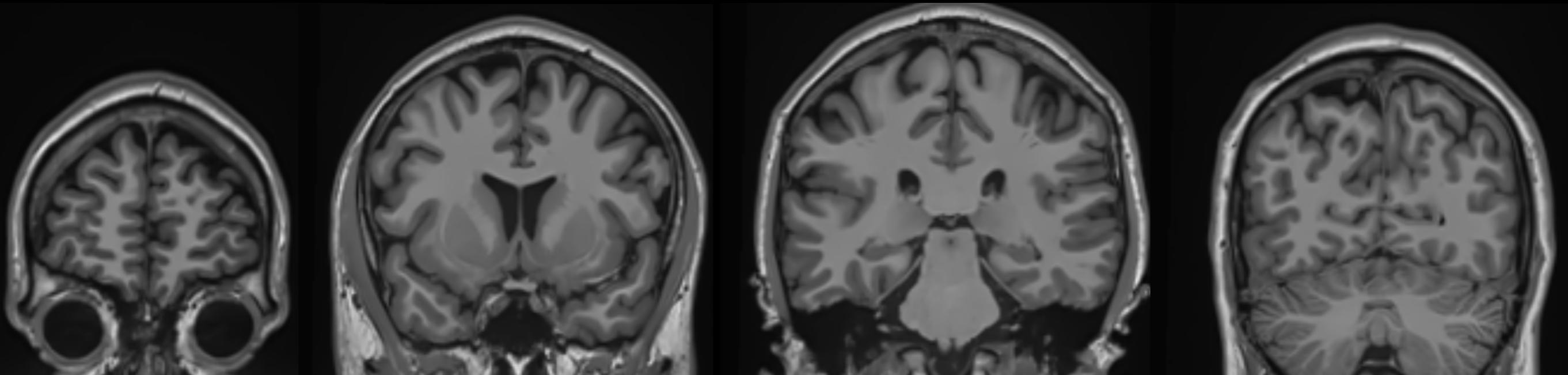


THE MISSED CLASS

NEUR-608

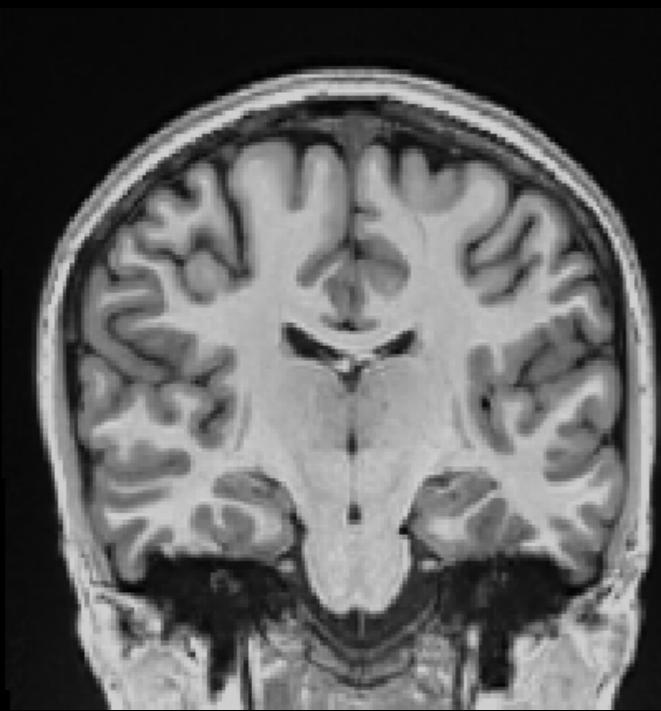


# structural MRI

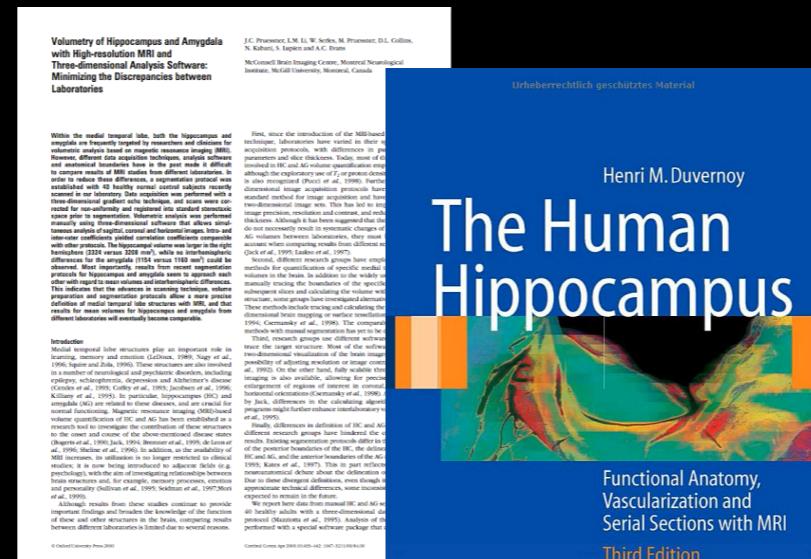


T1-weighted MRI

# VOLUMETRY PIPELINE



input

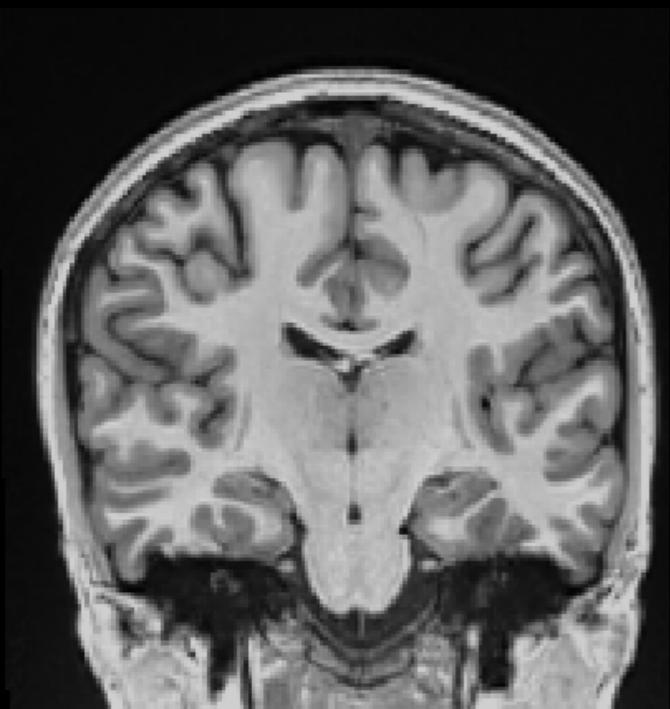


read and become expert



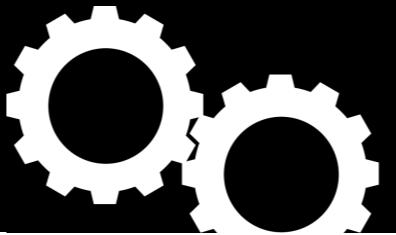
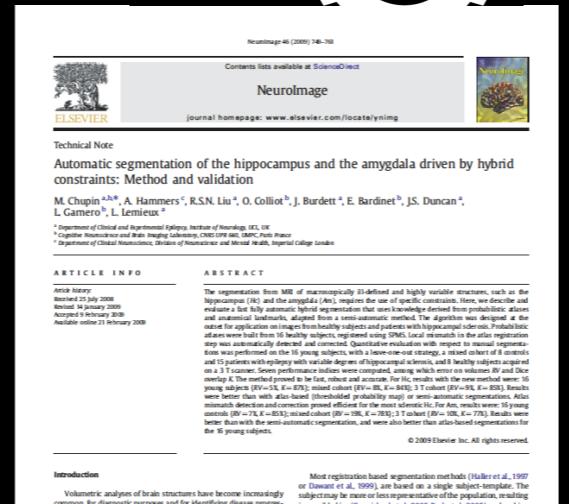
segment

# AUTOMATATION



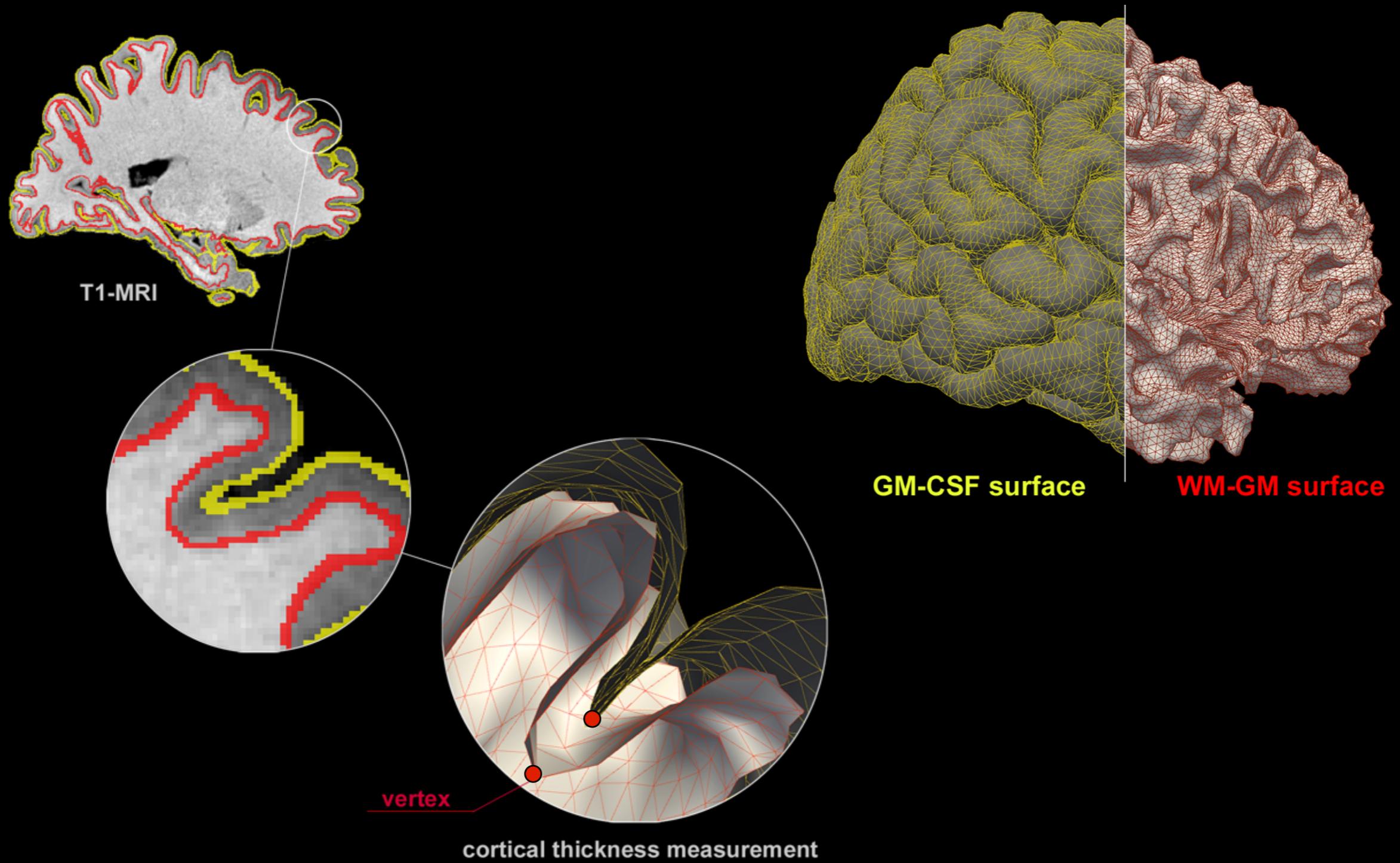
input

automatic segmentation approaches

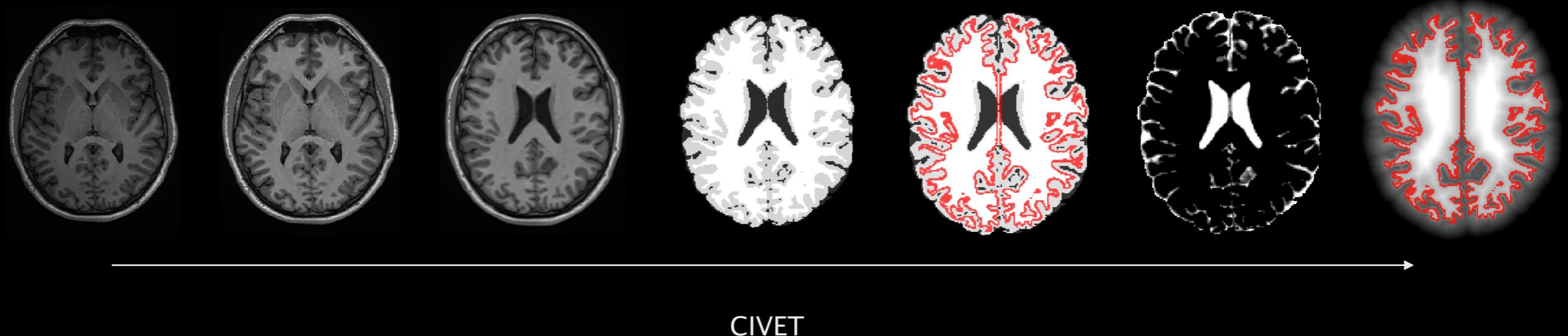


segment

# CORTICAL THICKNESS MEASUREMENT



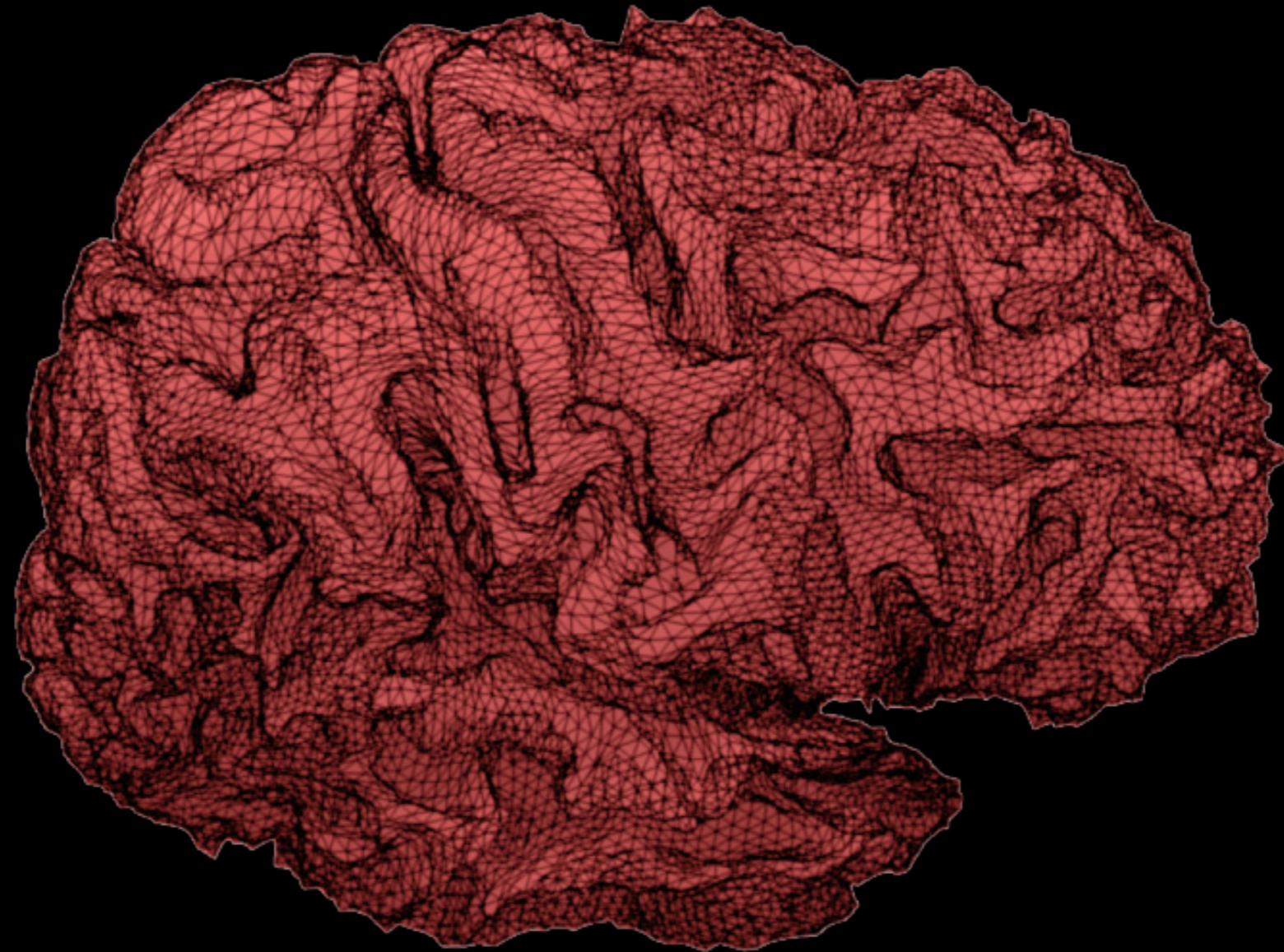
# PROCESSING PIPELINES



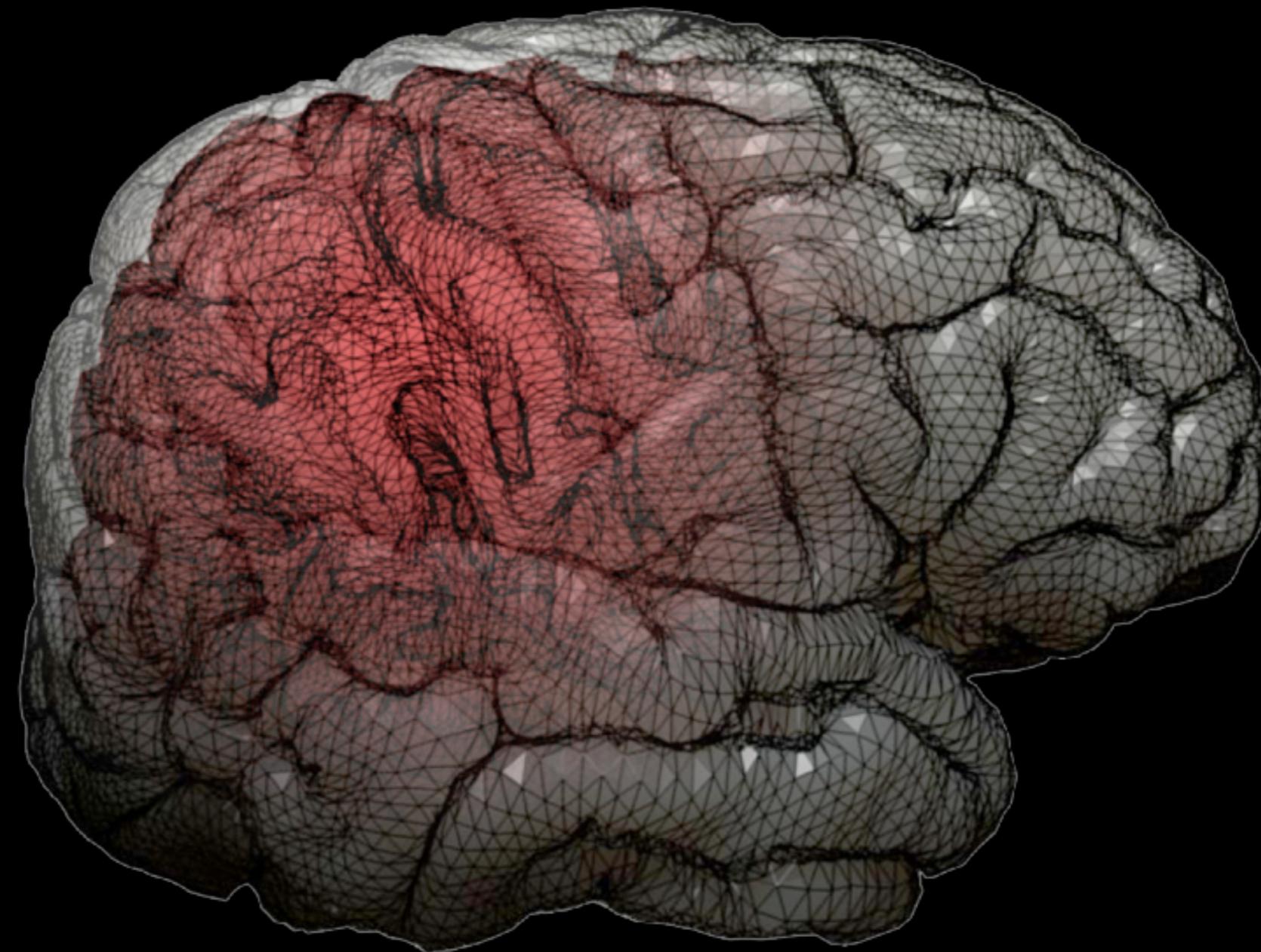
MacDonald et al. (2000) *NeuroImage*

Kim et al. (2005) *NeuroImage*

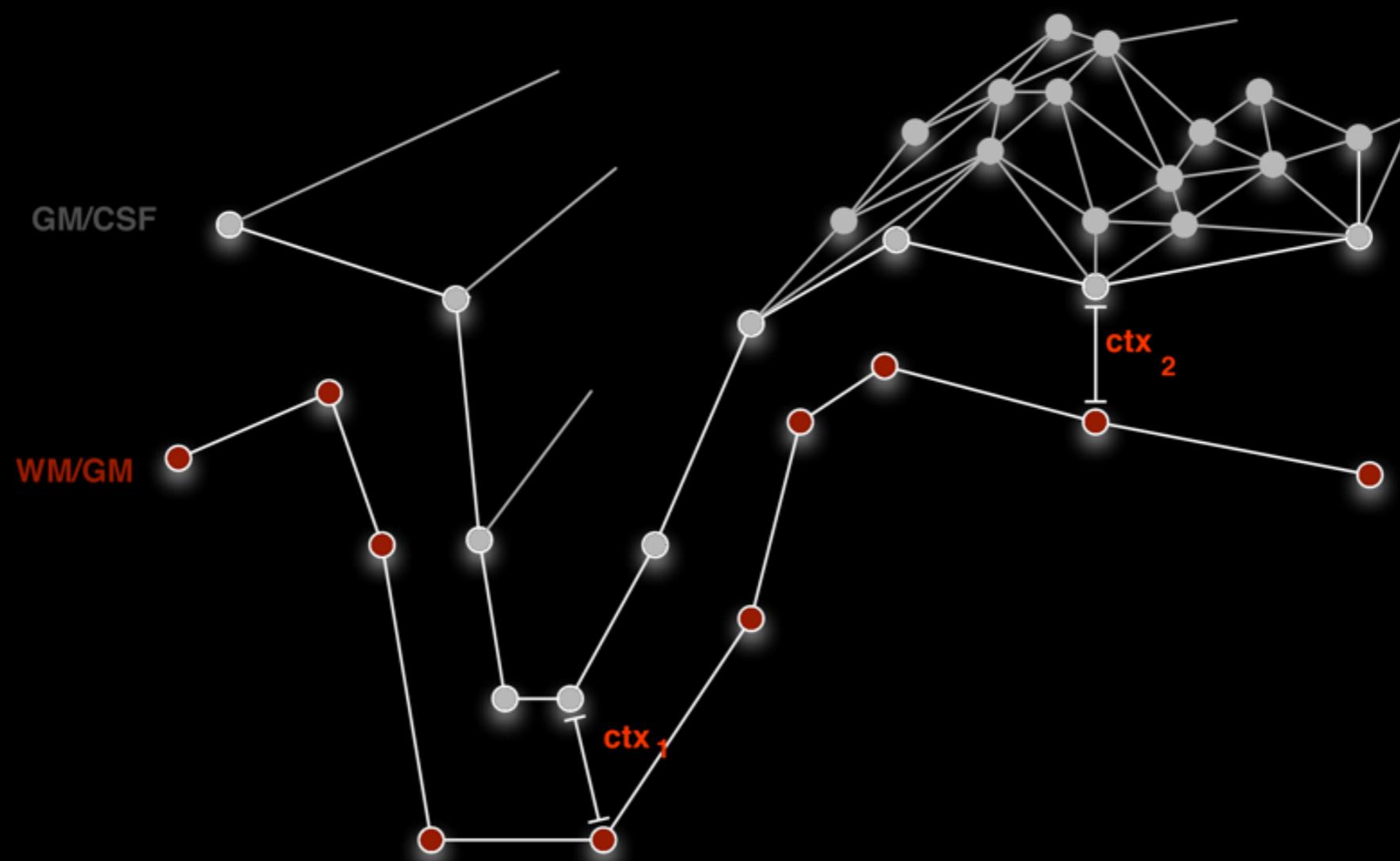
WM



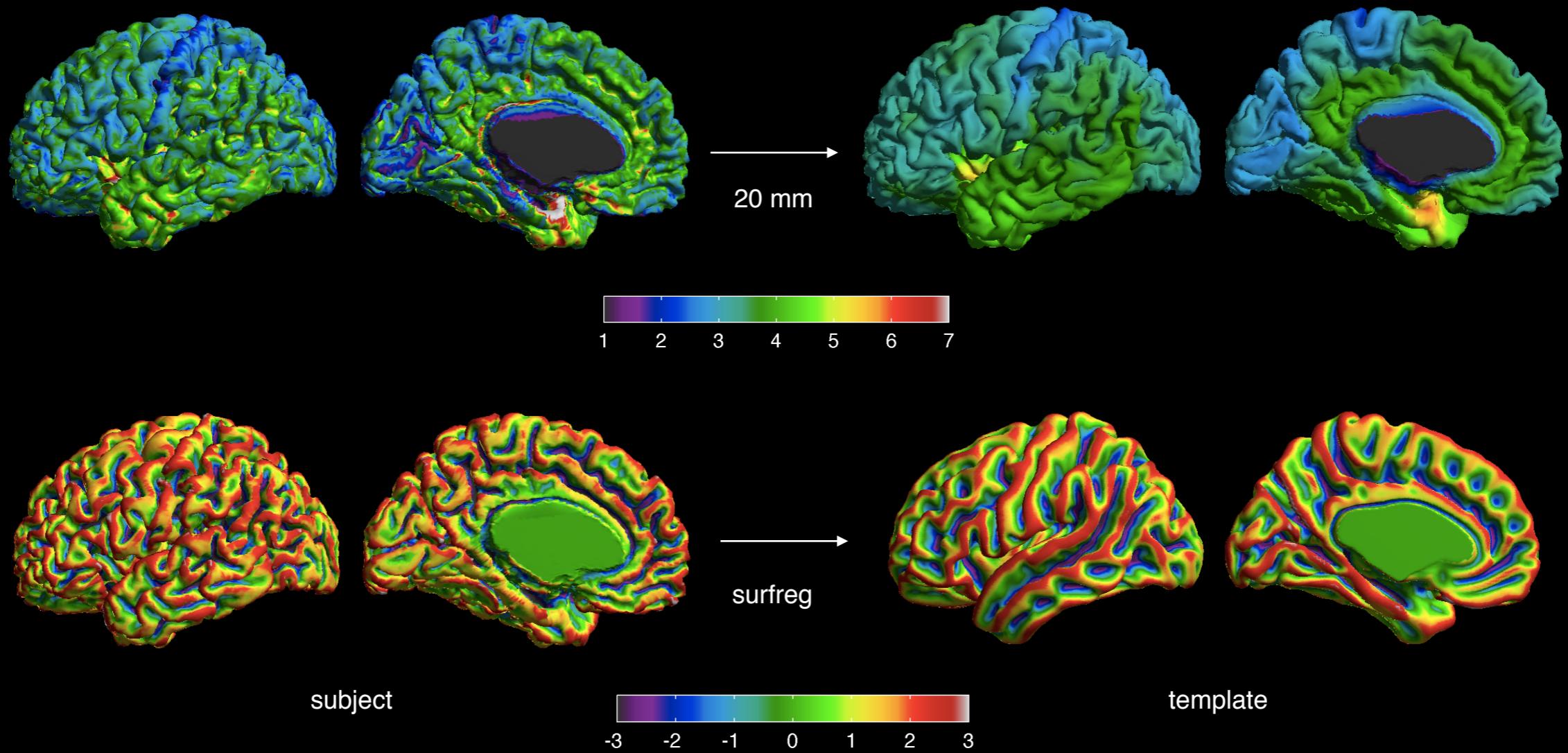
GM SURFACE



# 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,...

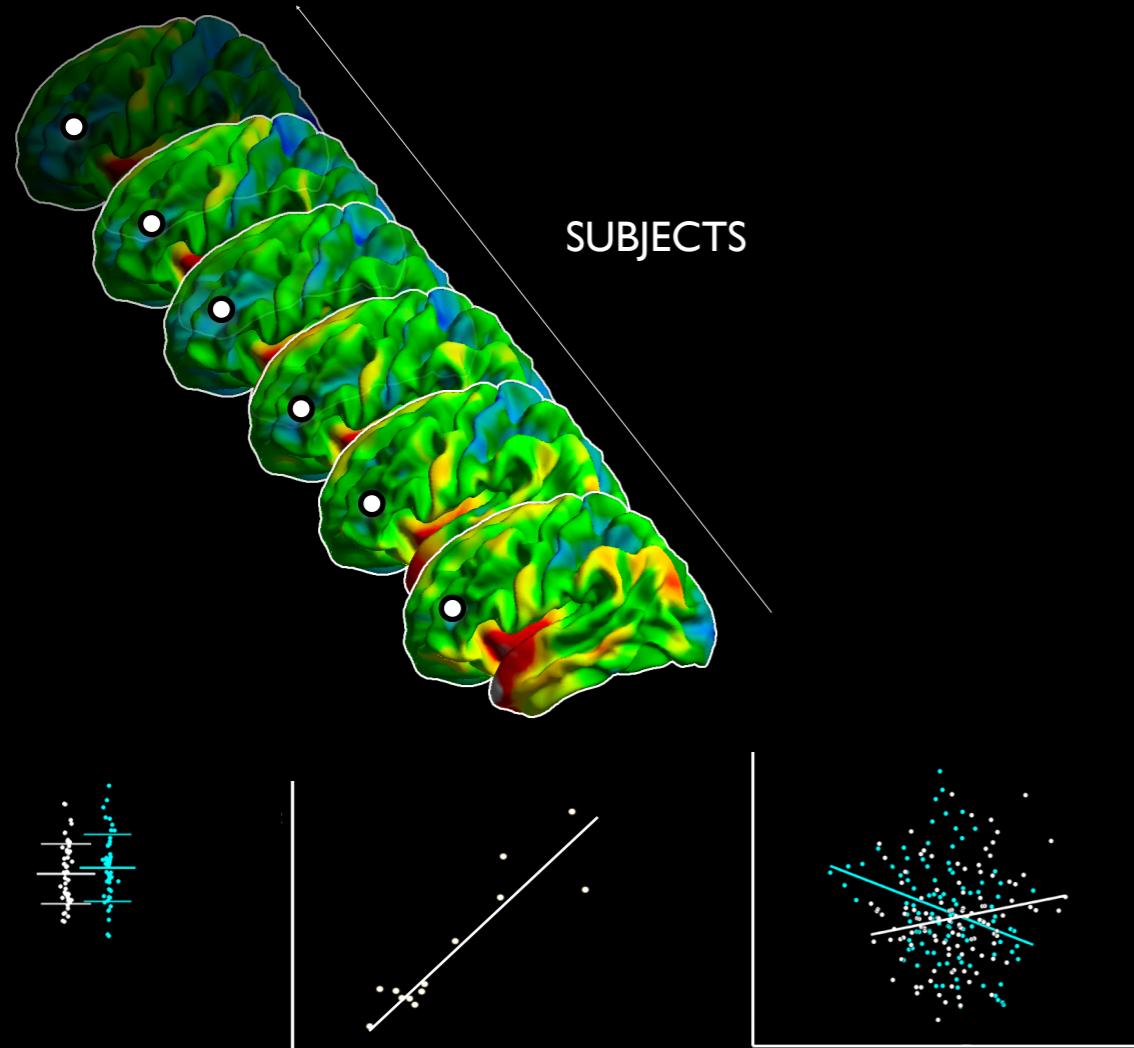
are just specific instances of the linear model of the form

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

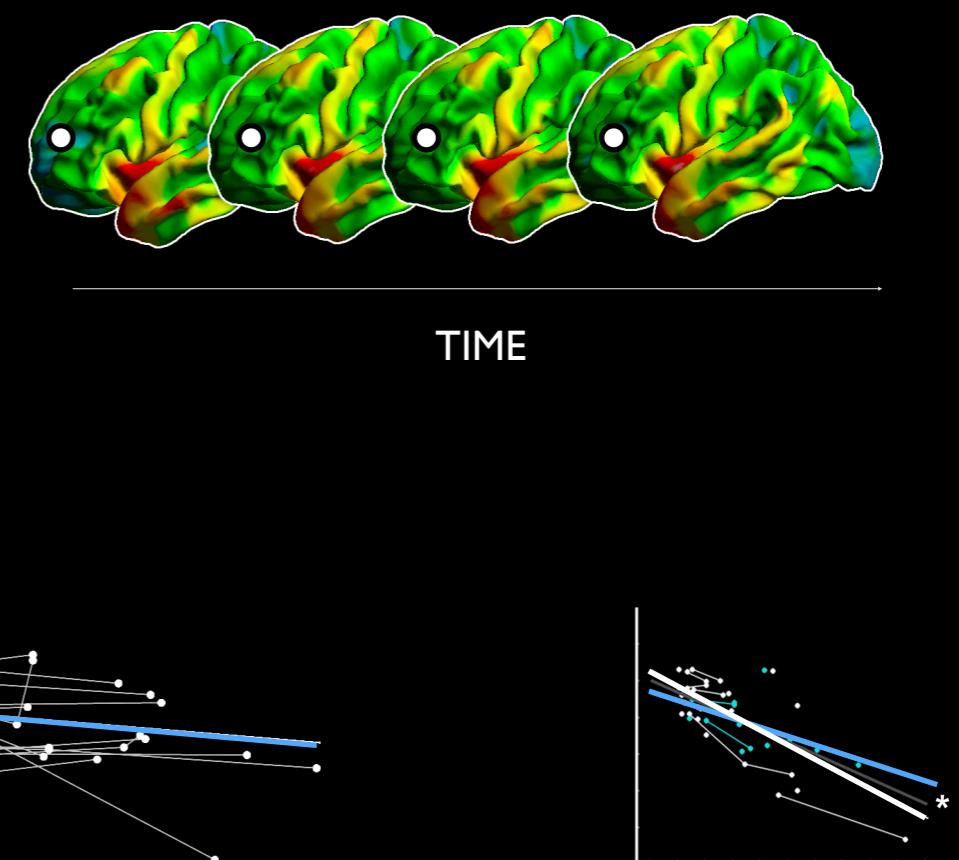
The diagram illustrates the components of a linear model equation. At the top is the equation  $Y \sim \beta_0 + \beta_1 * x_1 + \beta_2 * x_2 + \beta_3 * x_1 * x_2 \dots + \epsilon$ . Below the equation, four green arrows point upwards from the words "DATA", "INTERCEPT", "SIMPLE EFFECTS", and "INTERACTIONS" respectively. The word "DATA" has a single arrow pointing to the first term  $\beta_0$ . The word "INTERCEPT" has a single arrow pointing to the term  $\beta_0$ . The word "SIMPLE EFFECTS" has two arrows pointing to the terms  $\beta_1 * x_1$  and  $\beta_2 * x_2$ . The word "INTERACTIONS" has one arrow pointing to the term  $\beta_3 * x_1 * x_2$ .

## NOW WE CAN FINALLY BUILD THE FIRST LINEAR MODELS

CROSS-SECTIONAL ANALYSES

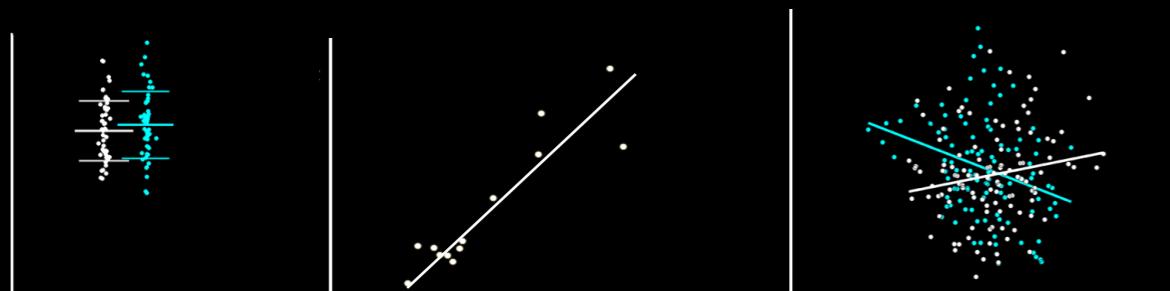
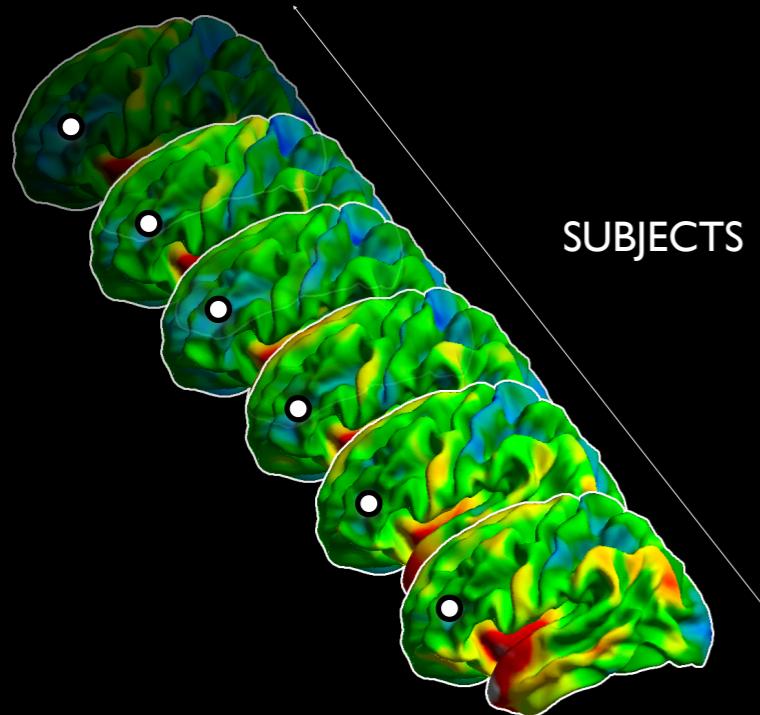


LONGITUDINAL ASSESSMENTS



# NOW WE CAN FINALLY BUILD THE FIRST LINEAR MODELS

CROSS-SECTIONAL ANALYSES

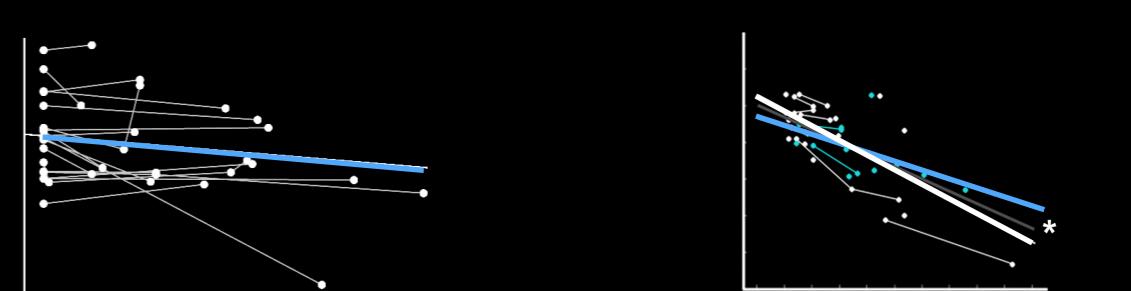
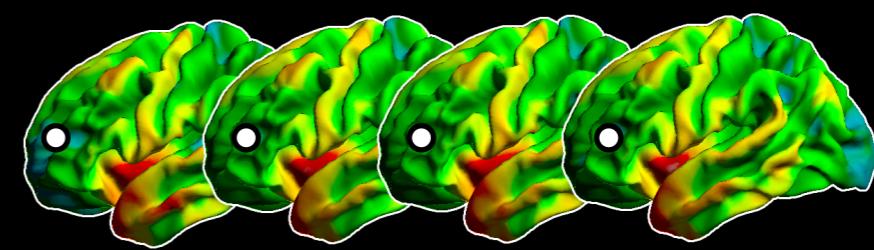


$$Y = I + G$$

$$Y = I + A$$

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

LONGITUDINAL ASSESSMENTS

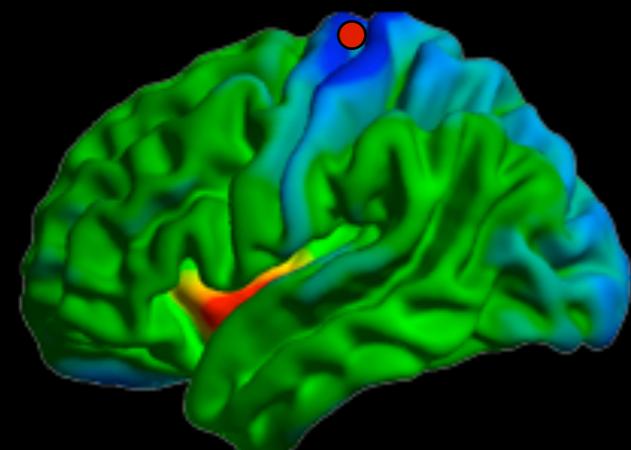


$$Y = I + r(S) + ISI$$

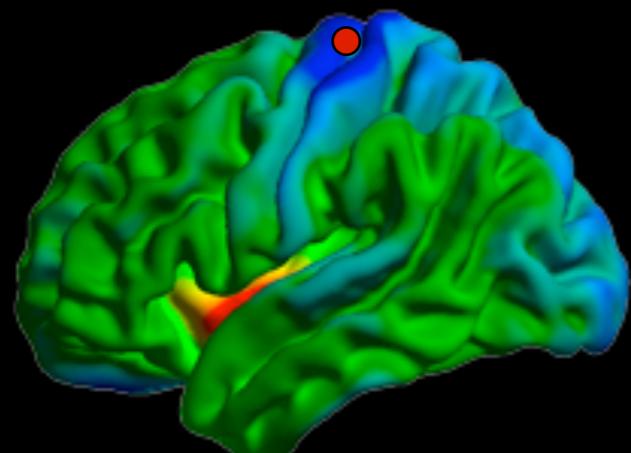
$$Y = I + r(S) + ISI + G + ISI \times G$$

# SURFSTAT

Controls

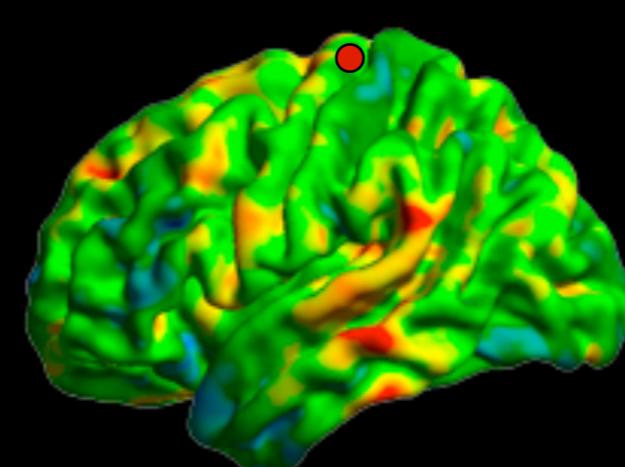


Patients

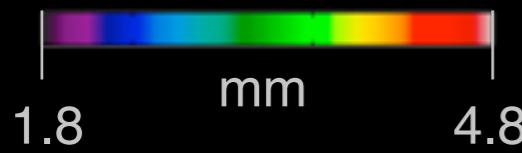
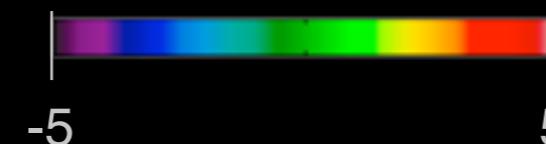
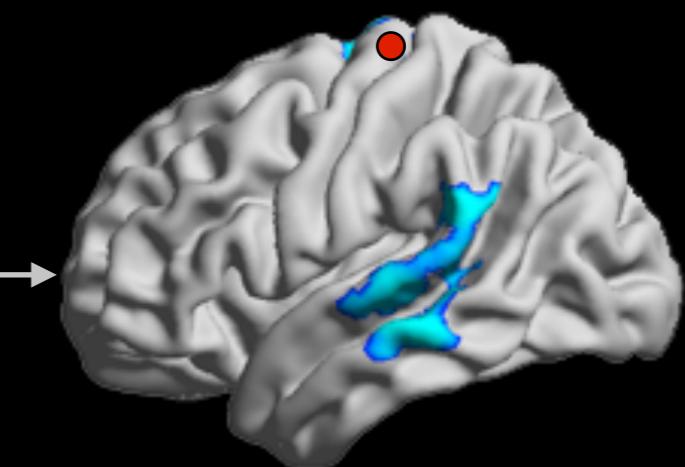


GLM

t-map

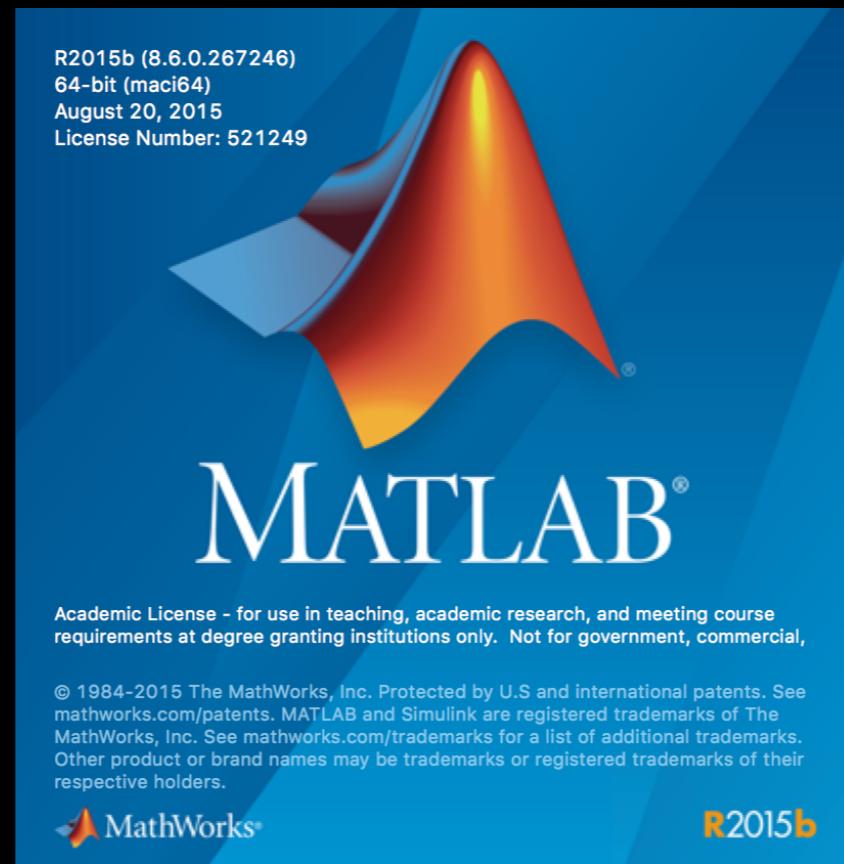


p-values



# 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
```

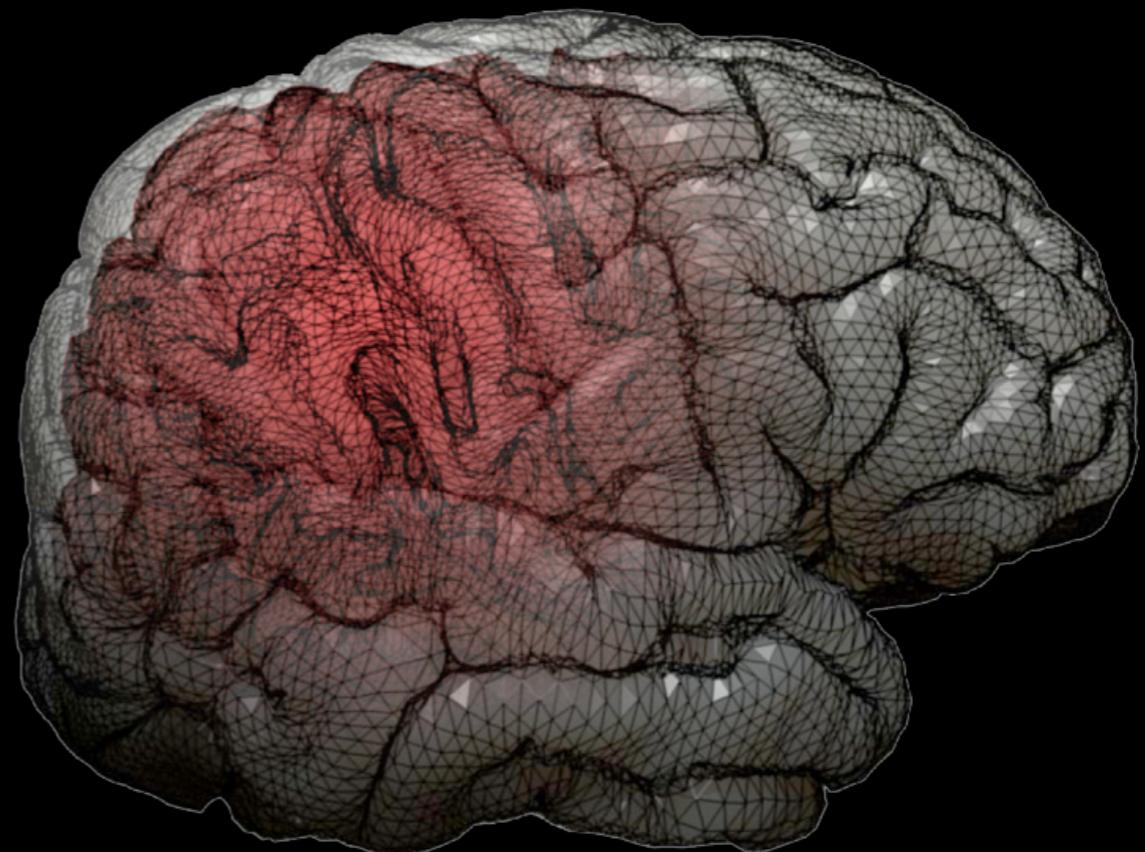
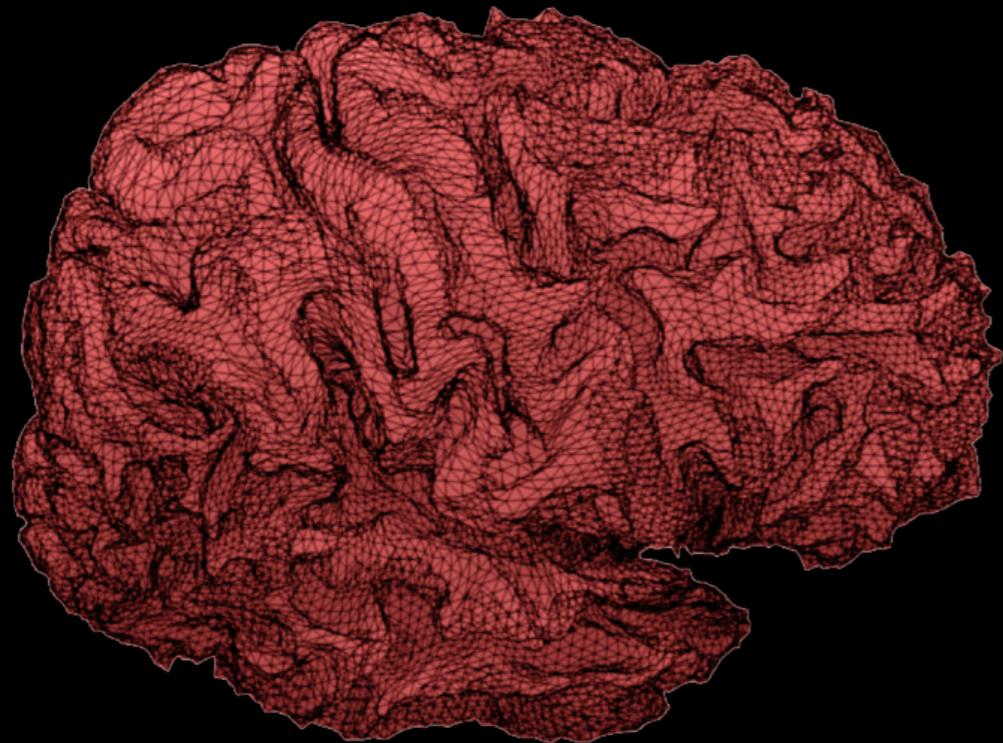
```
SW =
```

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

```
>> SP
```

```
SP =
```

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



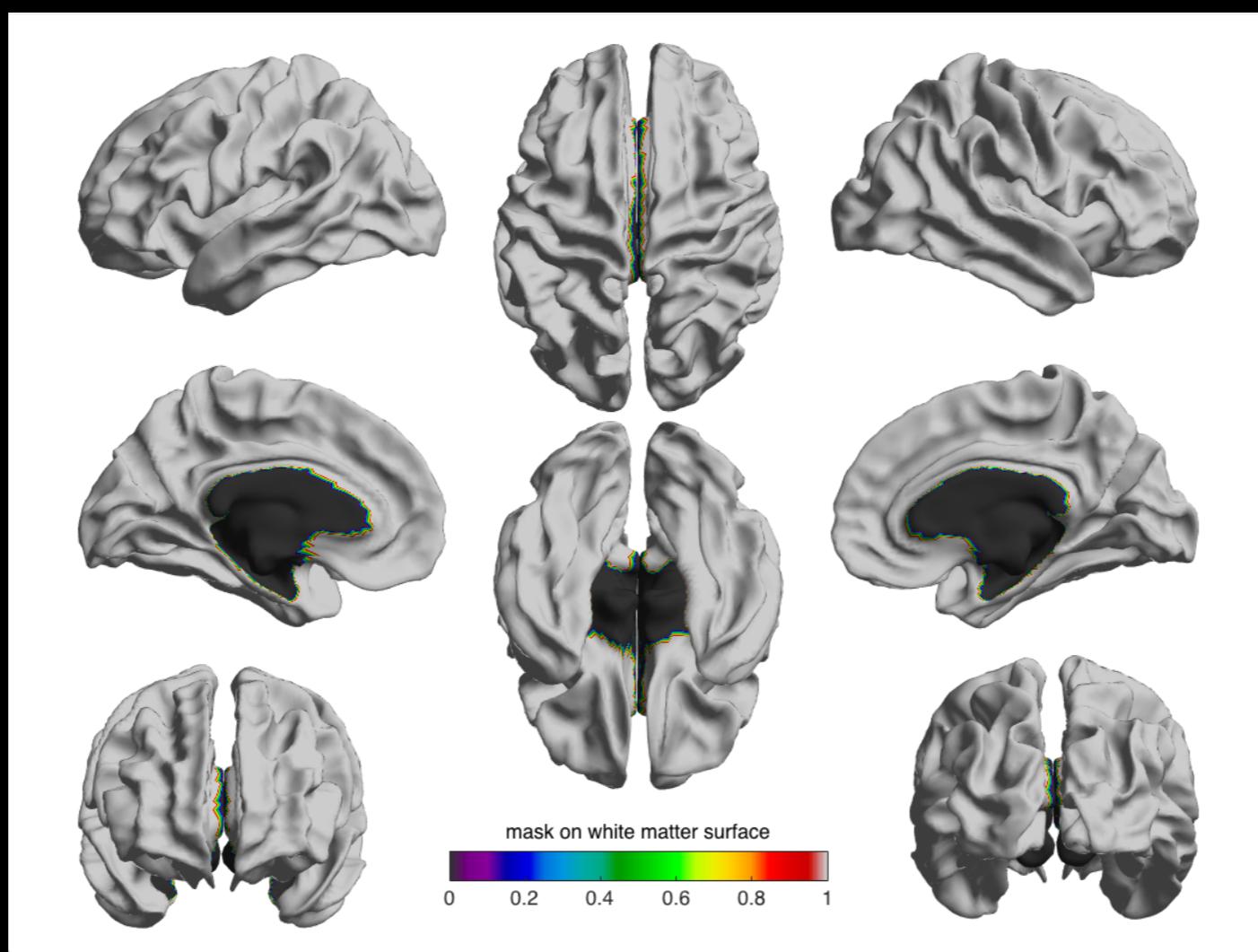
## LOAD SOME USEFUL FEATURES

```
%% 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
```

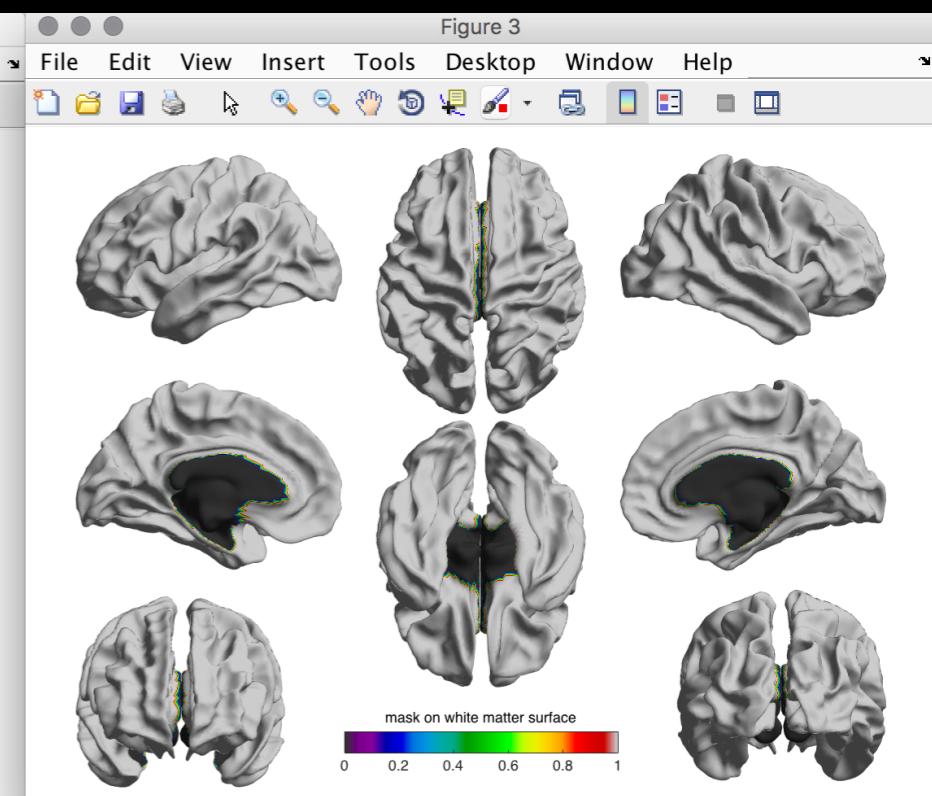
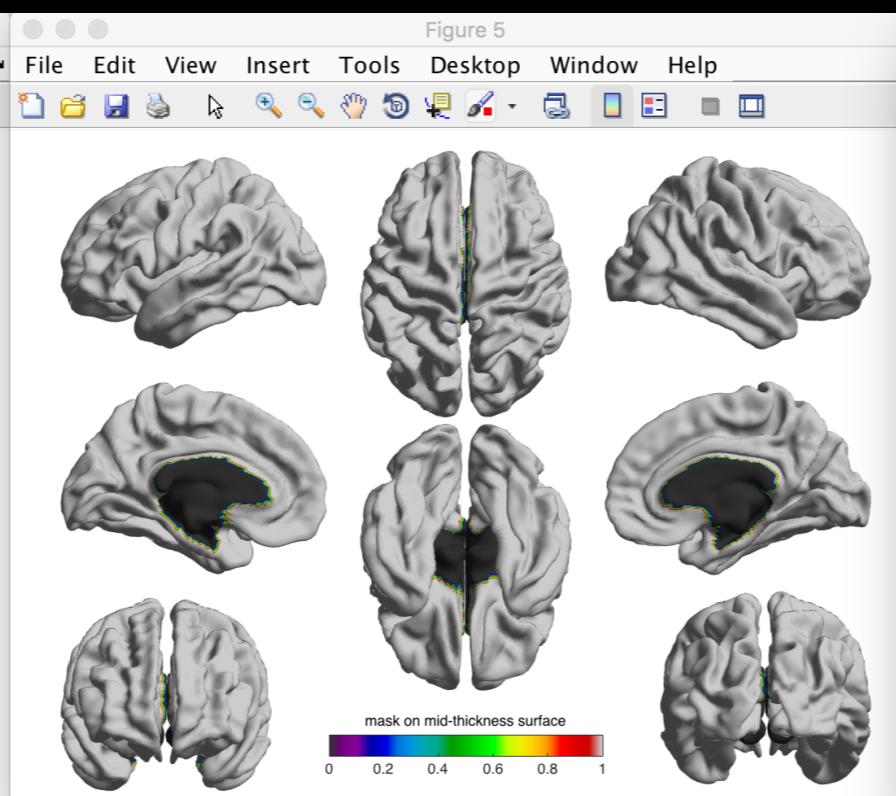
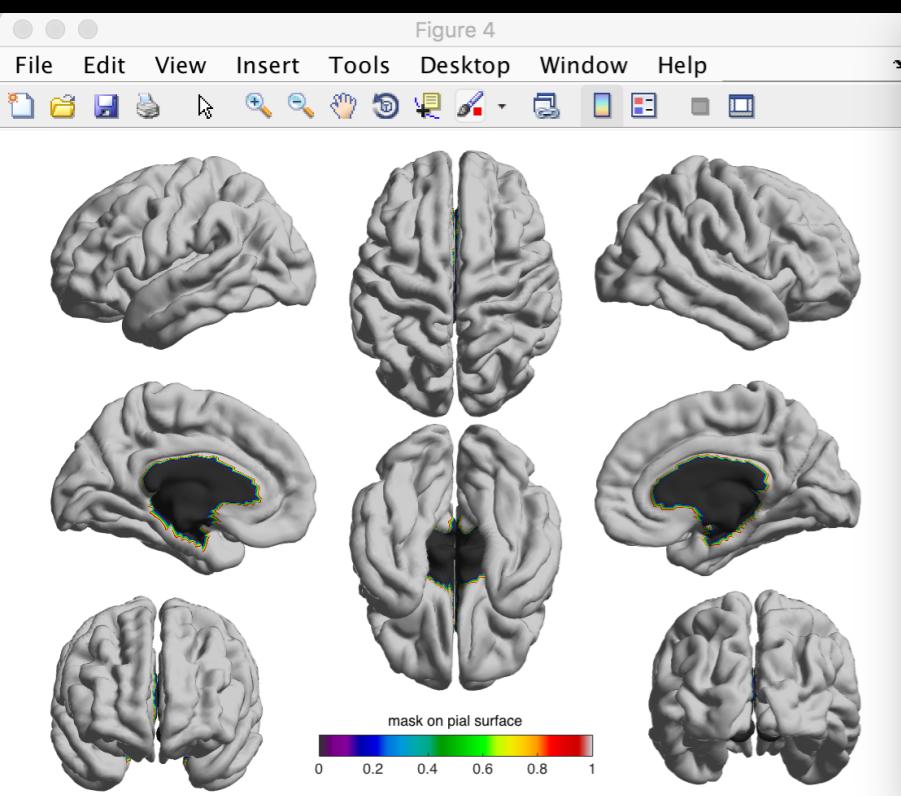
```
% first the brain mask
```

```
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

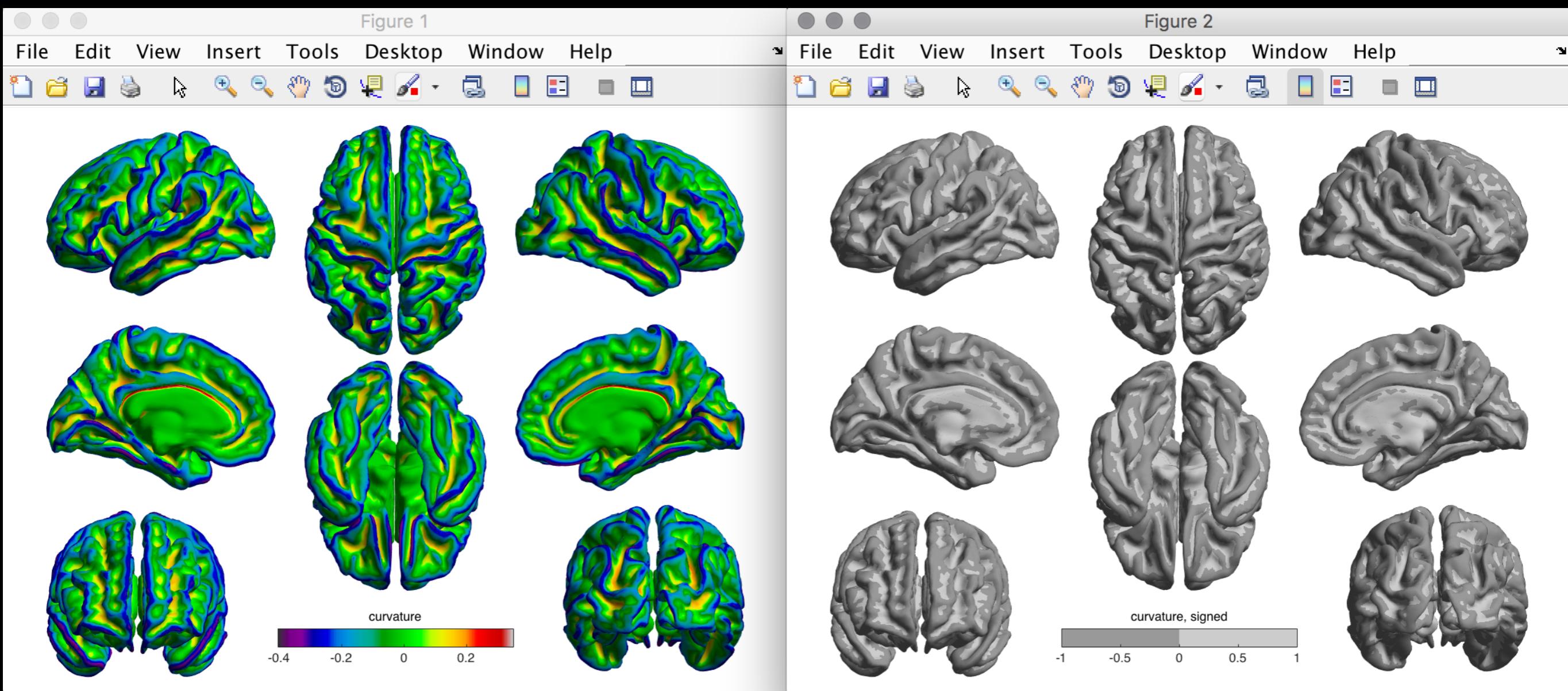


## WE CAN ALSO DISPLAY THE OTHER DATA

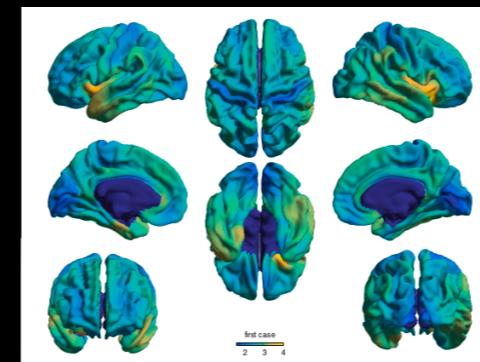
```
%% then the curvature data
f=figure,
SurfStatViewData(curv, SM, 'curvature')

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

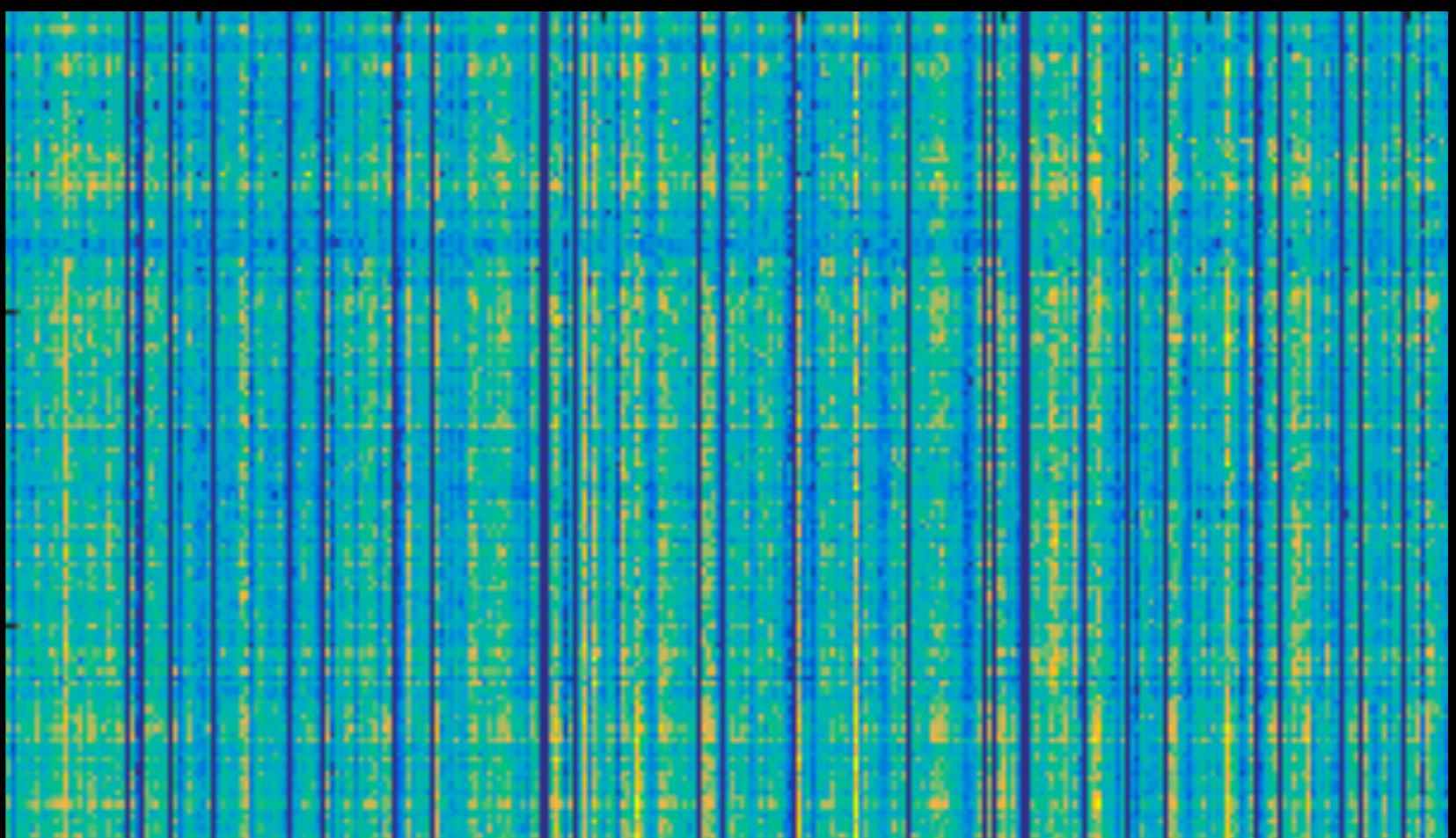
## WE CAN ALSO DISPLAY OTHER DATA



# 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



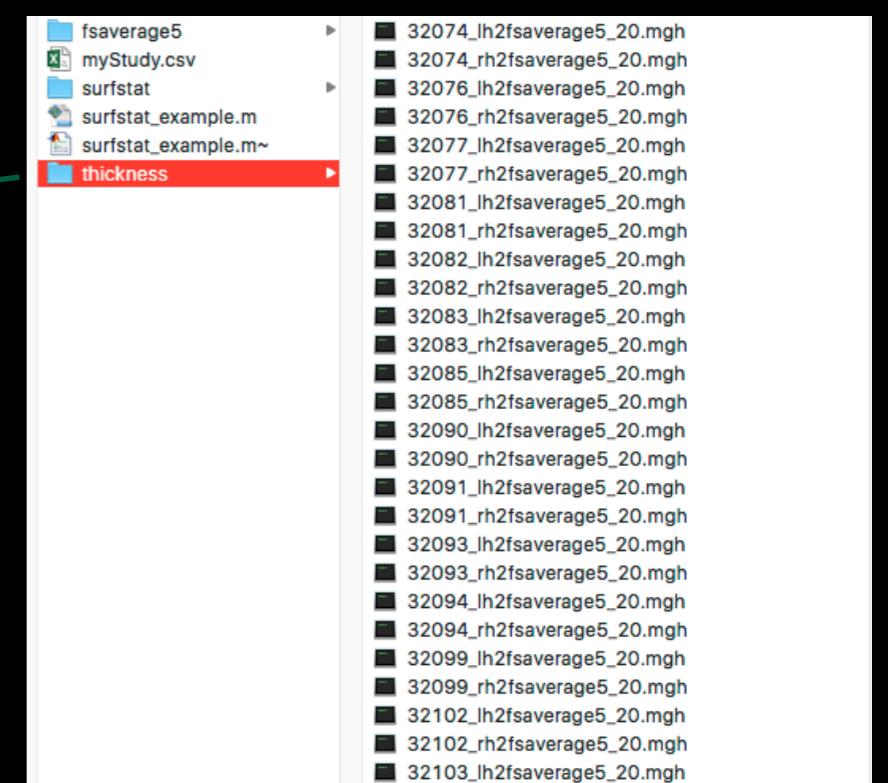
# NOW ITS TIME TO LOAD THE DATA OF THE INDIVIDUAL SUBJECTS

```
%% 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]);
```



A screenshot of a file browser window showing a list of files in a folder. The folder path is 'thickness'. The files listed are:

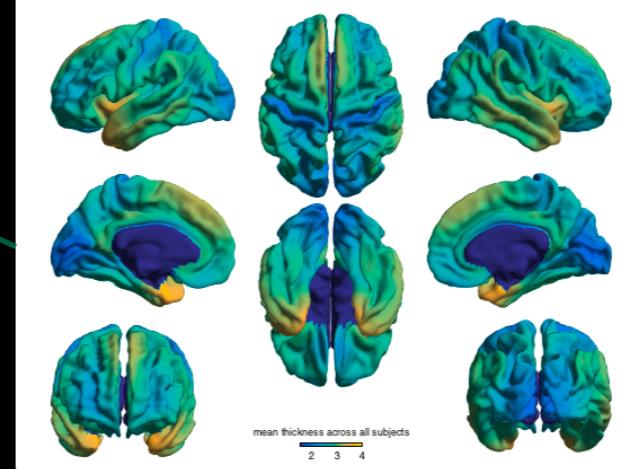
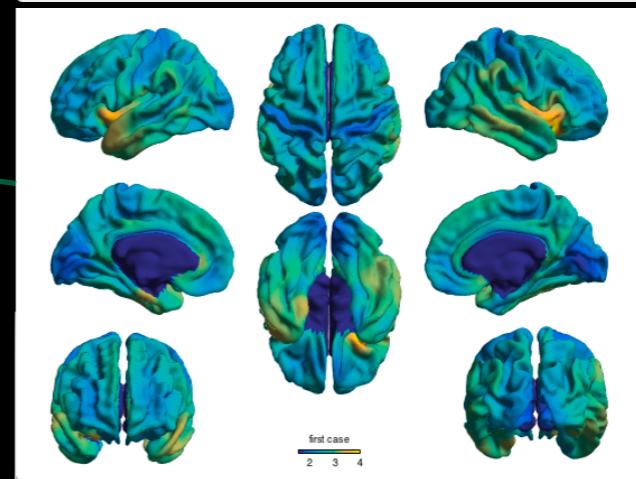
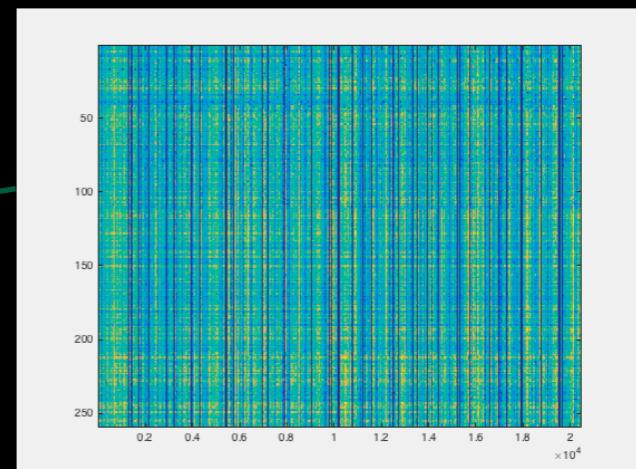
- fsaverage5
- myStudy.csv
- surfstat
- surfstat\_example.m
- surfstat\_example.m~
- thickness
- 32074\_lh2fsaverage5\_20.mgh
- 32074\_rh2fsaverage5\_20.mgh
- 32076\_lh2fsaverage5\_20.mgh
- 32076\_rh2fsaverage5\_20.mgh
- 32077\_lh2fsaverage5\_20.mgh
- 32077\_rh2fsaverage5\_20.mgh
- 32081\_lh2fsaverage5\_20.mgh
- 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,...

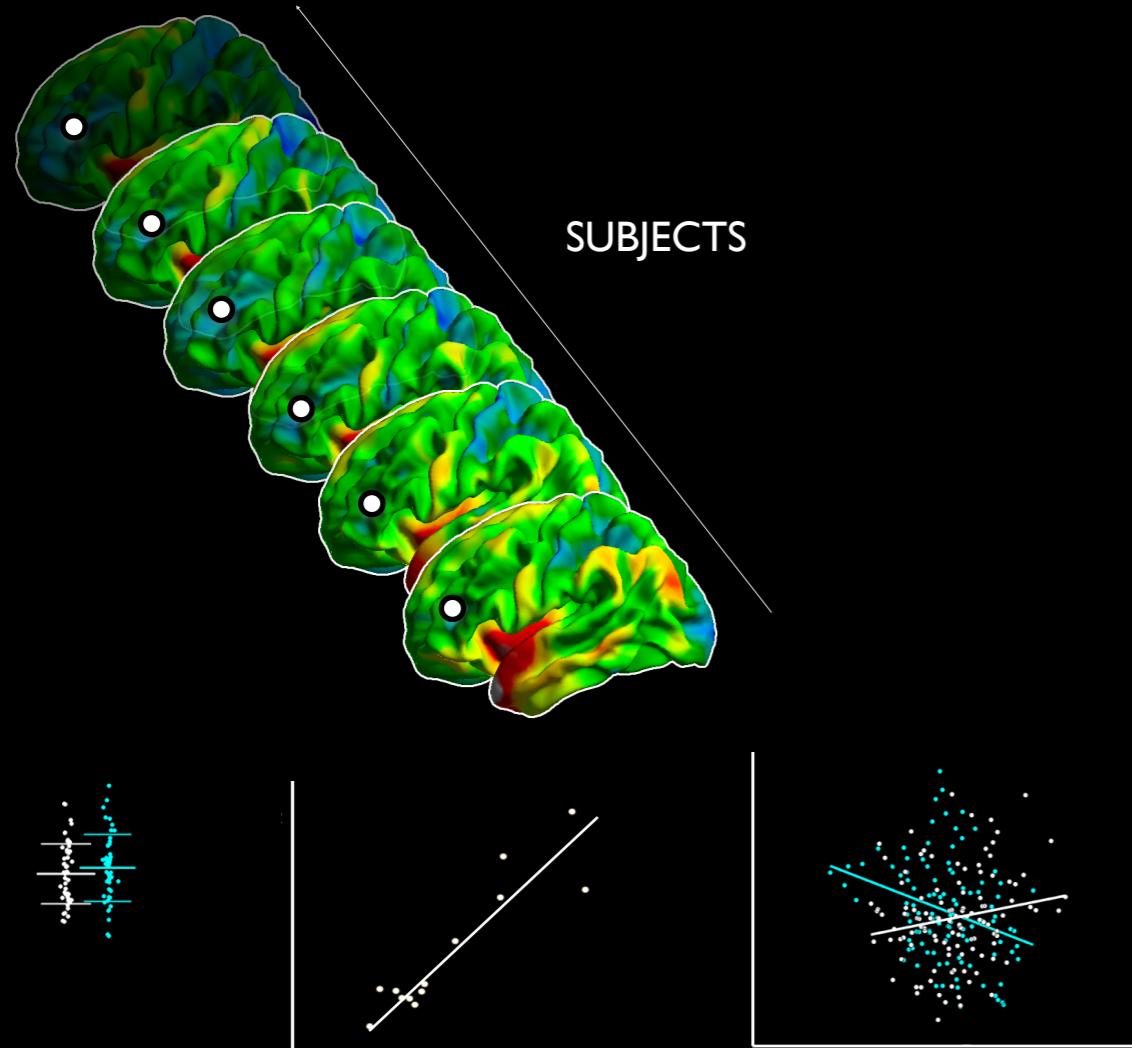
are just specific instances of the linear model of the form

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

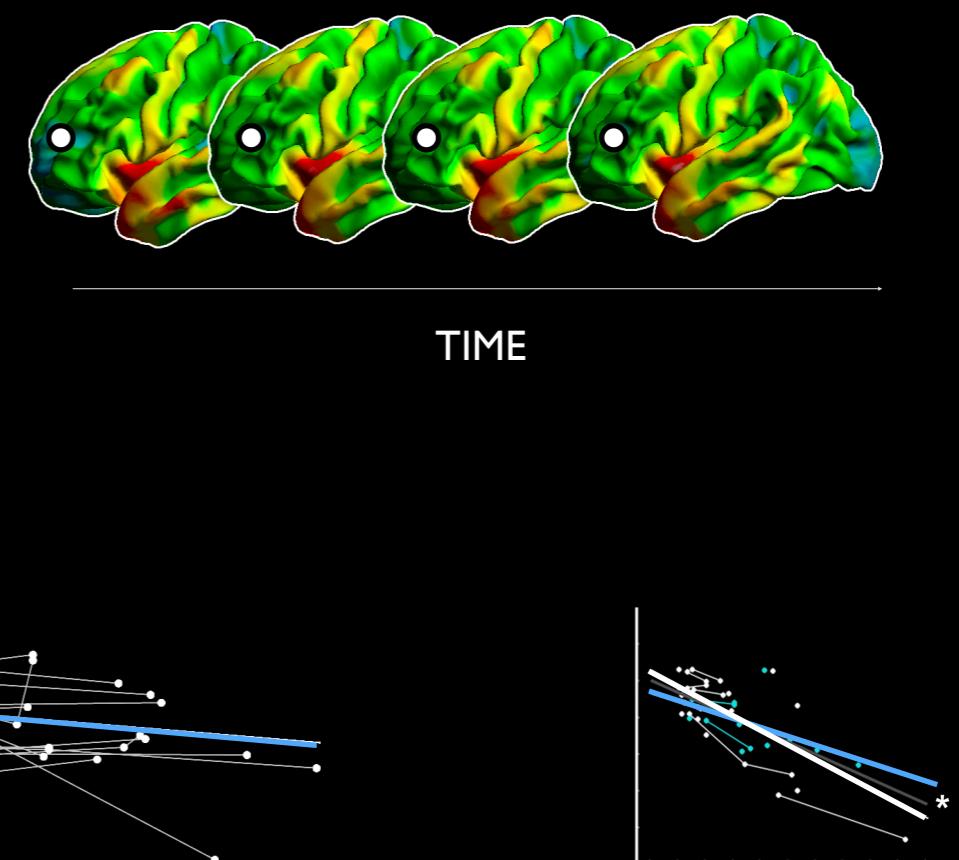
The diagram illustrates the components of a linear model equation. At the top is the equation  $Y \sim \beta_0 + \beta_1 * x_1 + \beta_2 * x_2 + \beta_3 * x_1 * x_2 \dots + \epsilon$ . Below the equation, four green arrows point upwards from the words "DATA", "INTERCEPT", "SIMPLE EFFECTS", and "INTERACTIONS" respectively. The word "DATA" has a single arrow pointing to the first term  $\beta_0$ . The word "INTERCEPT" has a single arrow pointing to the term  $\beta_0$ . The word "SIMPLE EFFECTS" has two arrows pointing to the terms  $\beta_1 * x_1$  and  $\beta_2 * x_2$ . The word "INTERACTIONS" has one arrow pointing to the term  $\beta_3 * x_1 * x_2$ .

# NOW WE CAN FINALLY BUILD THE FIRST LINEAR MODELS

CROSS-SECTIONAL ANALYSES

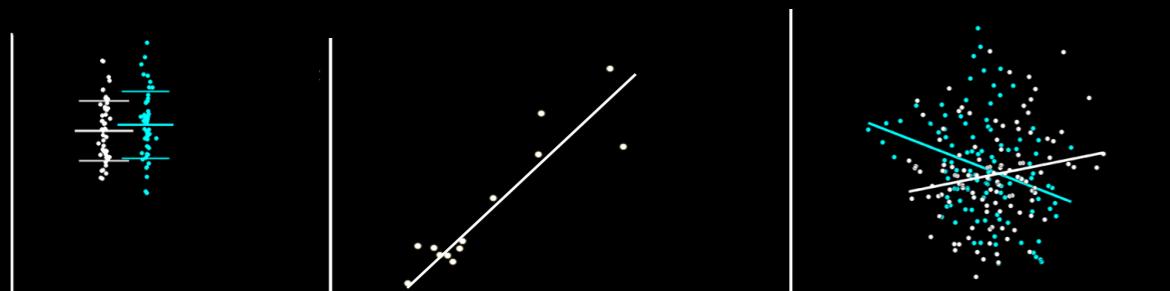
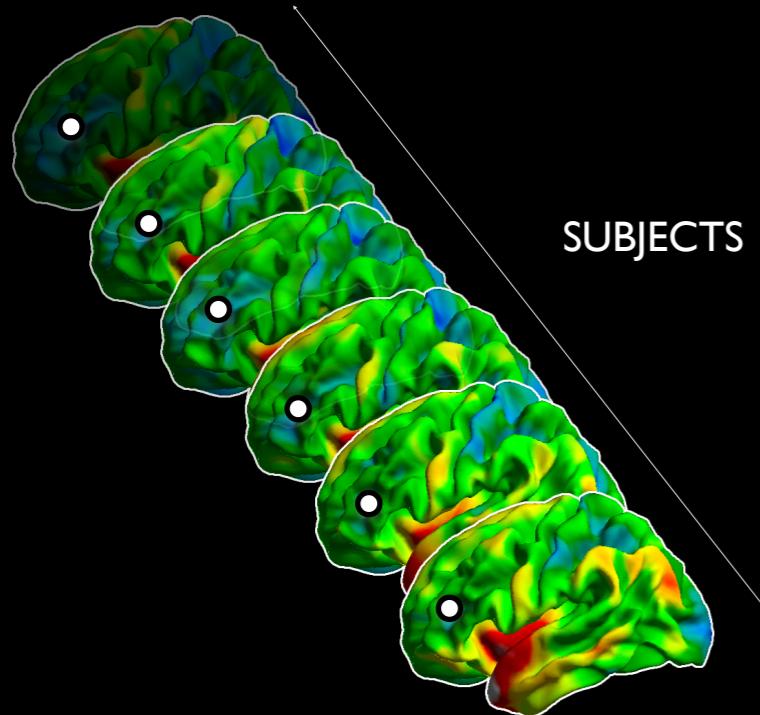


LONGITUDINAL ASSESSMENTS



# NOW WE CAN FINALLY BUILD THE FIRST LINEAR MODELS

CROSS-SECTIONAL ANALYSES

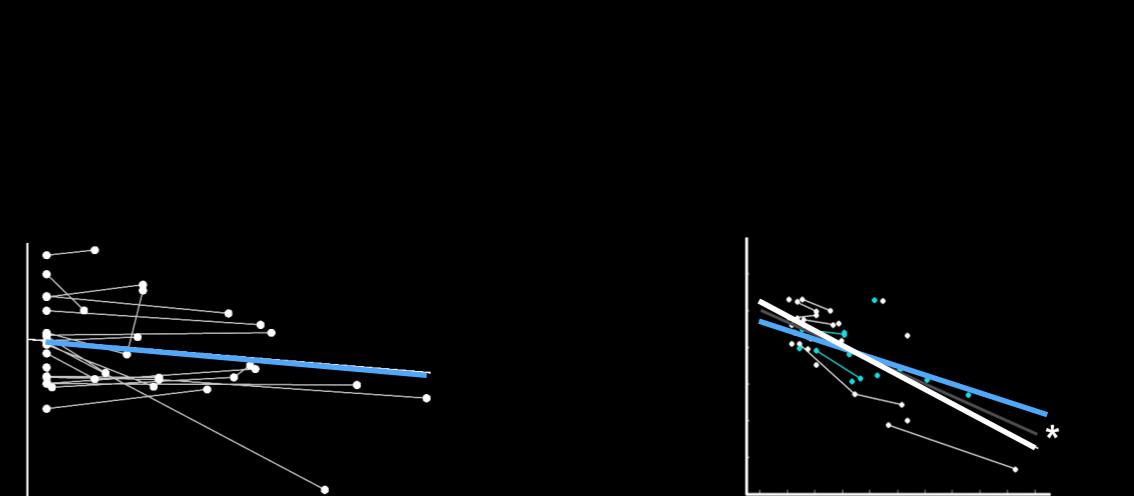
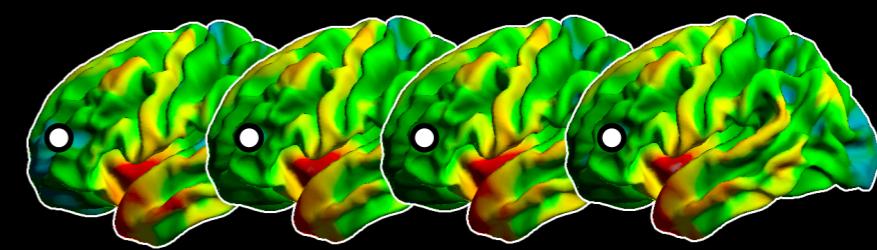


$$Y = I + G$$

$$Y = I + A$$

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

LONGITUDINAL ASSESSMENTS



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

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

# NOW WE CAN FINALLY BUILD THE FIRST LINEAR MODELS

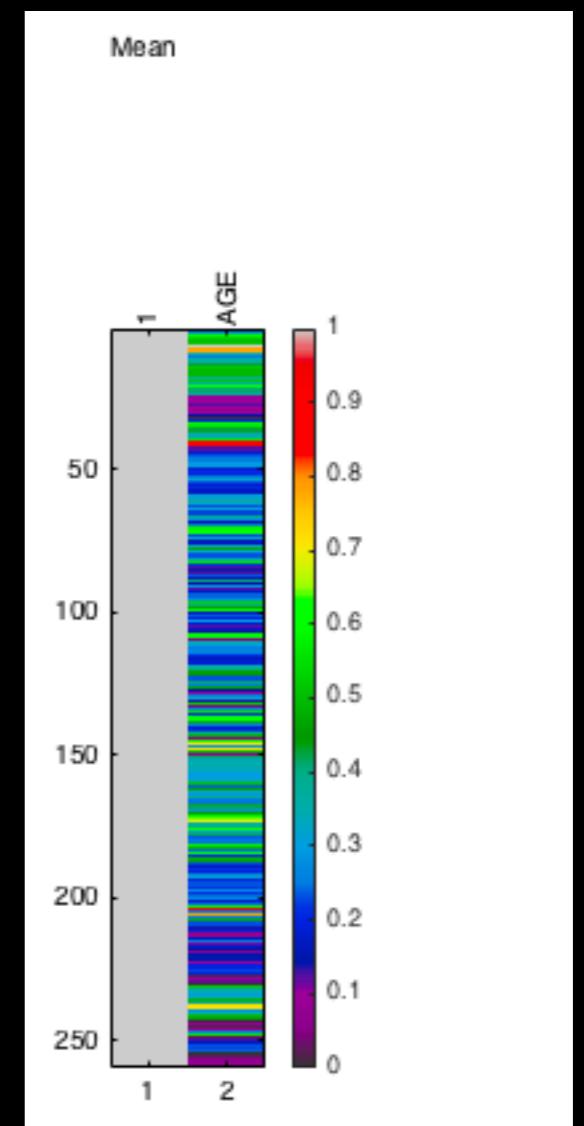
```
%% now we can finally do a firs models: lets look at effects of age

% 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);

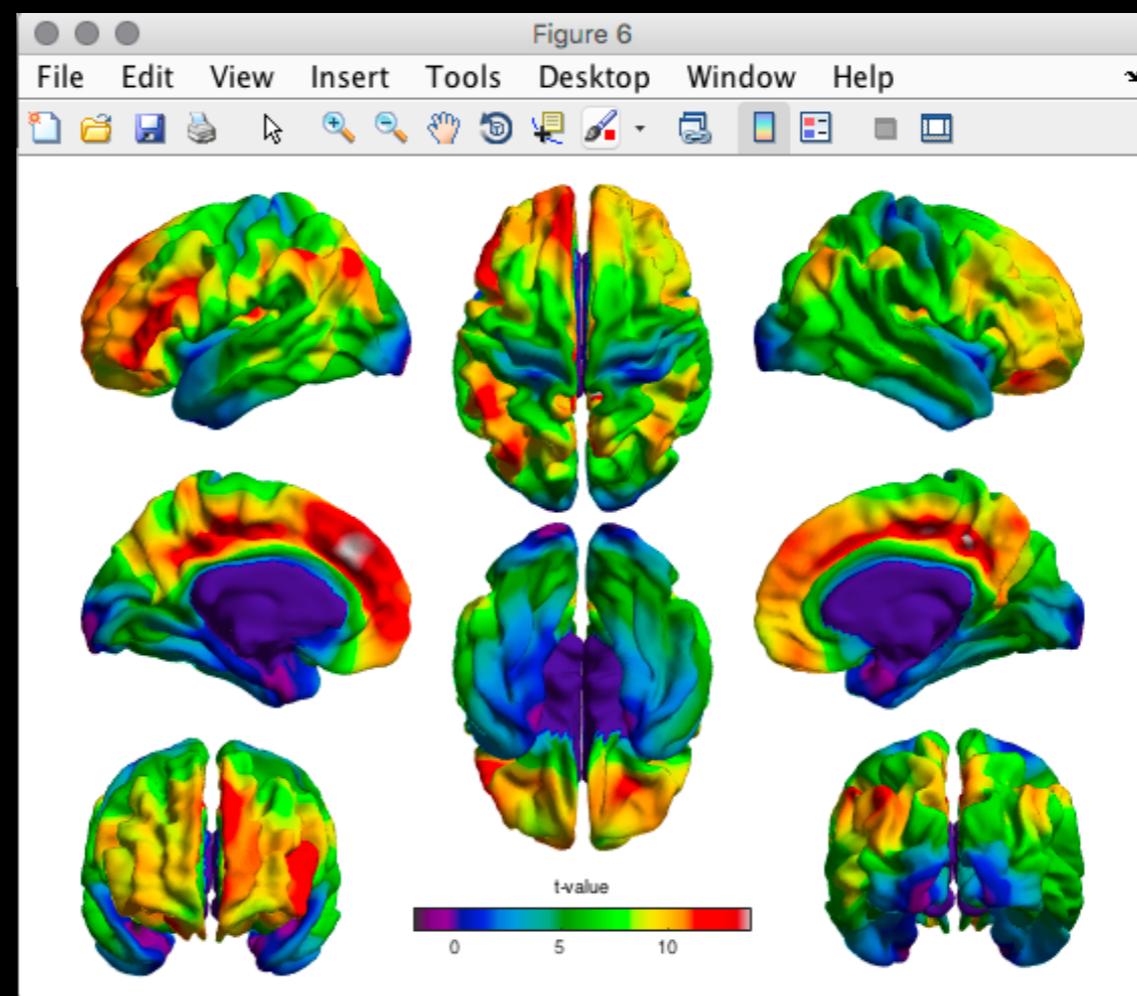
% specifiy contrast
slm = SurfStatT(slm, -AGE)
```



# 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])

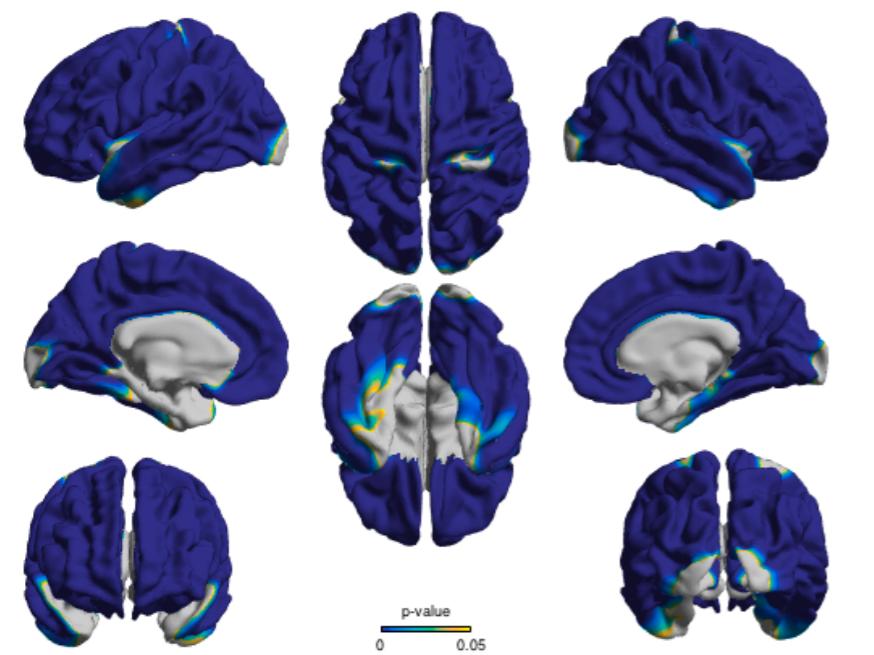
% multiple comparison correction: Bonferroni
p = 1-tcdf(slm.t,slm.df);
p = p*size(p,2);
f=figure
    SurfStatViewData(p, SM, 'Bonferroni p-value')
    SurfStatColLim([0 0.05])
    colormap([parula; .8 .8 .8])

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

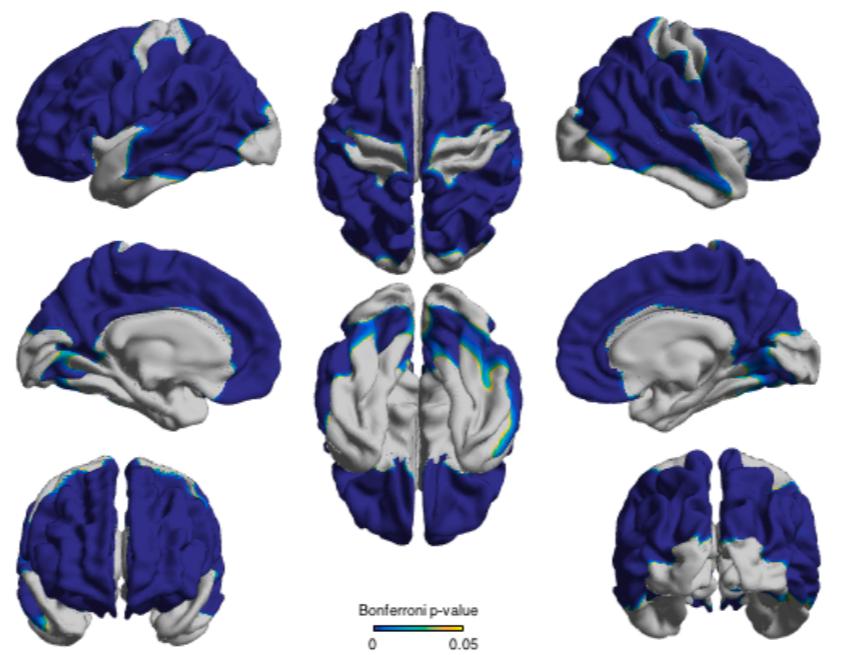
% multiple comparions using random field theory
pval = SurfStatP(slm,mask);
f=figure
    SurfStatView(pval, SM, 'rft')
```

# AND CORRECT FOR MULTIPLE COMPARISONS

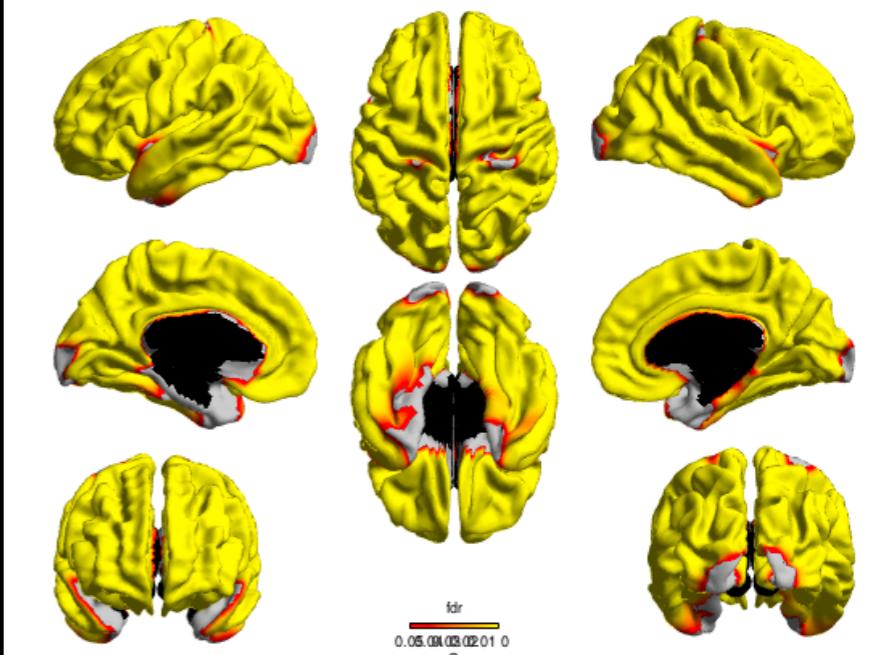
uncorrected



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

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

## 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

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

## 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
```

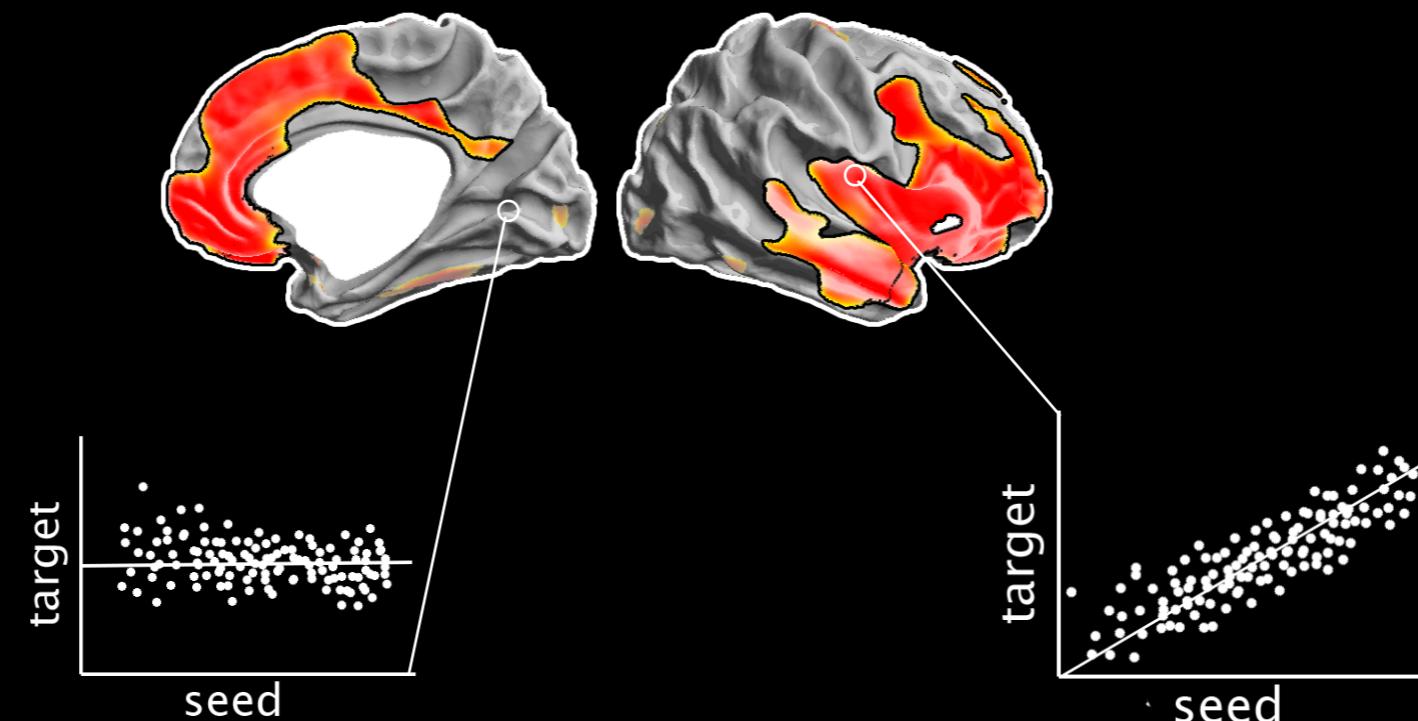
```
slm = SurfStatLinMod(T, M, SW)
```

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

interaction model, assumes different age effect across groups

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

# STRUCTURAL COVARIANCE ANALYSIS



Lerch et al. (2006) *NeuroImage*

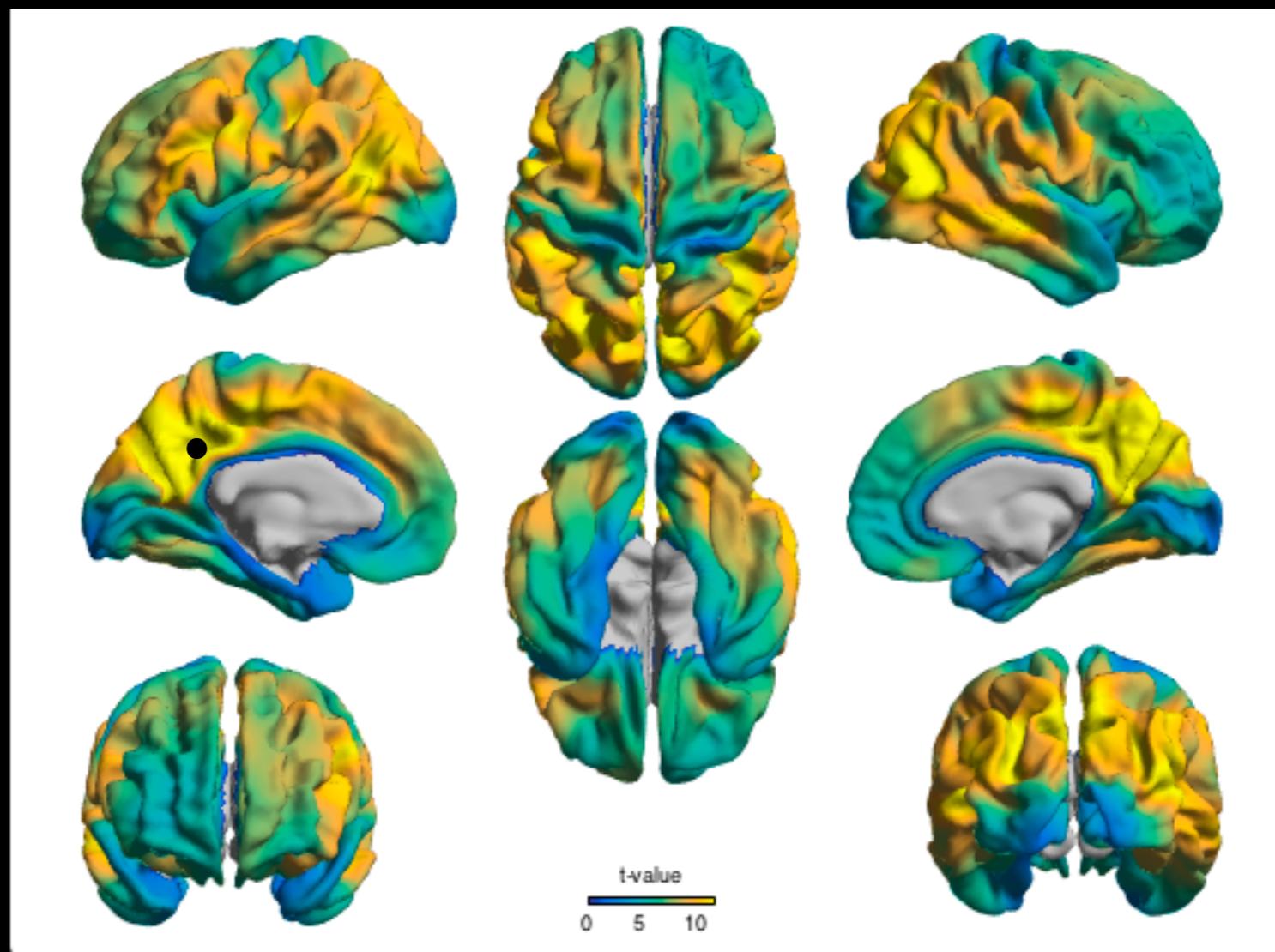
Alexander-Bloch et al. (2013) *Nat Rev Neurosci*

## 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>

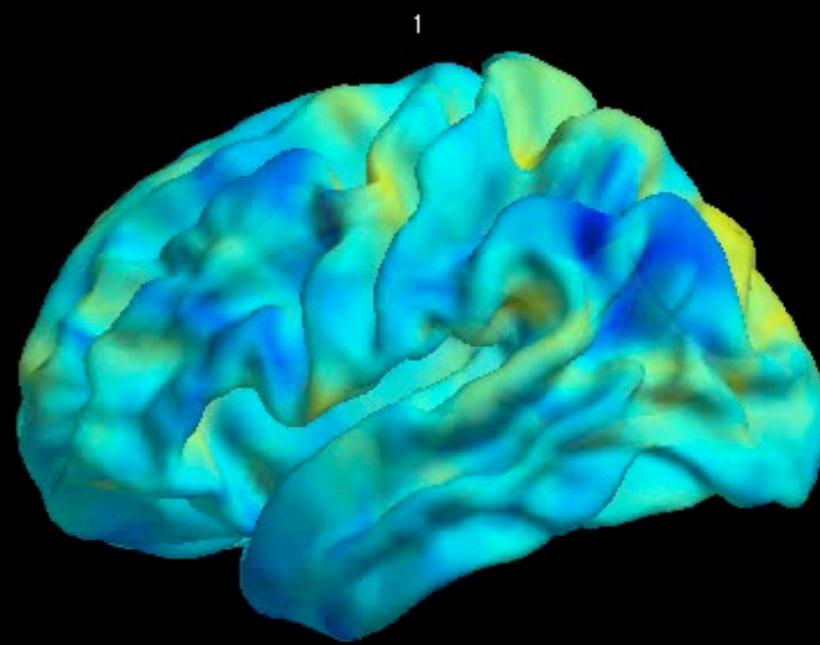
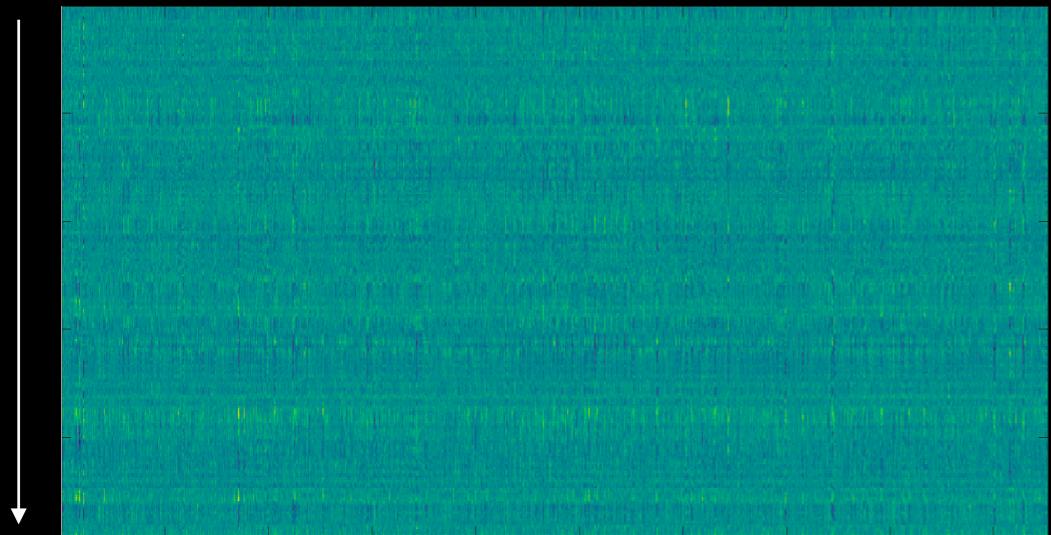
# CLUSTERING TECHNIQUES

NEUR-608

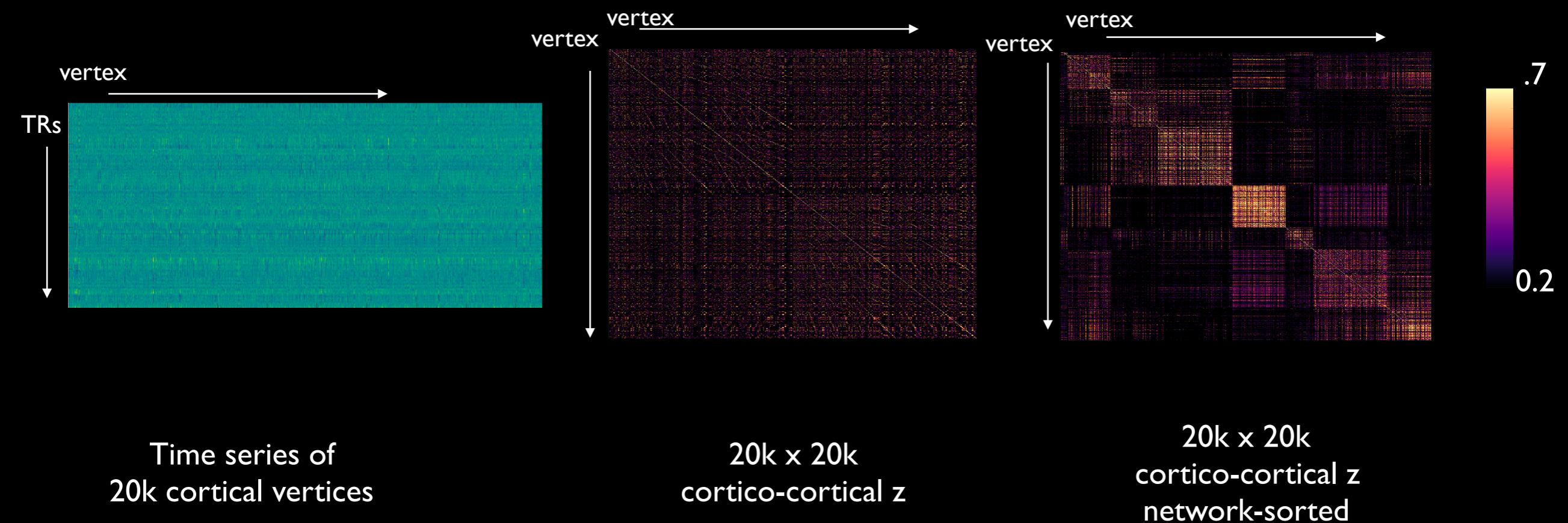


# BRAIN AT REST

Space/Vertices  
Time/TRs



# IMAGING DATA: HIGH DIMENSIONAL AND MULTIVARIATE



# HOW TO IDENTIFY NETWORKS HOW TO REDUCE DATA DIMENSIONALITY

## COMPRESSION

LINEAR (PCA, ICA, FA, MDS,...)

NON-LINEAR (LE, DME,.....)

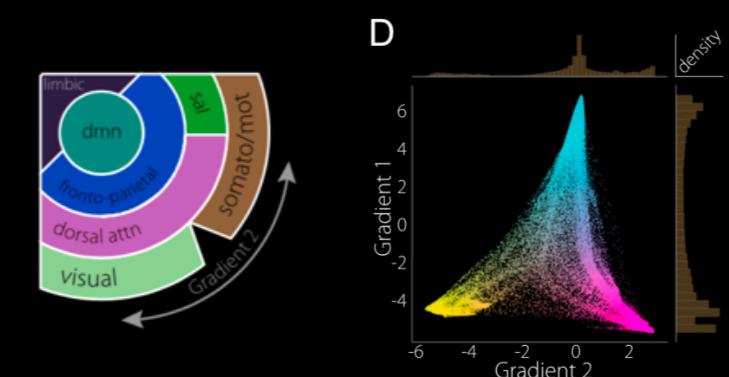
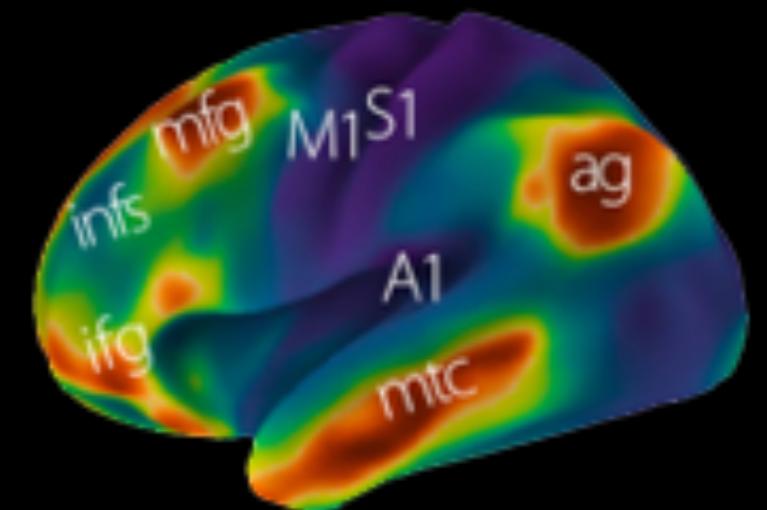
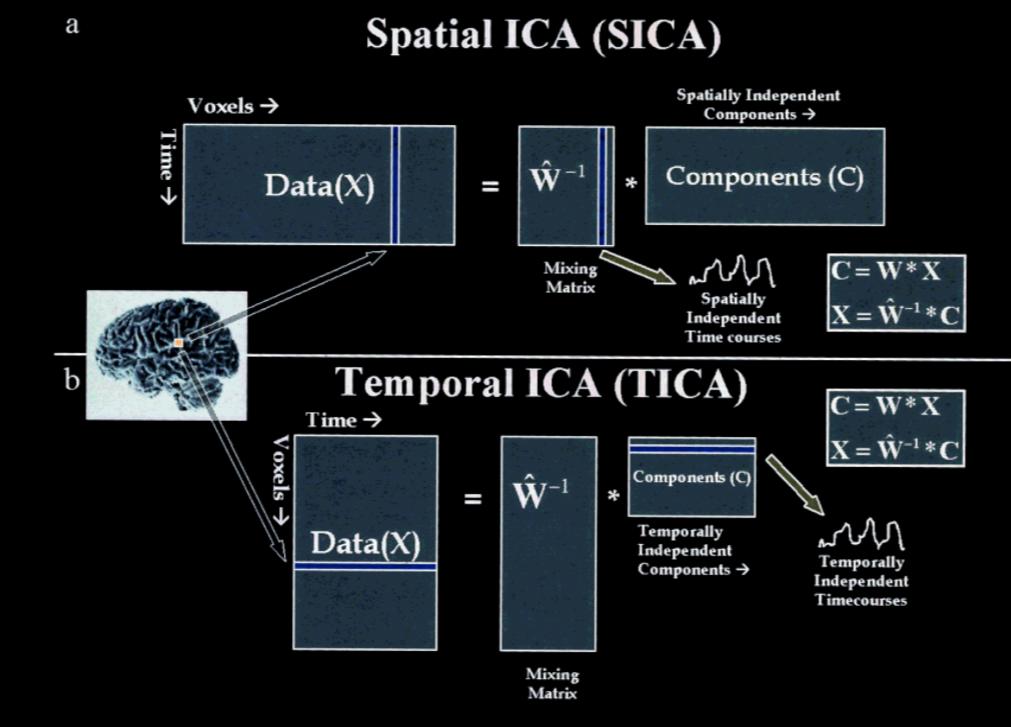
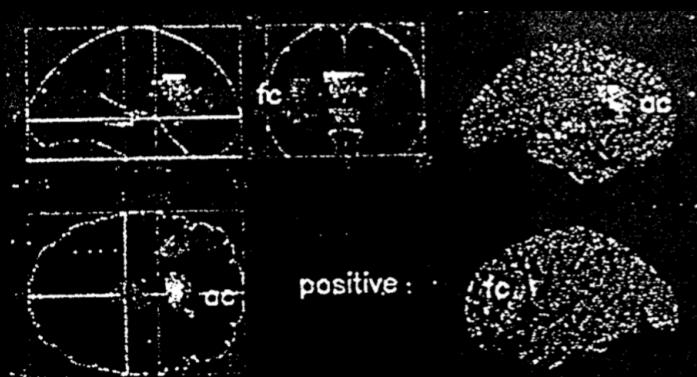
## CLUSTERING

K-MEANS

HIERARCHICAL

SPECTRAL

# COMPRESSION

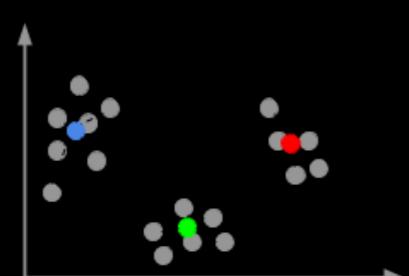
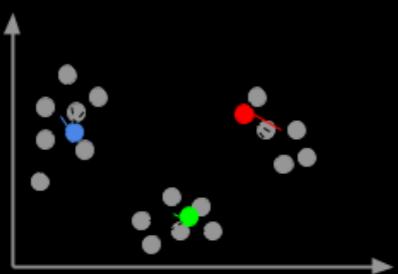
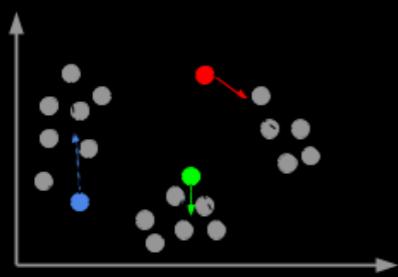
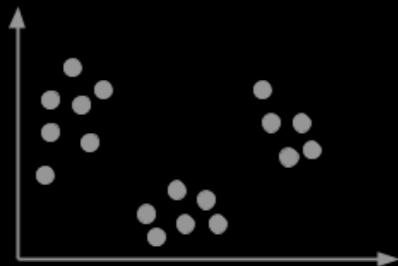


Friston 1993

McKeown et al 1998 HBM  
Calhoun 2001  
Beckmann 2012 NIMG

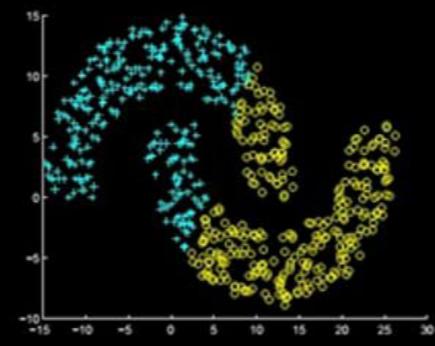
Margulies et al. 2016 PNAS

# CLUSTERING ALGORITHMS

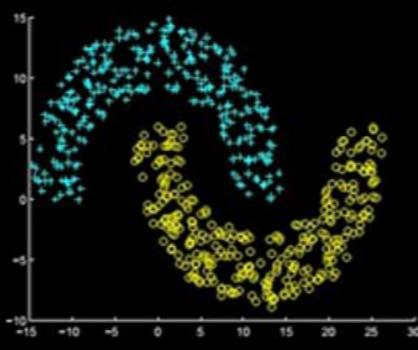


K-means clustering

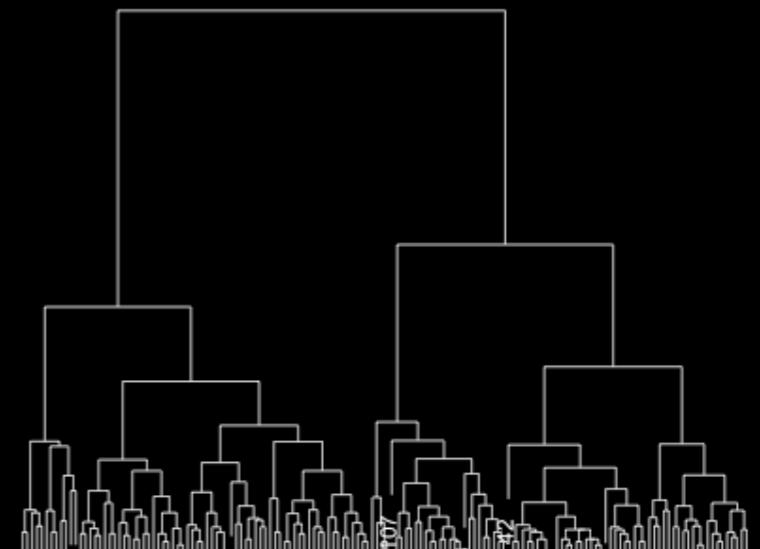
Spectral clustering  
Shi Malik 2000  
Von Luxburg 2007



(a) K-means

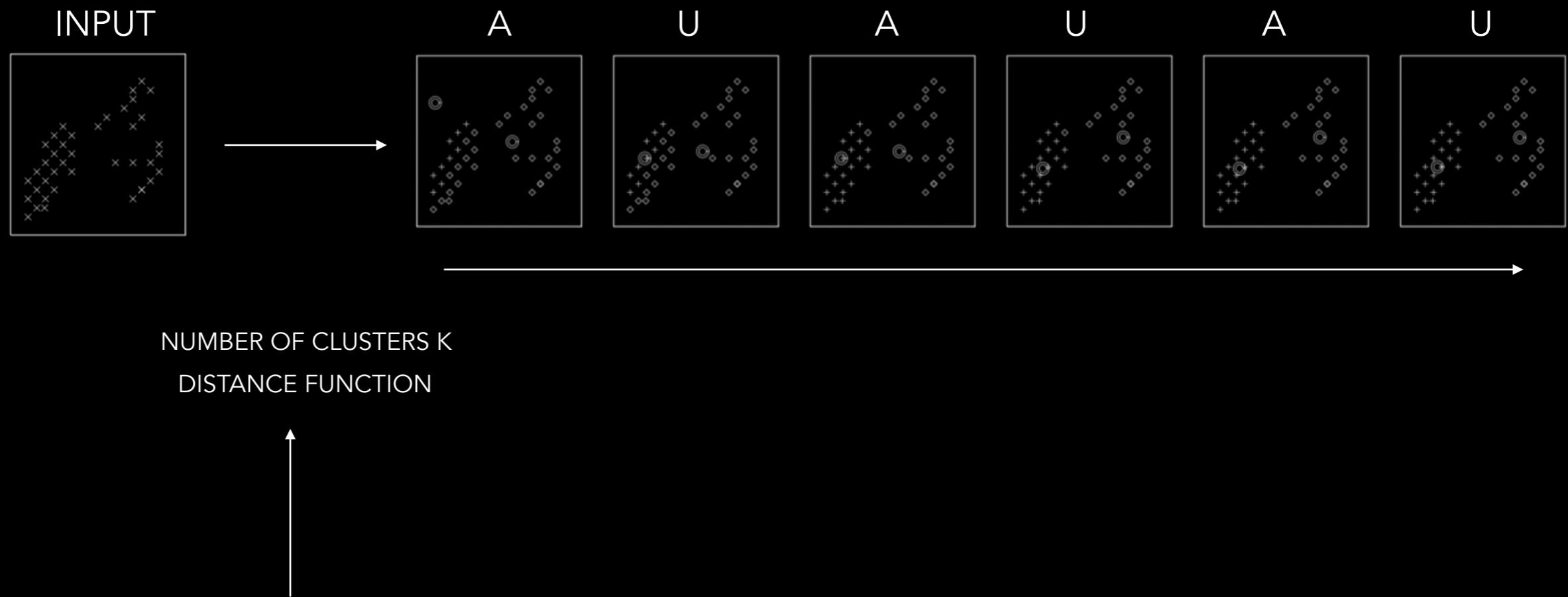


(b) Spectral Clustering



Hierarchical clustering

# K-MEANS

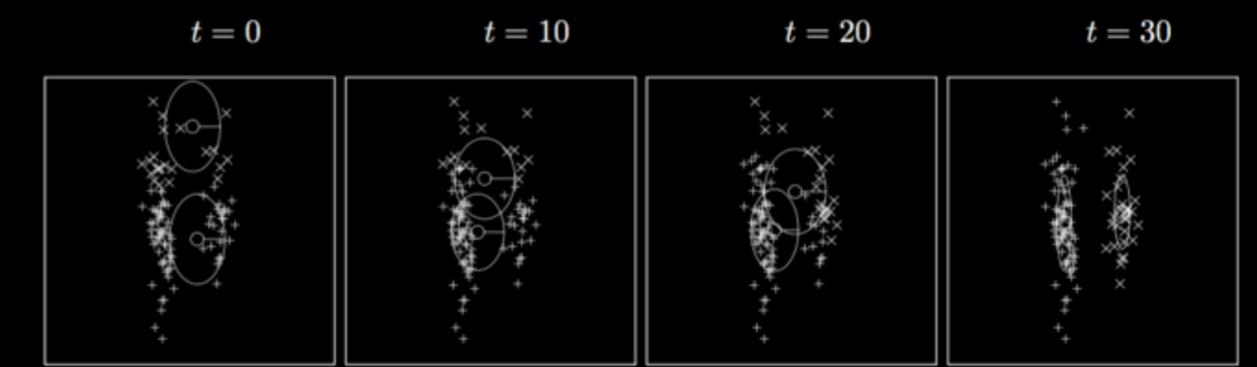
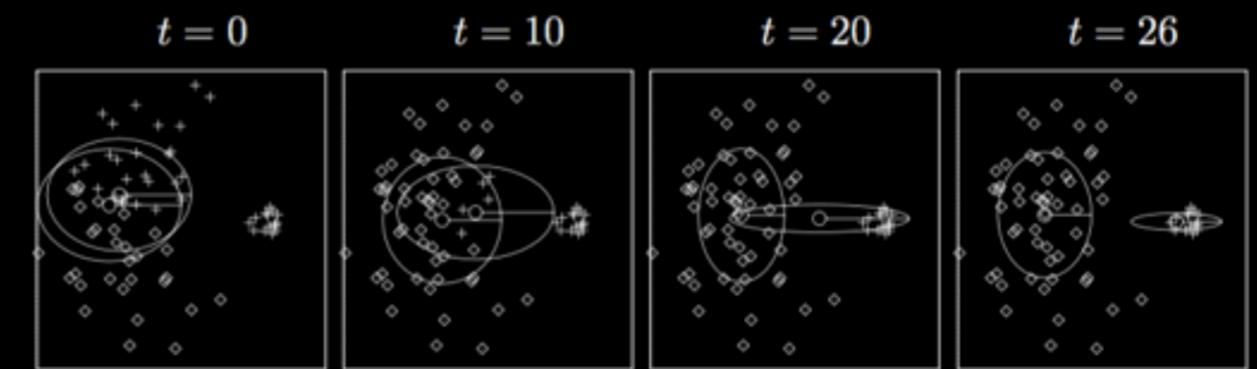
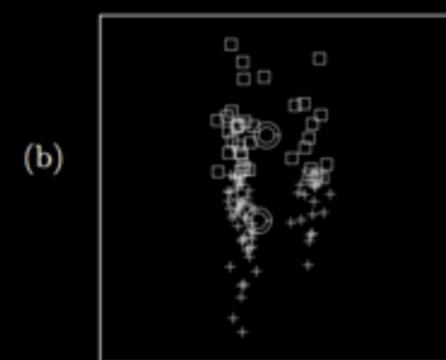
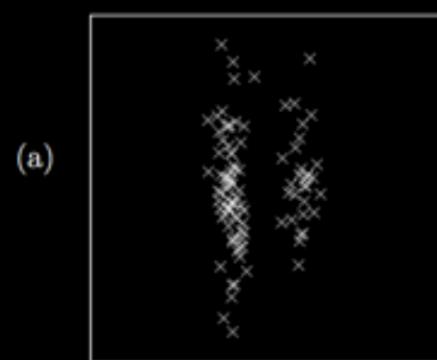
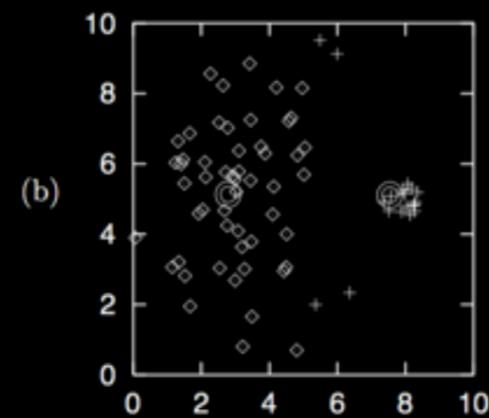
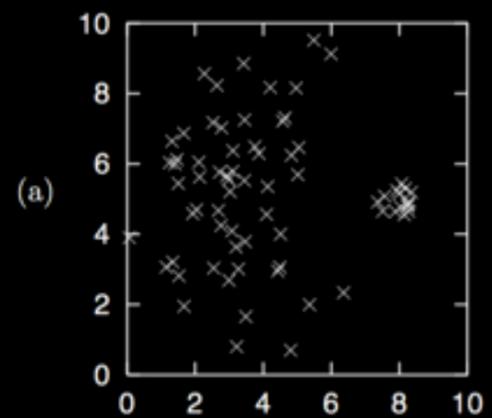


# CHALLENGE I: INITIALIZATION



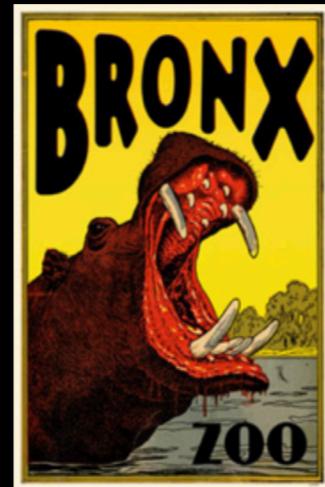
SOME SOLUTIONS: RUN MULTIPLE TIME AND IDENTIFY MOST STABLE SOLUTION; BOOTSTRAP

## CHALLENGE 2: VERY DISSIMILAR CLUSTERS



SOME SOLUTIONS: SOFTEN THE ASSIGNMENT RULES

# ON DISTANCE MEASURES



A ZOO OF DISTANCE MEASURES

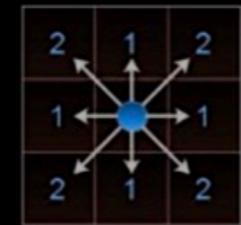
EUCLIDEAN, MANHATTAN, CHEBYCHEW  
MINKOWSKI, CANBERRA, MAXIMUM, COSINE

Euclidean Distance



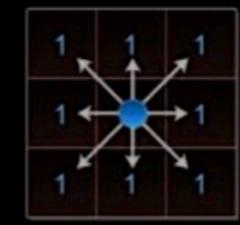
$$\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

Manhattan Distance



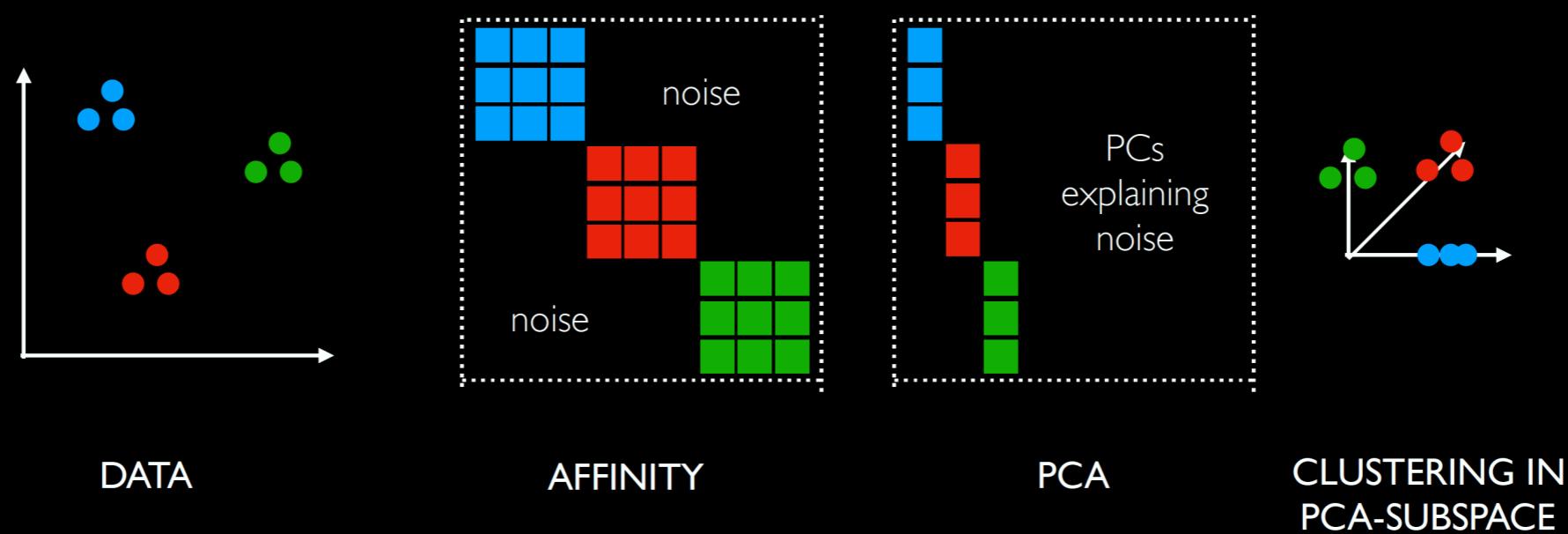
$$\|x_1 - x_2\| + |y_1 - y_2|$$

Chebyshev Distance

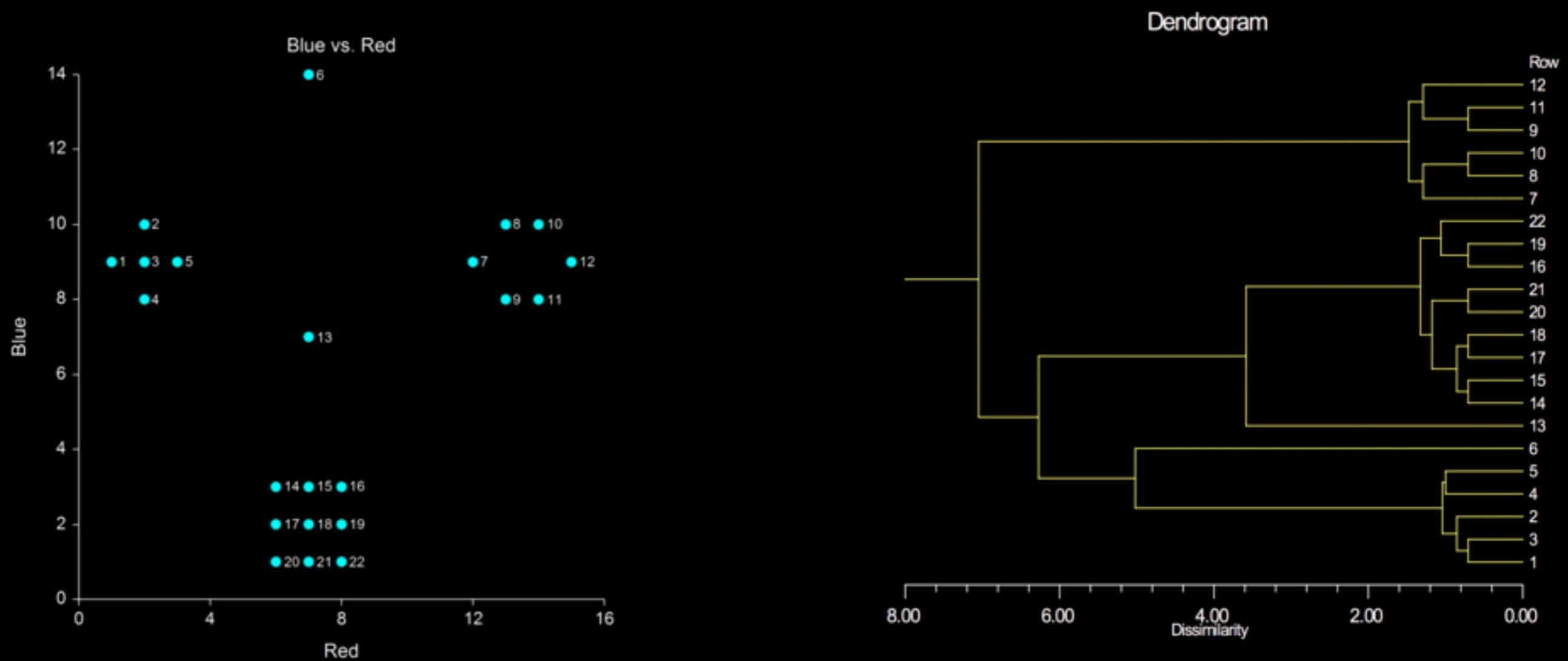


$$\max(|x_1 - x_2|, |y_1 - y_2|)$$

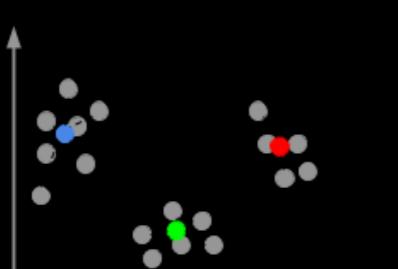
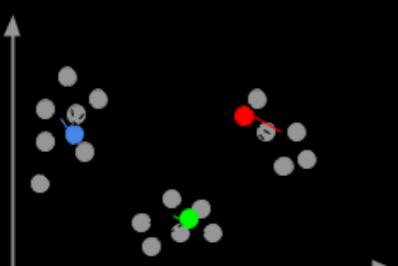
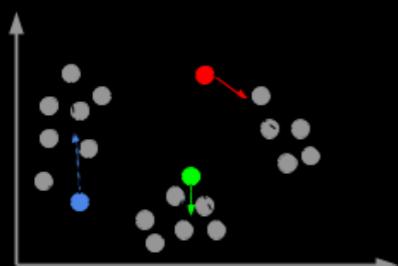
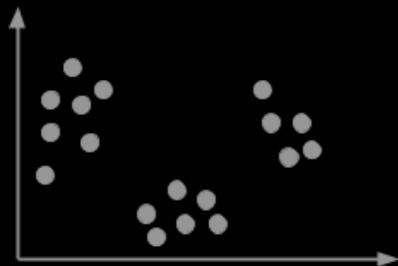
# SPECTRAL CLUSTERING



# HIERARCHICAL CLUSTERING

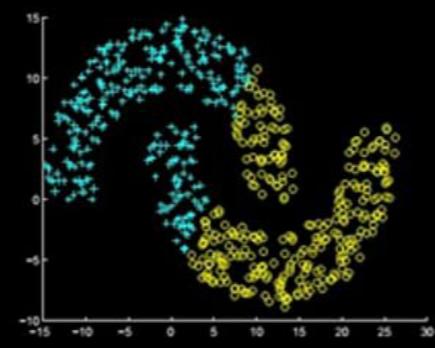


# CLUSTERING ALGORITHMS

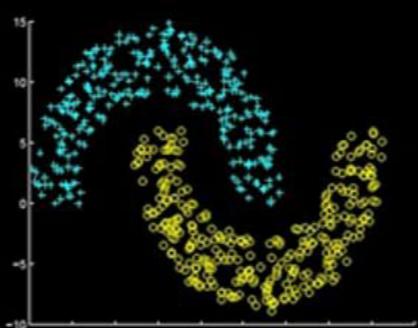


K-means clustering

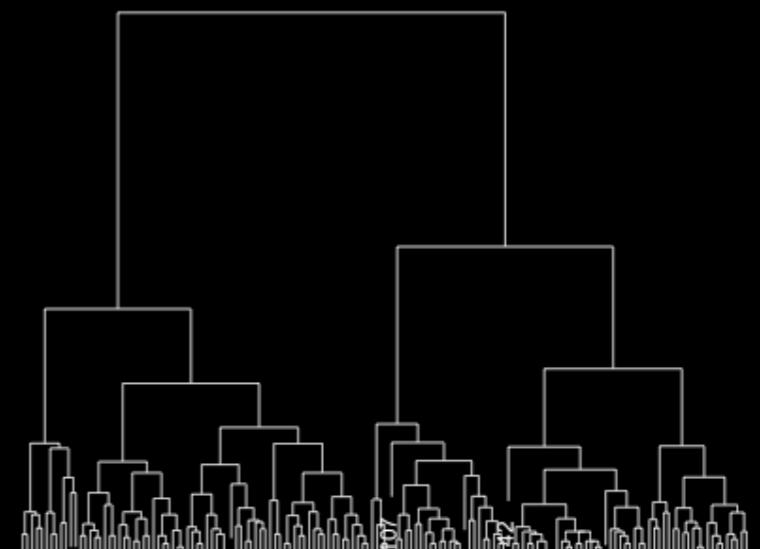
Spectral clustering  
Shi Malik 2000  
Von Luxburg 2007



(a) K-means

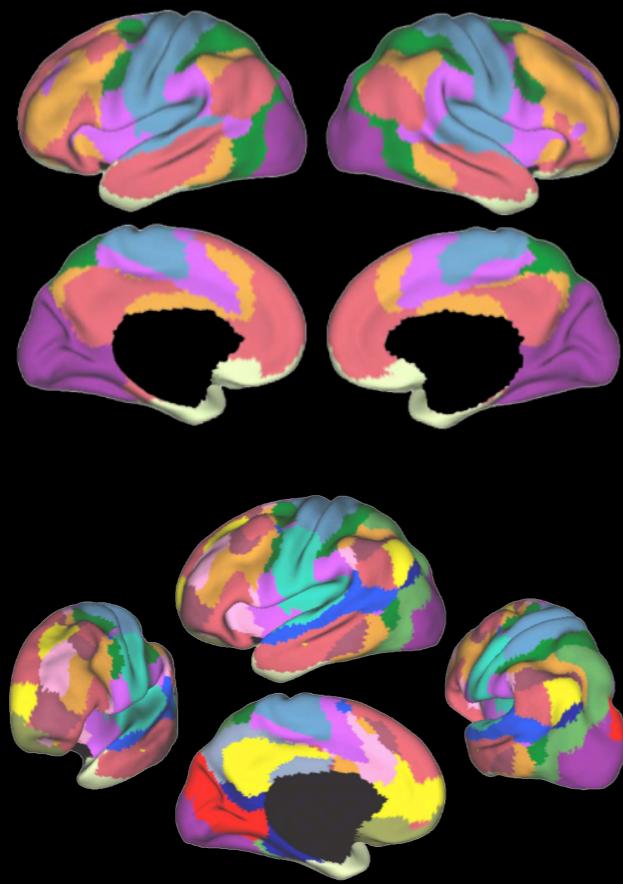
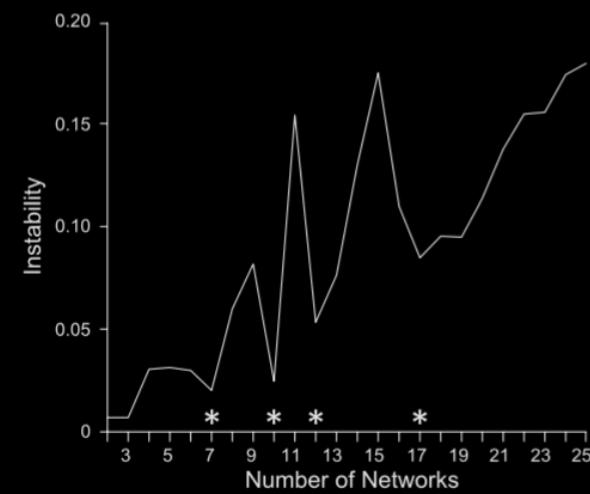


(b) Spectral Clustering

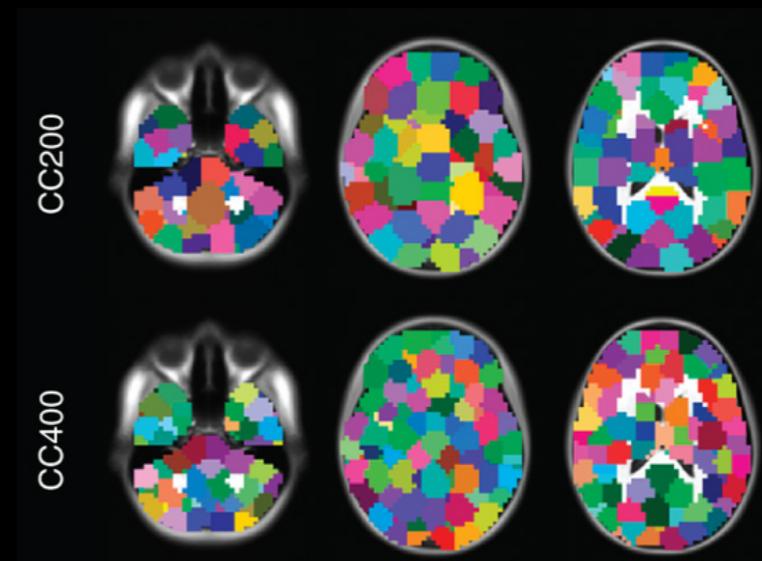


Hierarchical clustering

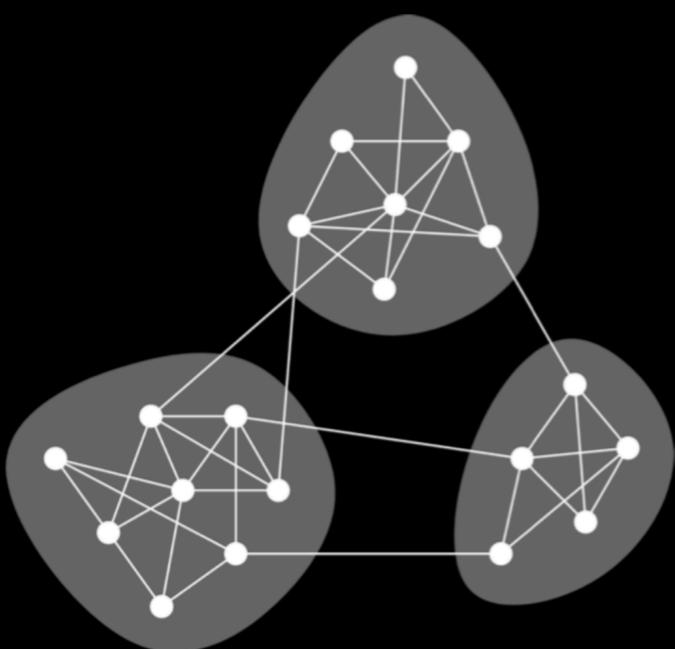
# NETWORK CLUSTERING APPLICATIONS



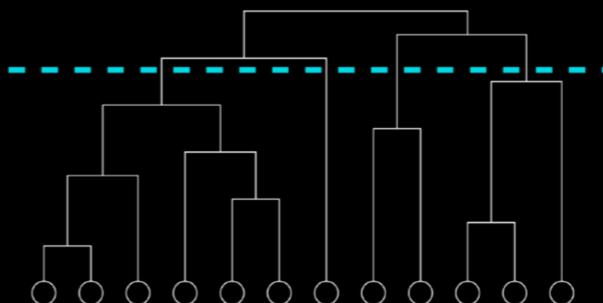
$$w_{ij} = \begin{cases} s(v_i, v_j) & d_{ij} \leq \varepsilon \\ 0 & d_{ij} > \varepsilon \end{cases}.$$



# MODULARITY DETECTION



**Fig. 1.** The vertices in many networks fall naturally into groups or communities, sets of vertices (shaded) within which there are many edges, with only a smaller number of edges between vertices of different groups.

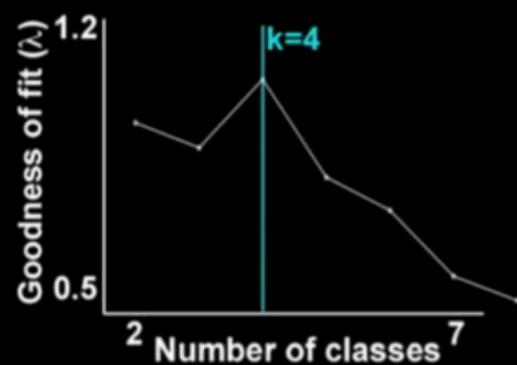


Girvan and Newman 2002, Newman 2006 PNAS, Blondel et al. 2008, Fortunato 2009

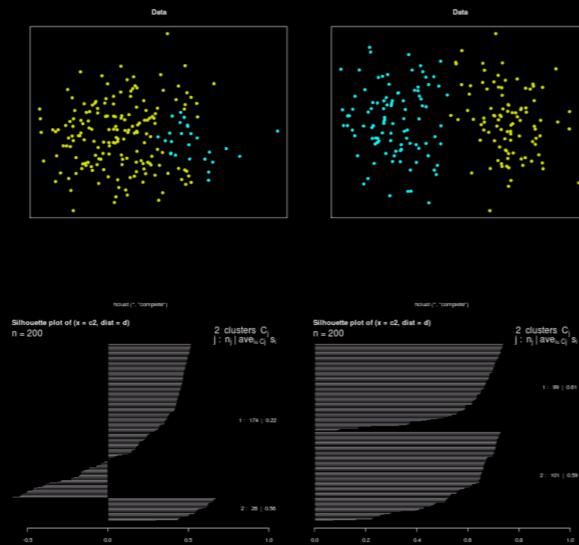
NEXT  
WEEK

# HOW MANY KS?

VARIANCE  
WITHIN/BETWEEN CLUSTERS



SILHOUETTE



MULTI-CRITERIA

**NbClust {NbClust}**  
NbClust Package for determining the best number of clusters

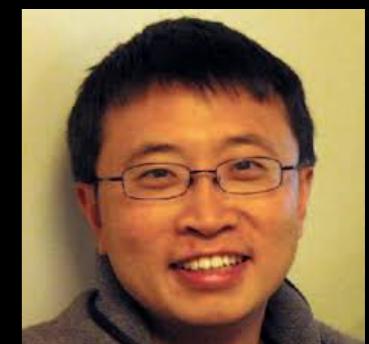
**Description**  
NbClust package provides 30 indices for determining the number of clusters and proposes to user the scheme from the different results obtained by varying all combinations of number of clusters, distance and clustering methods.

**Usage**  
`NbClust(data, diss = NULL, distance = "euclidean", min.nc = 2, max.nc = 15, method = "ward.D2", index = "all", alphaBeale = 0.1)`

**Arguments**

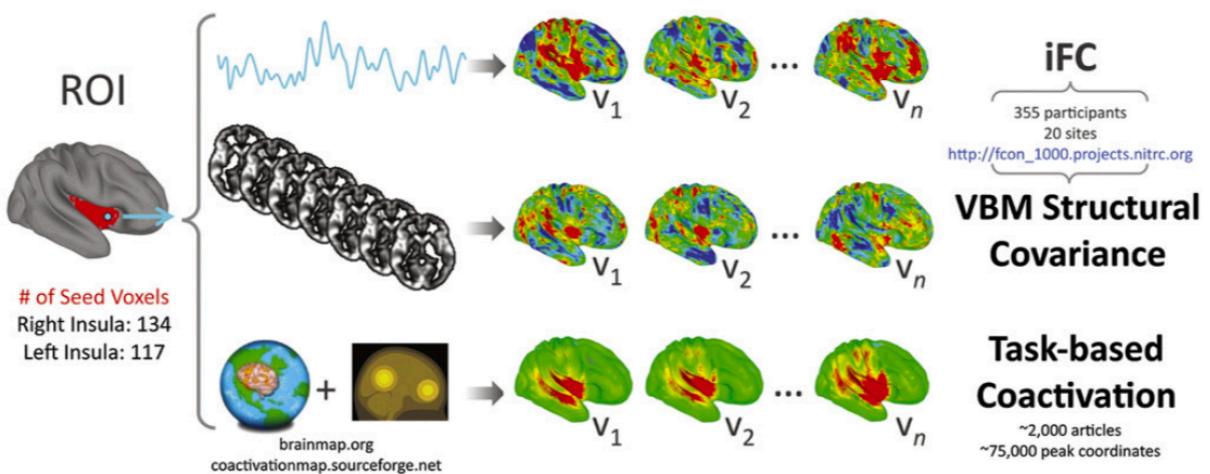
- data** matrix or dataset (the only mandatory argument)
- diss** dissimilarity matrix to be used. By default, `diss=NULL`, but if it is replaced by a dissimilarity matrix should be "NULL".
- distance** the distance measure to be used to compute the dissimilarity matrix. This must be one of "maximum", "manhattan", "canberra", "binary", "minkowski" or "NULL". By default, `d` If the distance is "NULL", the dissimilarity matrix (`diss`) should be given by the user. If "NULL", the dissimilarity matrix should be "NULL".
- min.nc** minimal number of clusters, between 1 and (number of objects - 1)
- max.nc** maximal number of clusters, between 2 and (number of objects - 1), optional

STABILITY  
REPLICABILITY

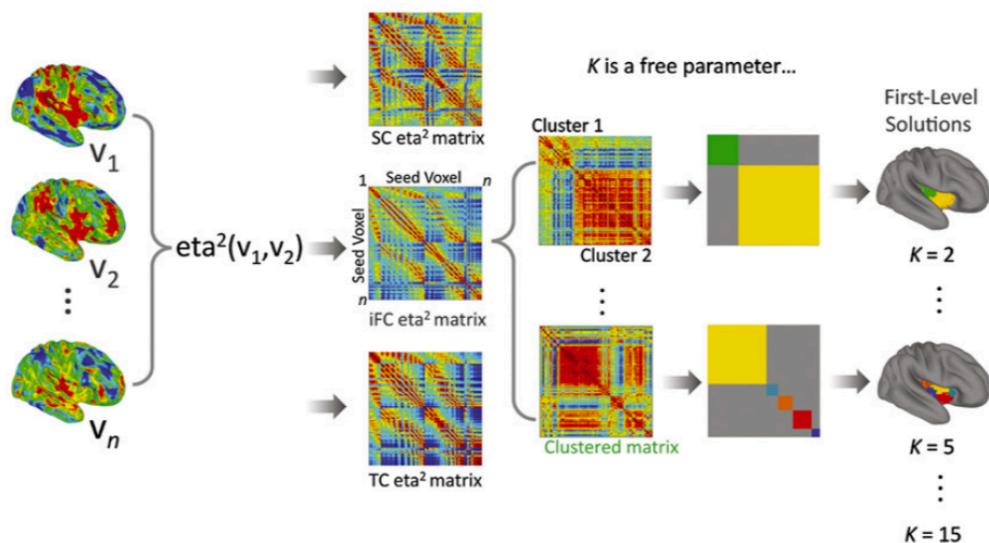


KELLY ET AL

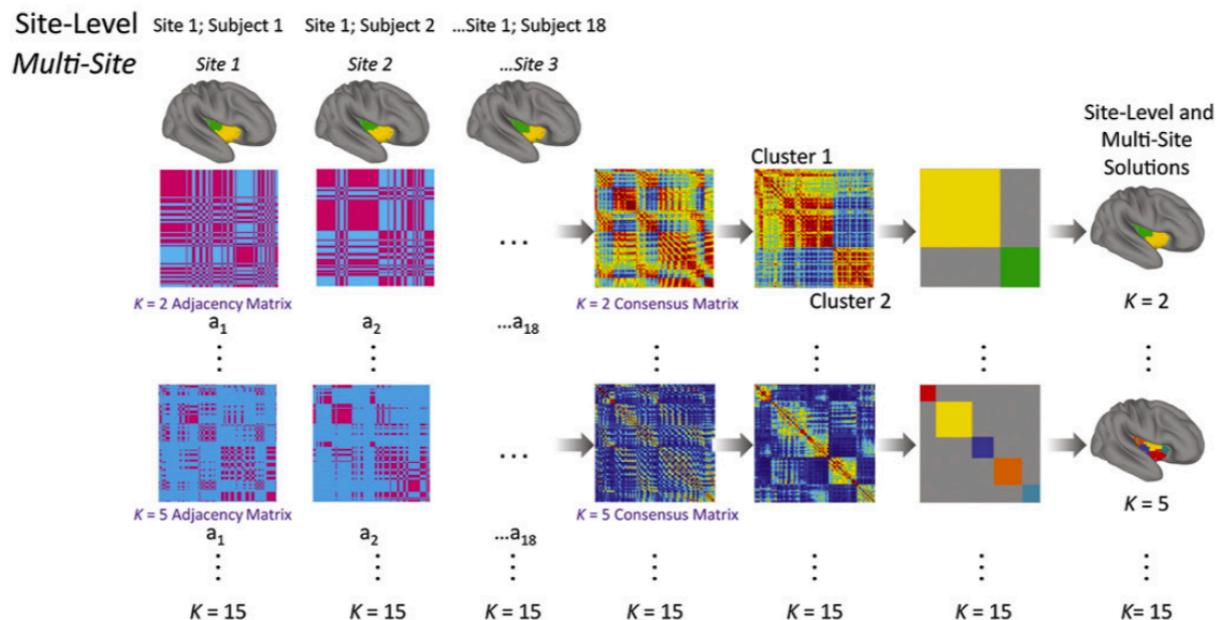
## Step 1: Covariance-Based Measures

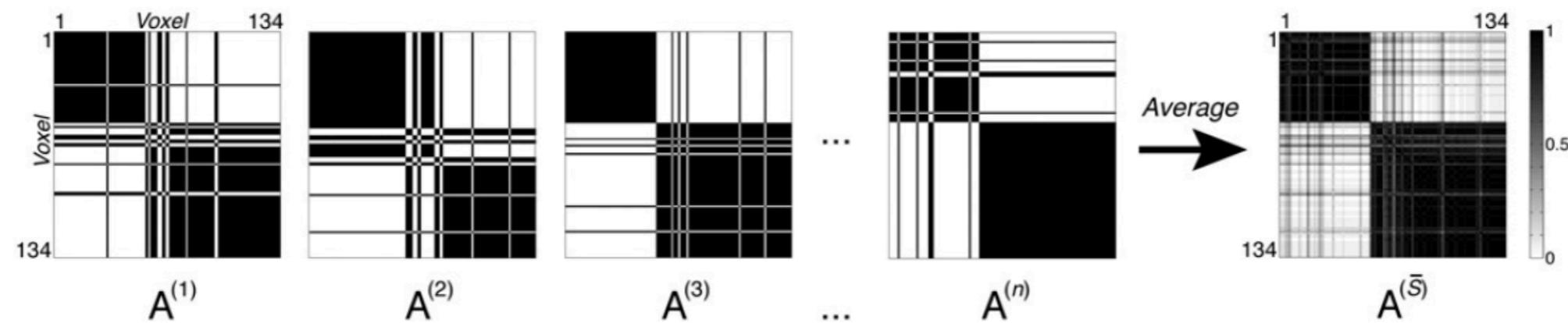


## Step 2: $\eta^2$ and First-Level Clustering

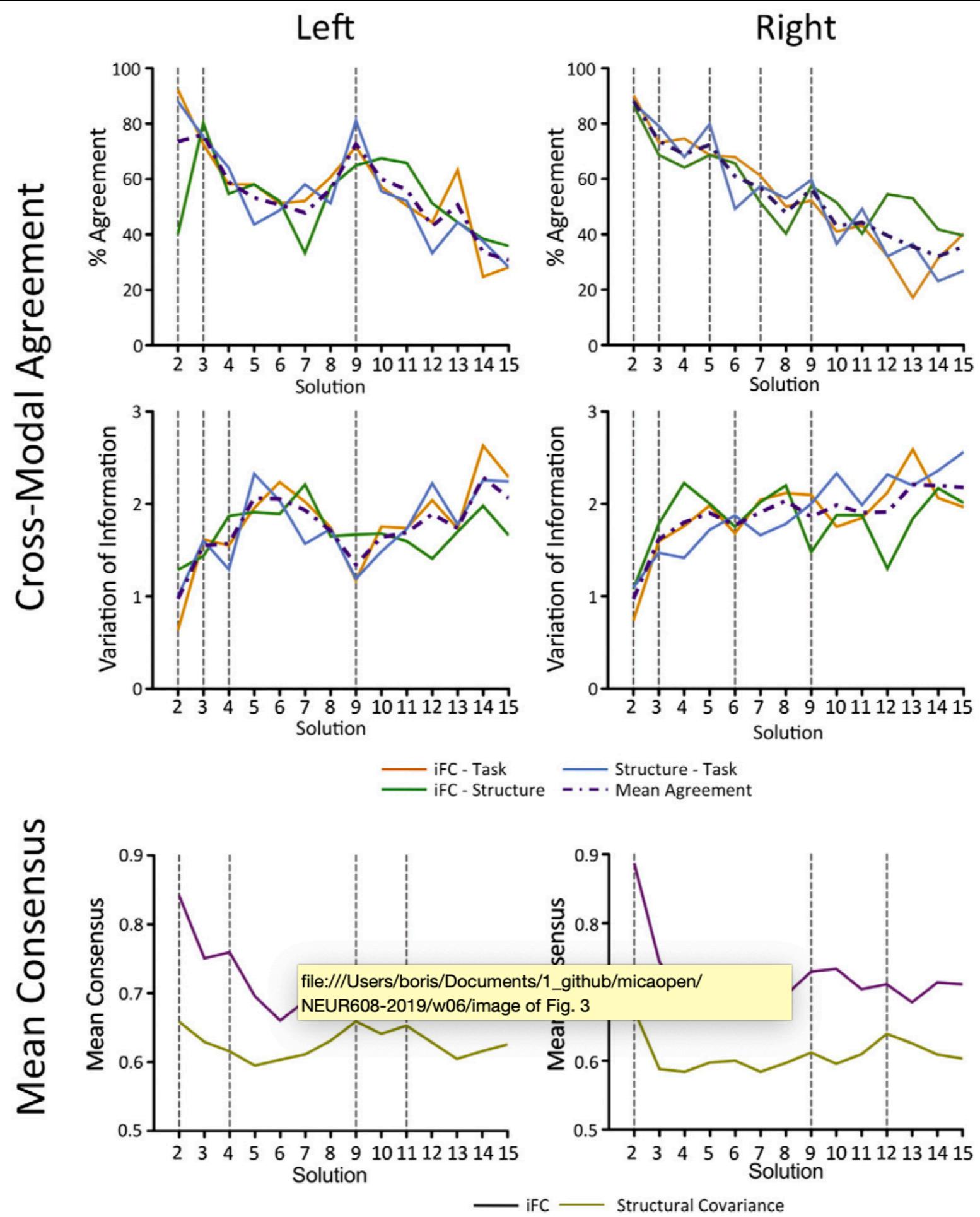


## Step 3: Consensus (Site-Level and Multi-Site) Clustering



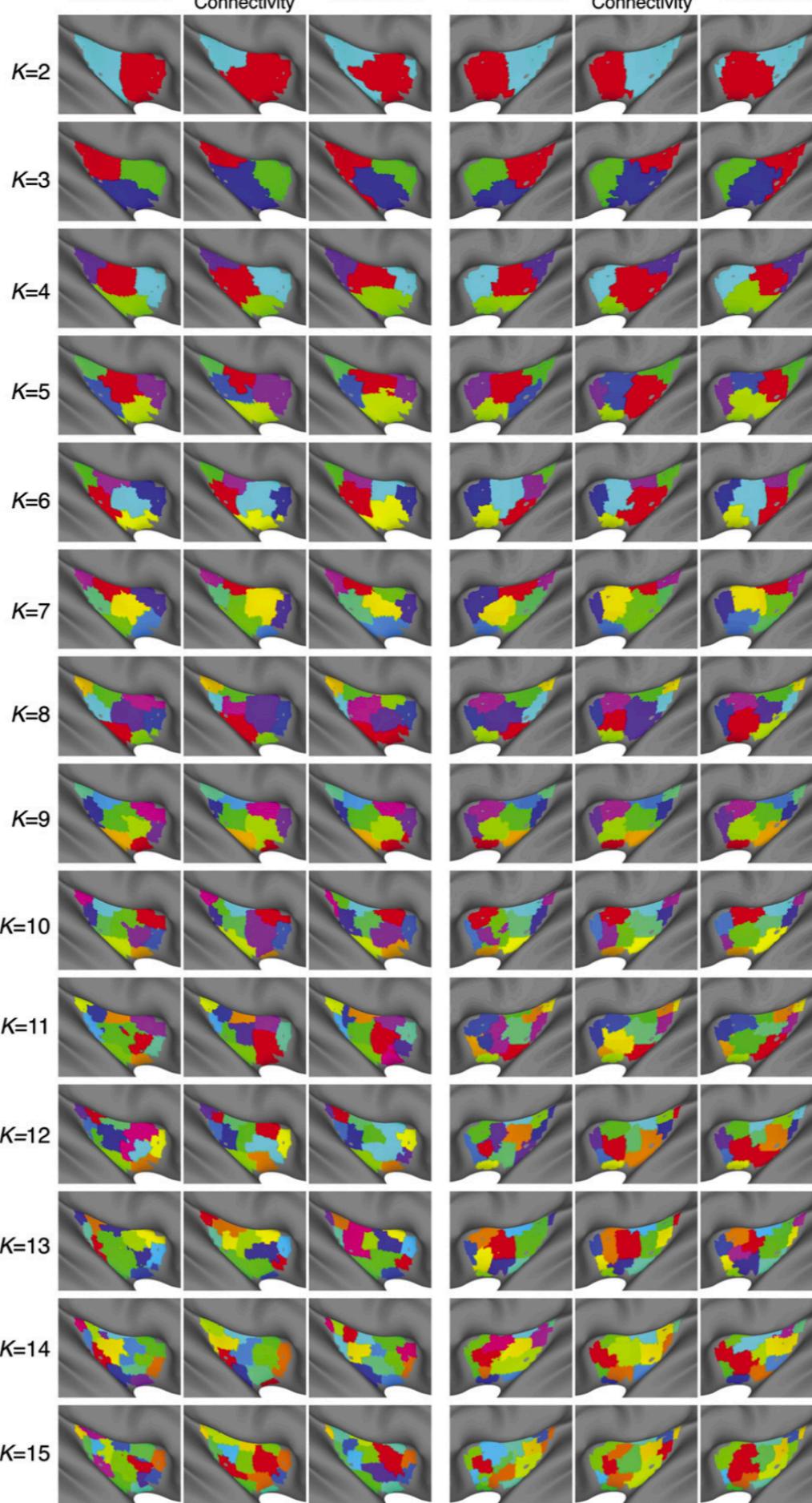


**Fig. 2.** Consensus clustering schematic for  $K=2$ . The schematic illustrates the consensus clustering process. For each scale  $K$ , each clustering instance contributes an adjacency matrix  $A^{(s)}$ , each element  $a_{ij}^{(s)}$  of which contains a value of 1 if voxels  $i$  and  $j$  are assigned to the same cluster  $k$ , and 0 otherwise. In this example, let each instance be a data collection site, so  $A^{(1)}$  is contributed by Bangor;  $A^{(2)}$  is contributed by Berlin, etc. A consensus matrix  $A^{(\bar{S})}$  is derived by averaging across adjacency matrices. Each element of the consensus matrix thus contains a number between 0 and 1, corresponding to the proportion of times a given pair of voxels appeared in the same cluster, across instances (here, data collection sites). The spectral clustering algorithm can then be applied to identify the most stable pattern of cluster assignments across instances, using the same scale  $K$  that was used to generate the consensus matrix (here,  $K=2$ ).



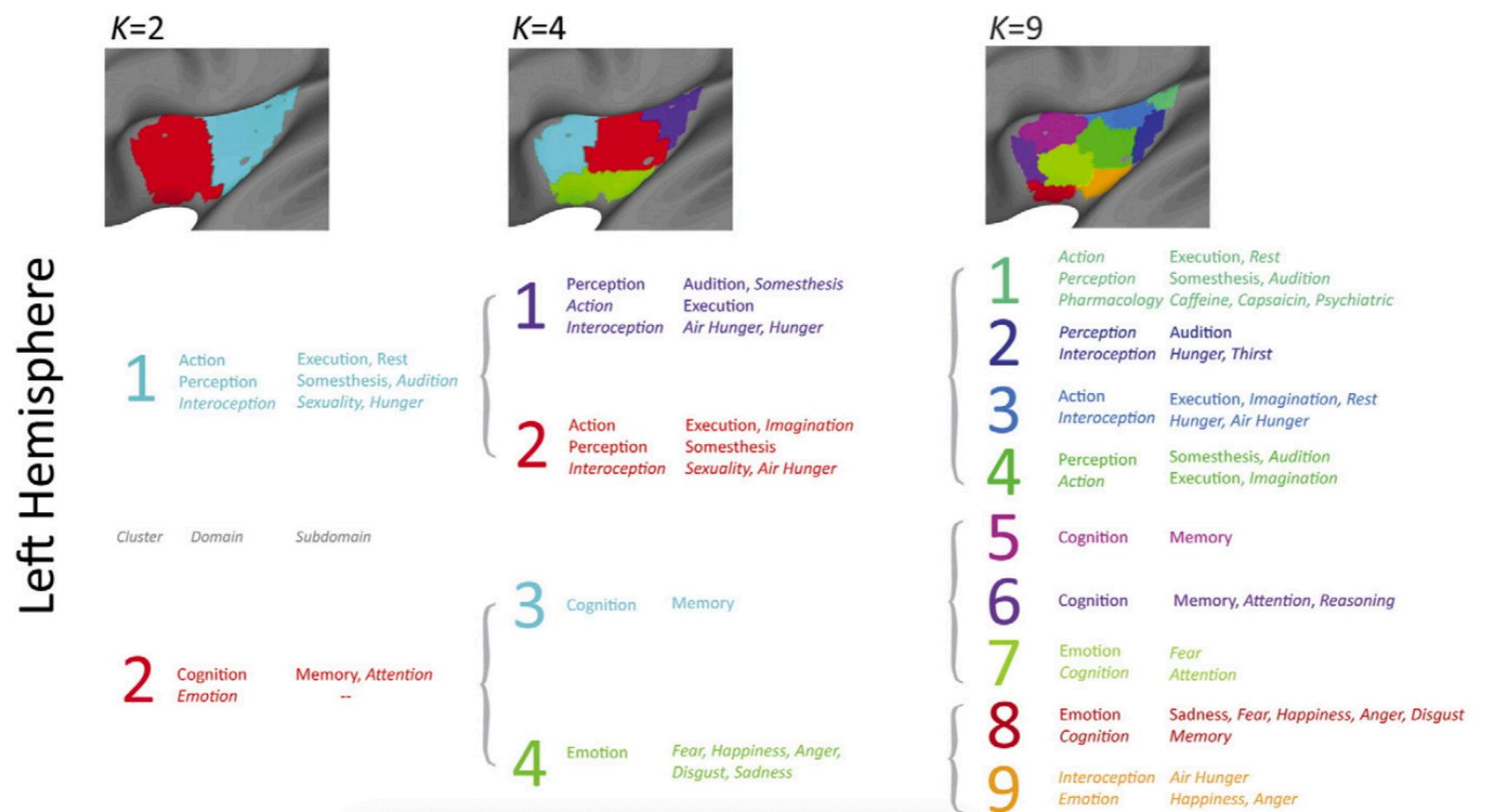
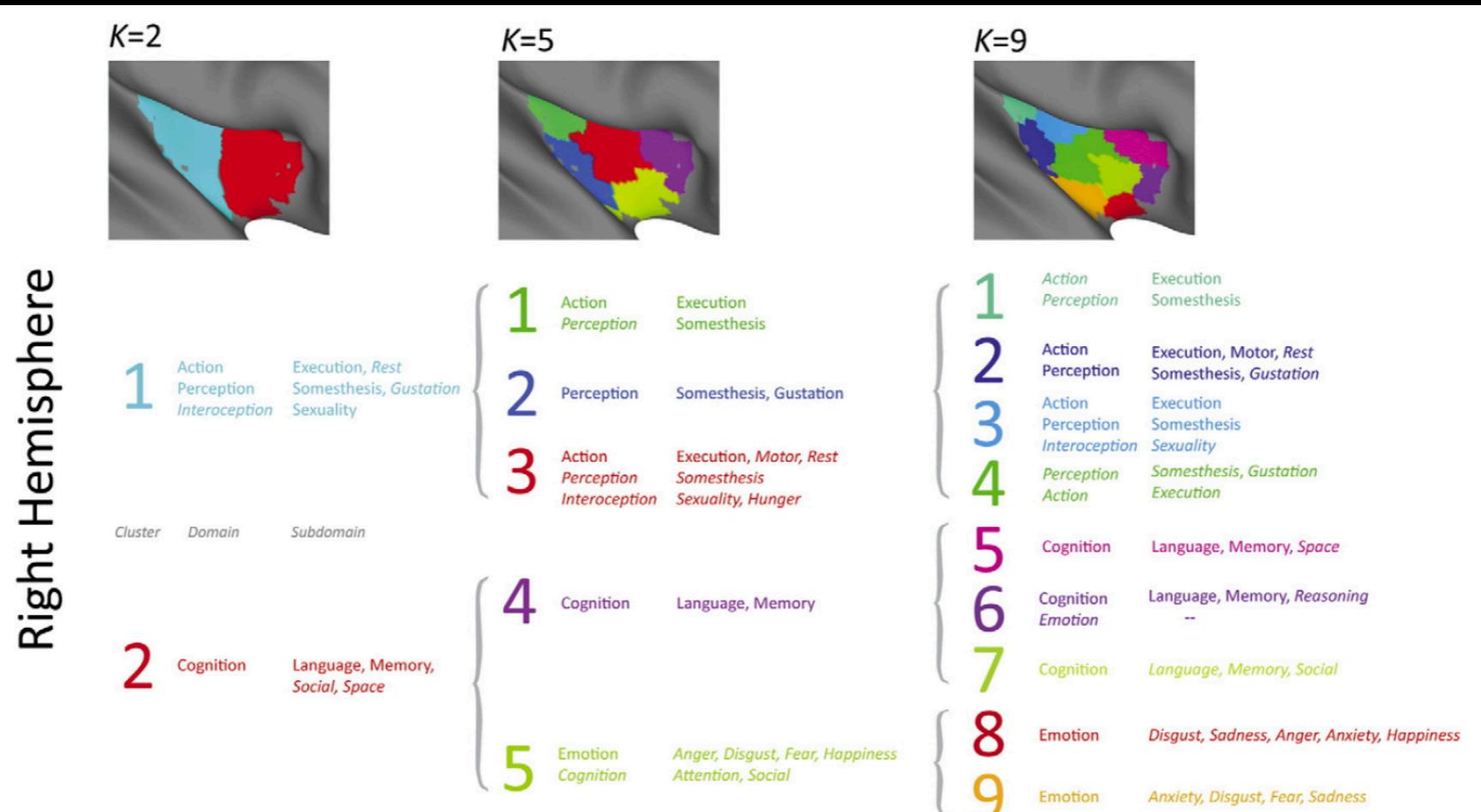
Right Hemisphere

Task Coactivation	Intrinsic Functional Connectivity	Structural Covariance
-------------------	-----------------------------------	-----------------------

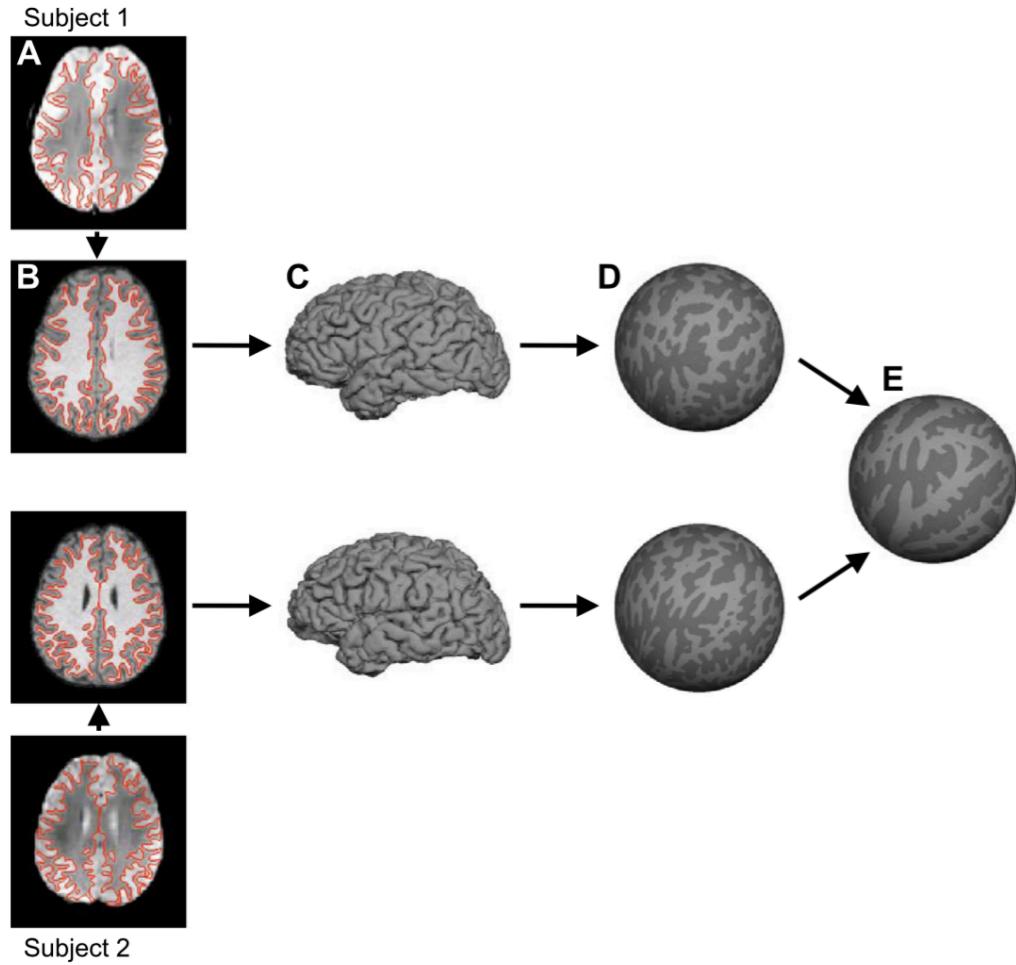


Left Hemisphere

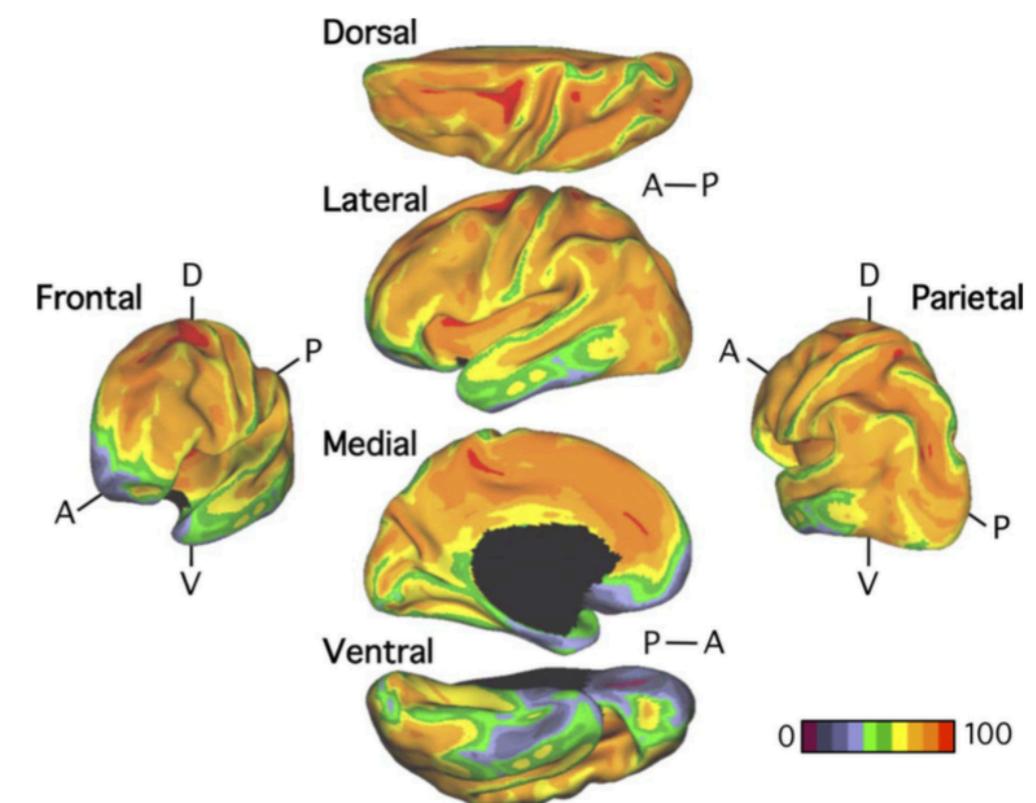
Task Coactivation	Intrinsic Functional Connectivity	Structural Covariance
-------------------	-----------------------------------	-----------------------



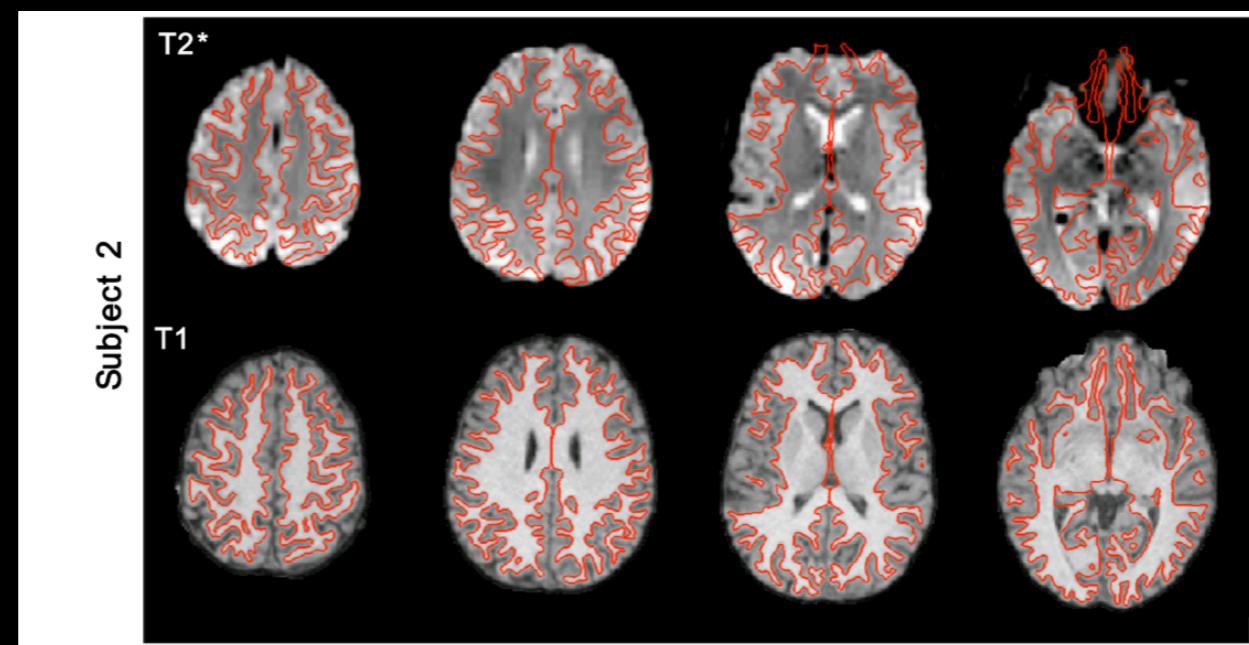
YEO, KRIENEN ET AL

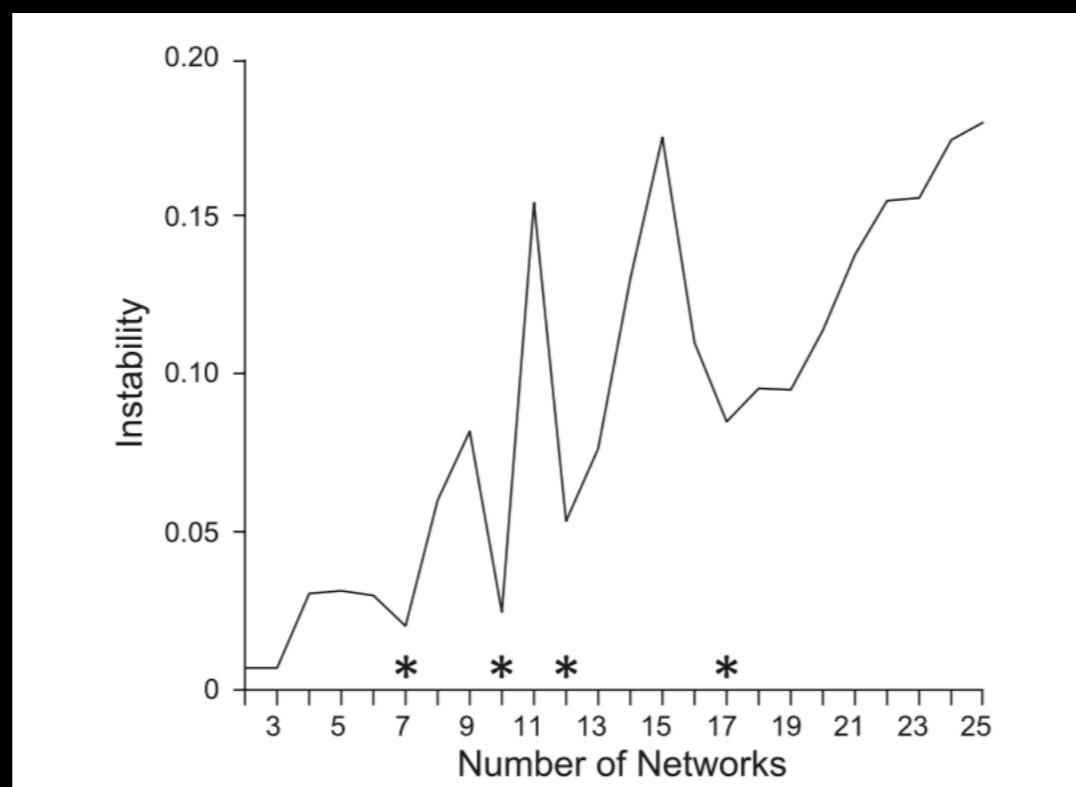
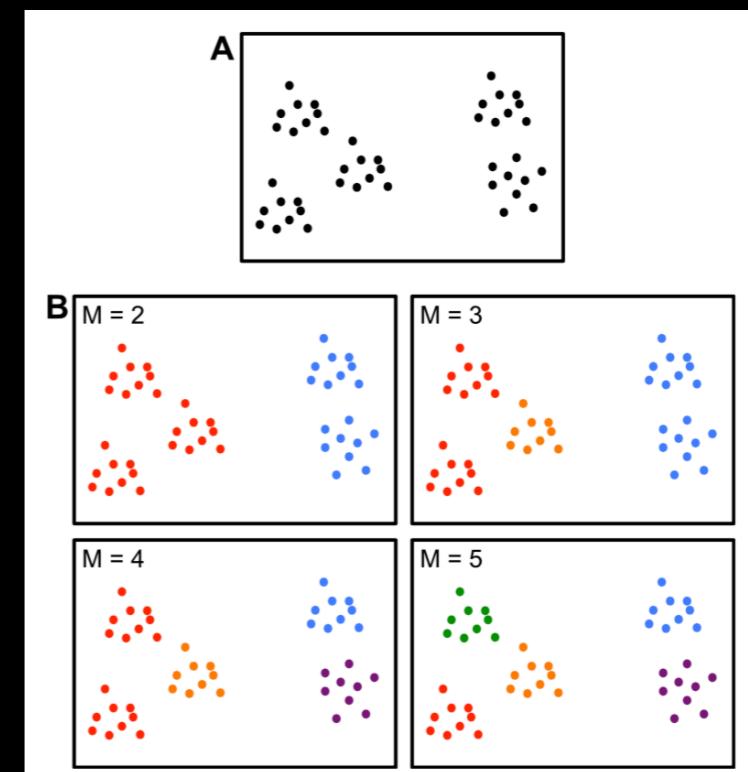
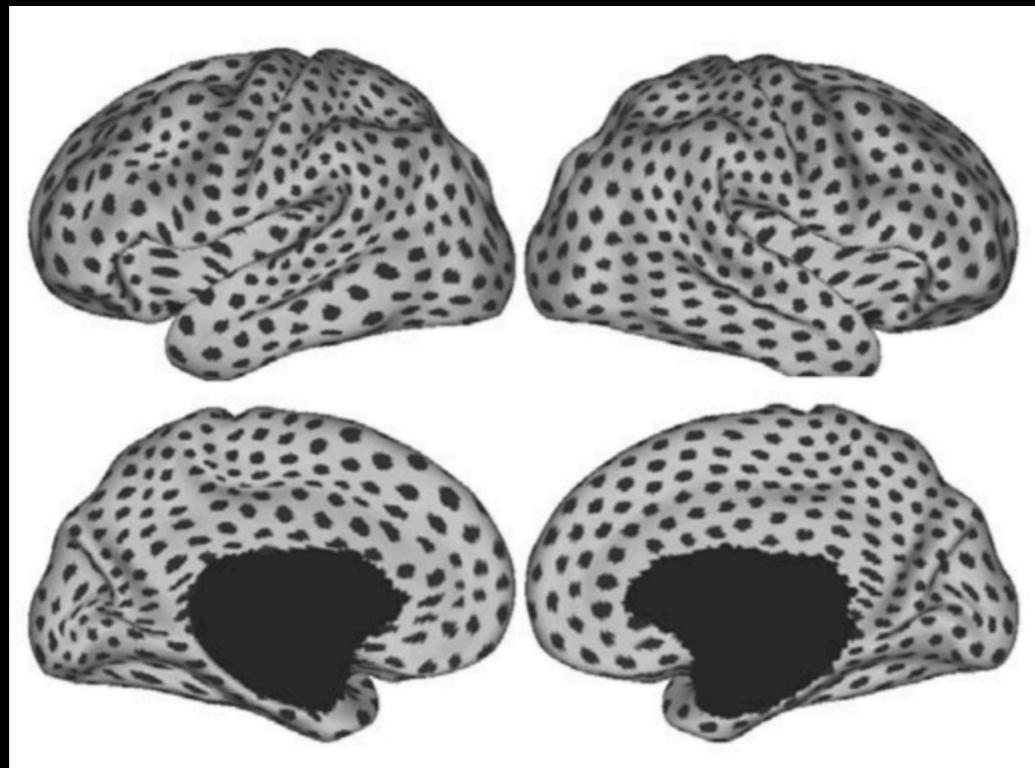


**Fig. 1.** Surface coordinate system for functional magnetic resonance imaging (fMRI) analysis. For each subject, the  $T2^*$  images yielding blood oxygenation level-dependent (BOLD) contrast fMRI data (*A*) were registered to the  $T1$ -weighted structural data (*B*). The cortical gray-white and pial surfaces were estimated from the structural data. The red lines show the estimated gray-white surface (*A* and *B*). Pial surface is shown in *C*. The gray-white surface was inflated into a sphere (*D*). The inflated spheres were then aligned across subjects using surface-based registration of the cortical folding pattern, resulting in a common spherical coordinate system (*E*). BOLD data of individual subjects (*A*) can then be projected onto the spherical coordinate system (*E*) in a single transformation step to reduce artifacts due to multiple interpolations.



**Fig. 3.** Signal-to-noise ratio (SNR) maps of the functional data from the full sample ( $N = 1,000$ ). The mean estimate of the BOLD fMRI data SNR is illustrated for multiple views of the left hemisphere in Caret PALS space. A, anterior; P, posterior; D, dorsal; V, ventral.





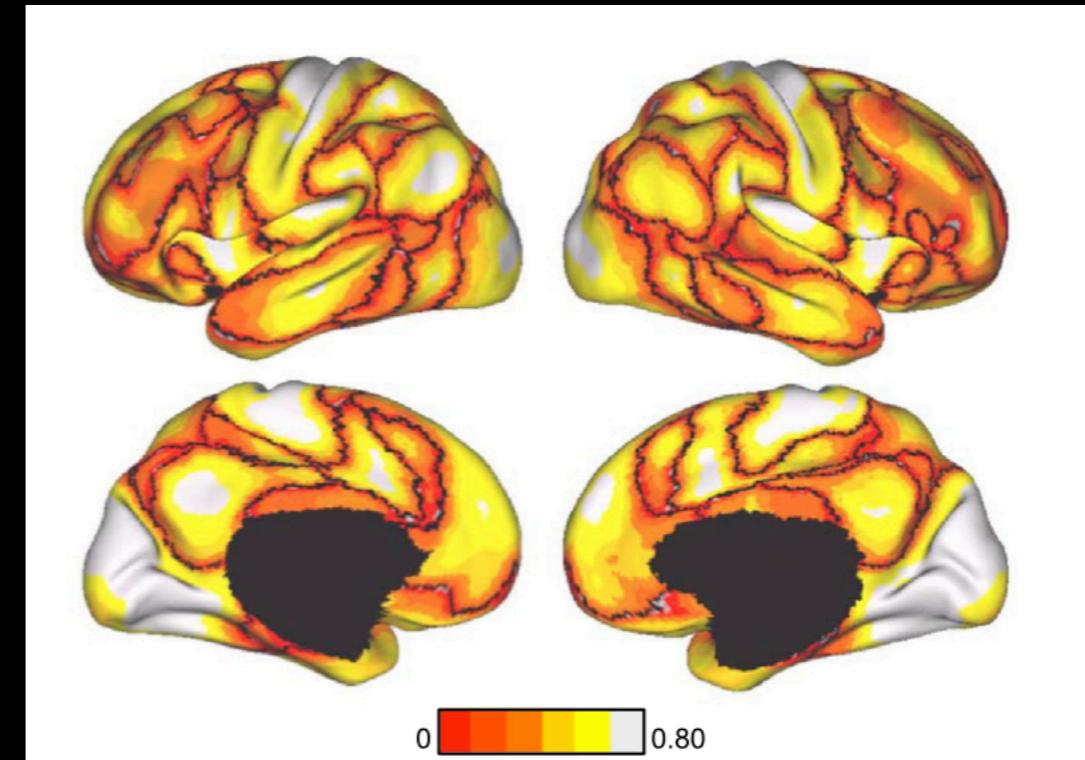
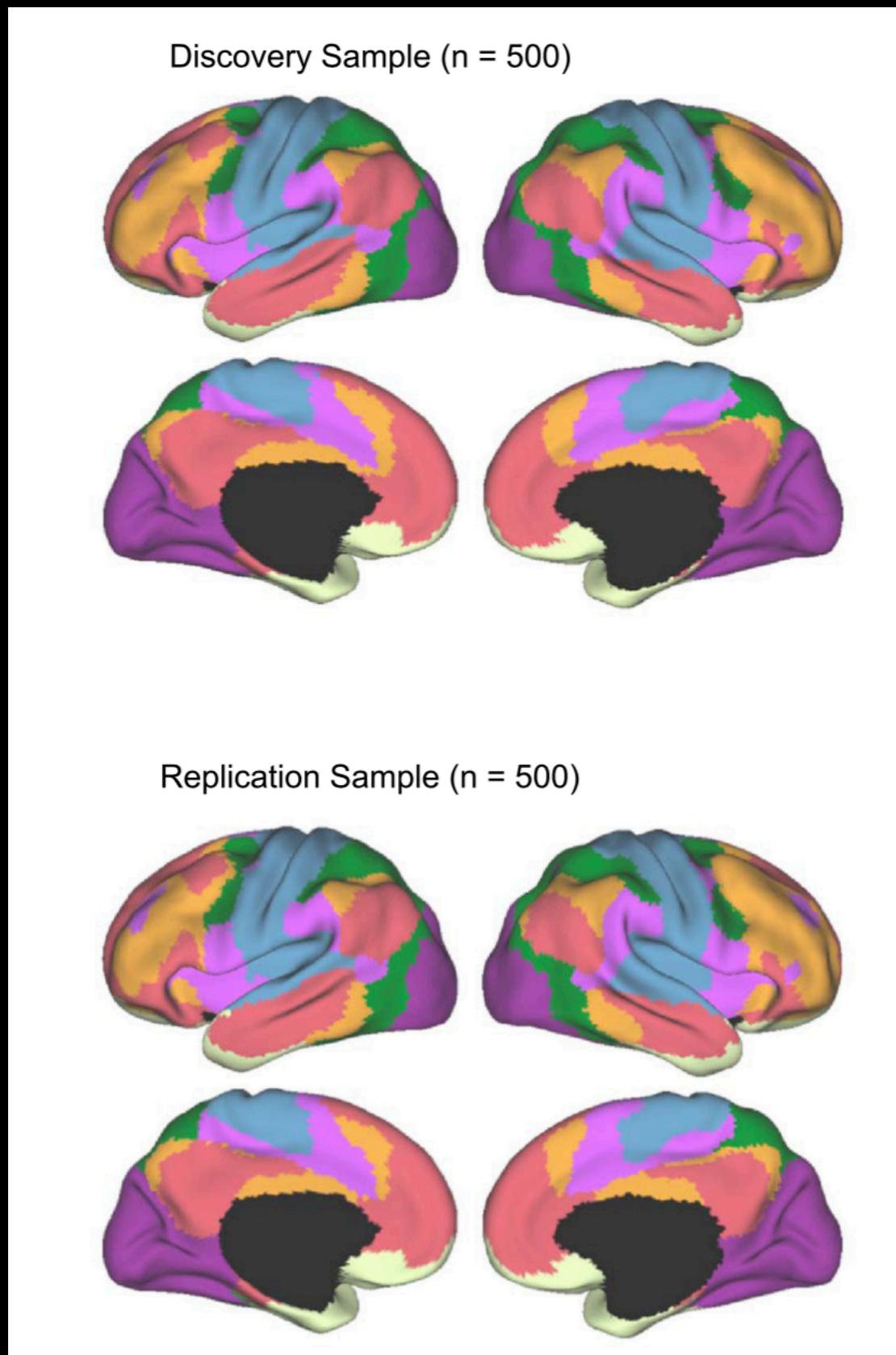
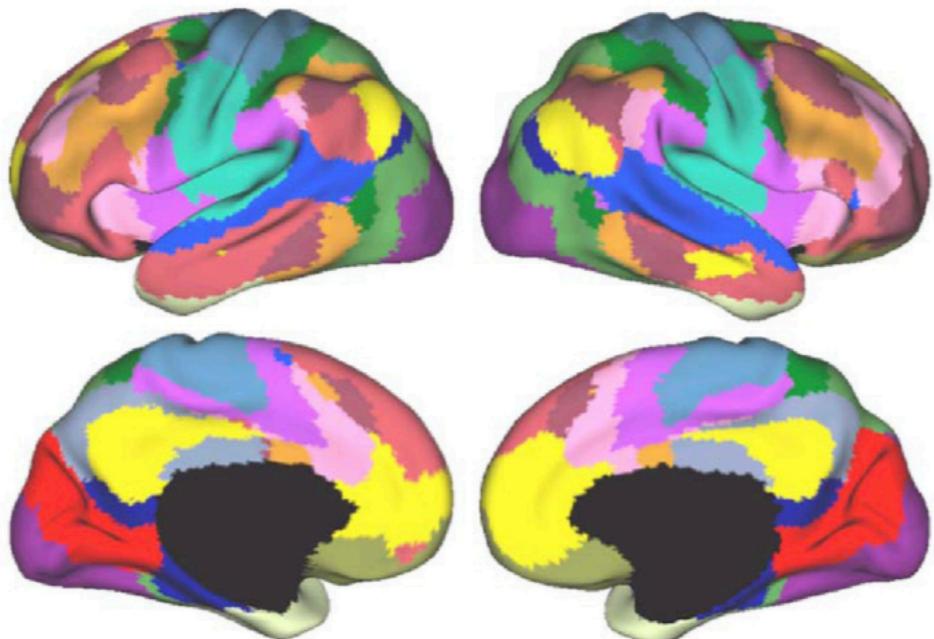


Fig. 8. Confidence of the 7-network estimate in the discovery data set. Confidence (silhouette) value for each vertex with respect to its assigned network is shown for the discovery data set. Regions close to the boundaries between networks were less confident of their assignment, although we also observed structured spatial variation within individual components of the estimated networks, such as lateral prefrontal cortex, which foreshadows its division in the 17-network estimate (see Fig. 9).

Discovery Sample (n = 500)



Replication Sample (n = 500)

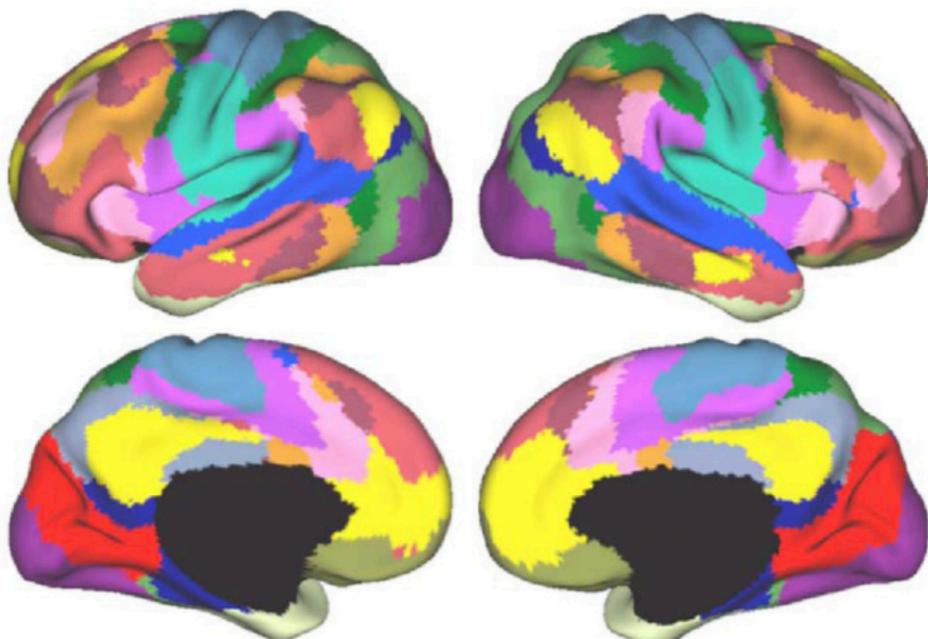
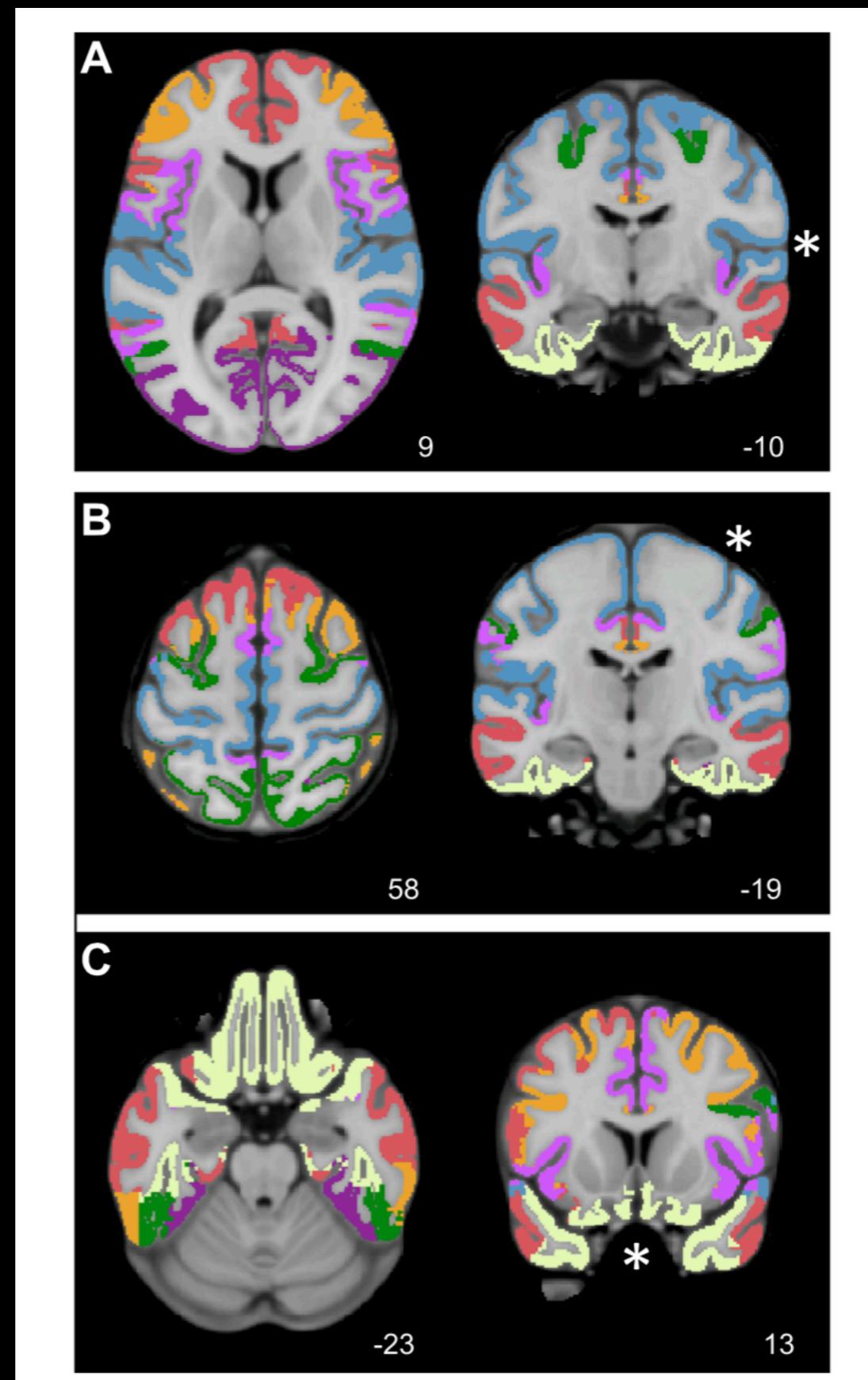
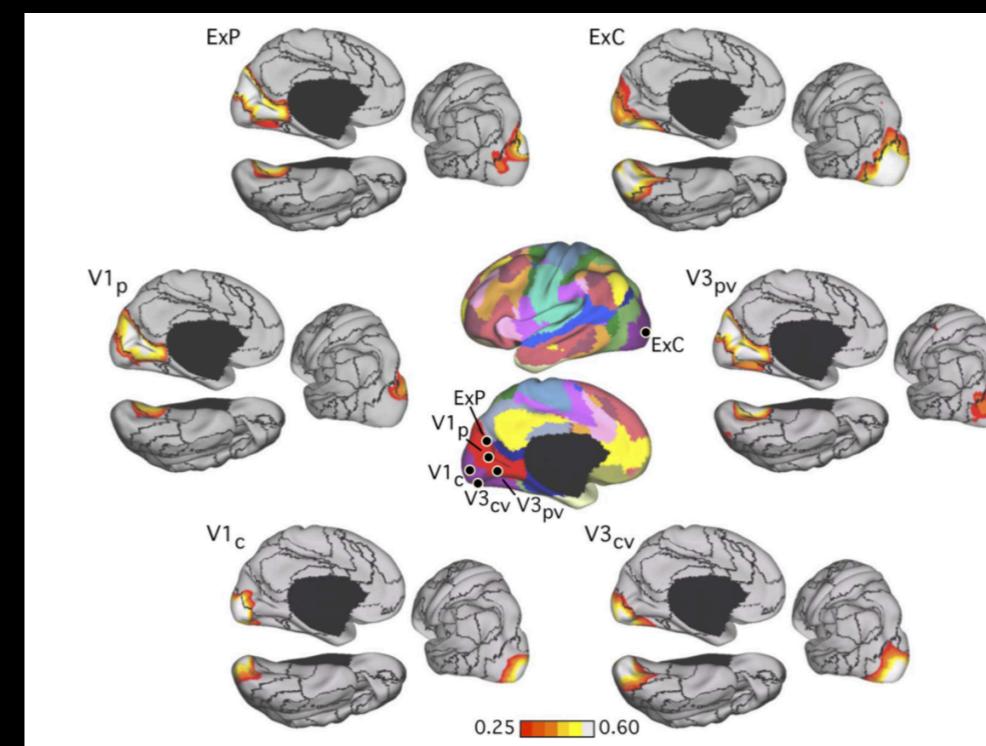
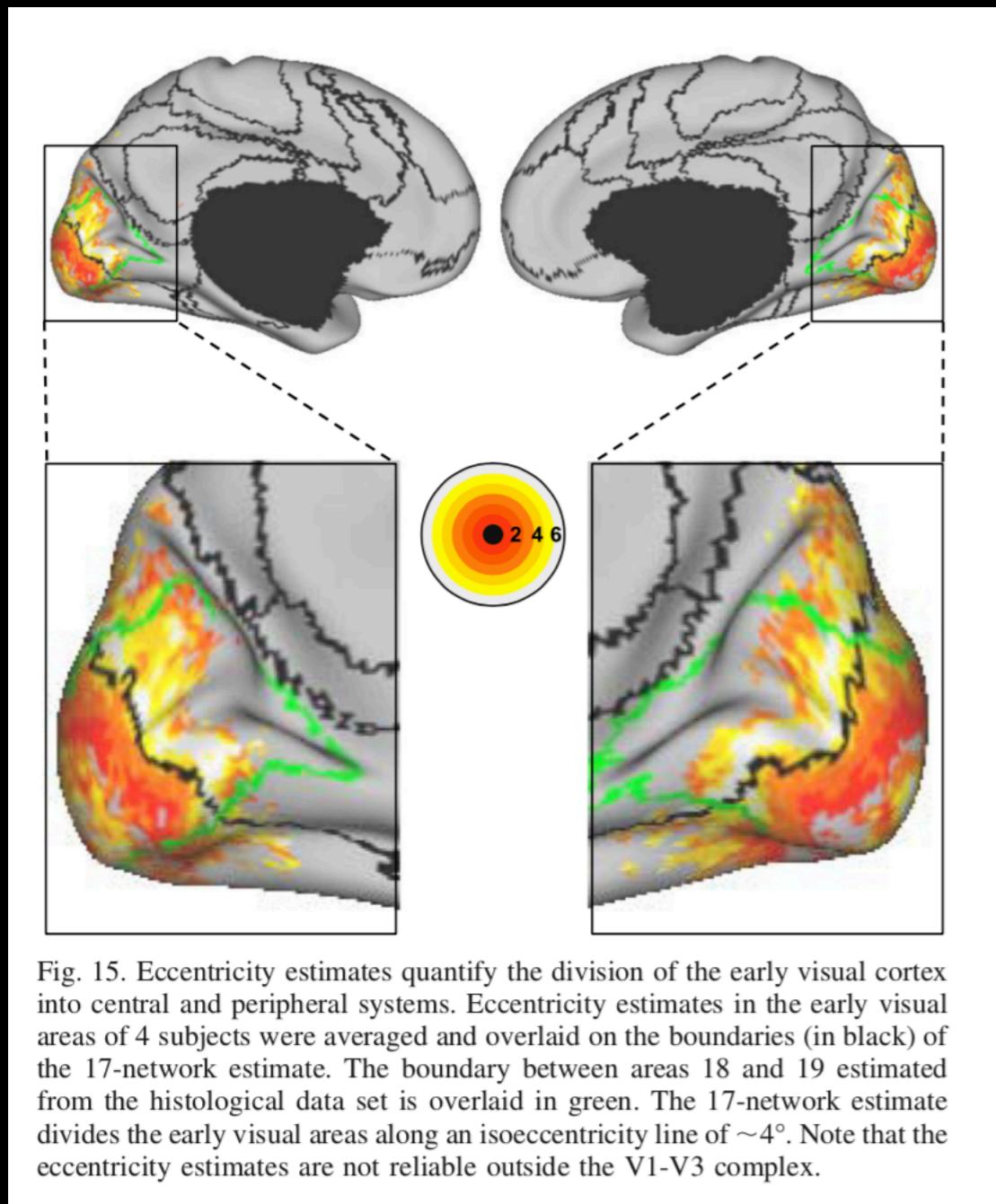


Fig. 10. Confidence of 17-network estimate in the discovery data set. Confidence (silhouette) value for each vertex with respect to its assigned network is shown for the discovery data set. Again, regions close to the boundaries between networks were less confident of their assignment.





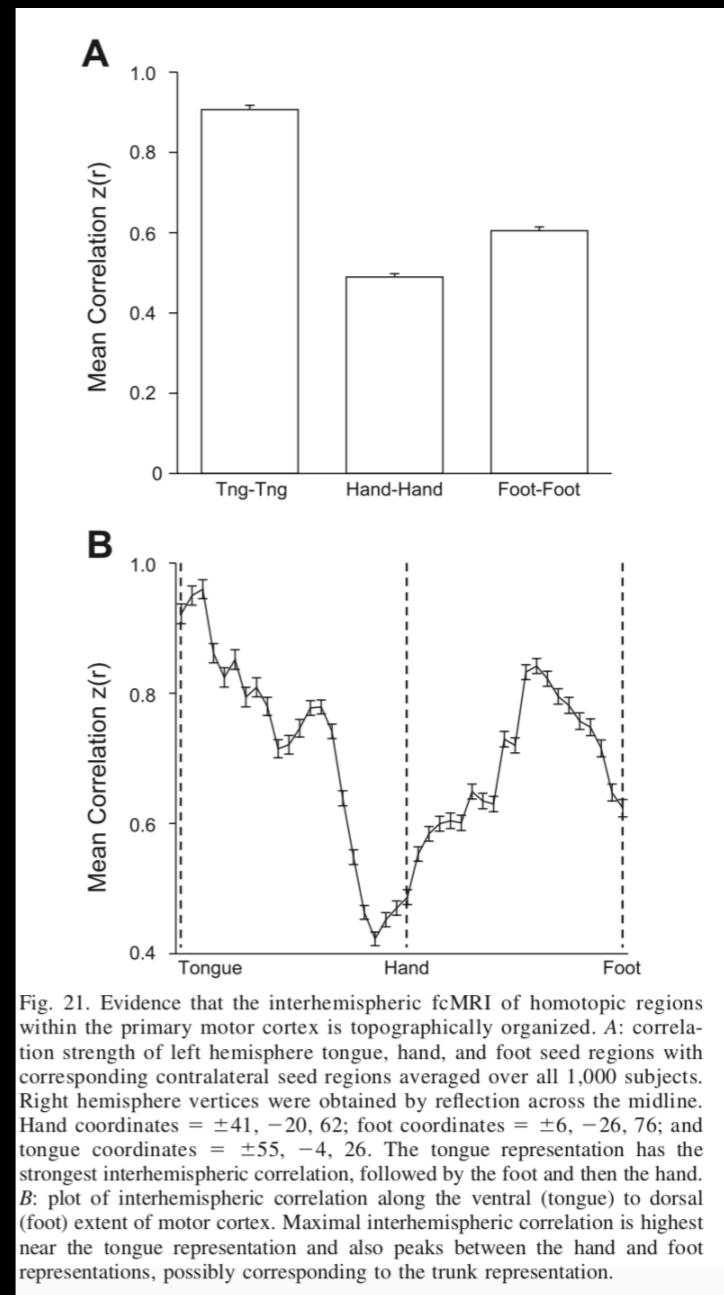
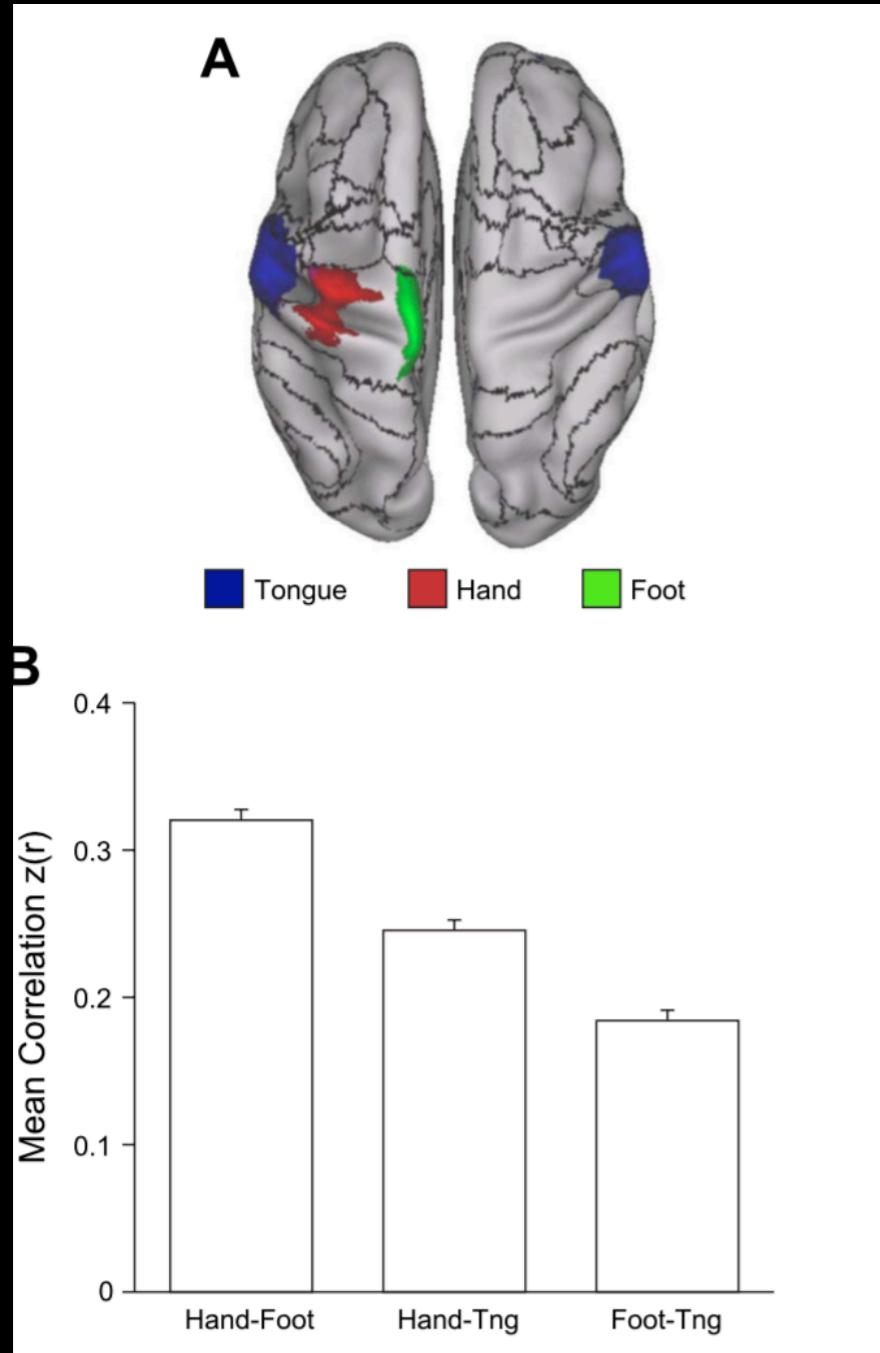


Fig. 21. Evidence that the interhemispheric fcMRI of homotopic regions within the primary motor cortex is topographically organized. **A:** correlation strength of left hemisphere tongue, hand, and foot seed regions with corresponding contralateral seed regions averaged over all 1,000 subjects. Right hemisphere vertices were obtained by reflection across the midline. Hand coordinates =  $\pm 41, -20, 62$ ; foot coordinates =  $\pm 6, -26, 76$ ; and tongue coordinates =  $\pm 55, -4, 26$ . The tongue representation has the strongest interhemispheric correlation, followed by the foot and then the hand. **B:** plot of interhemispheric correlation along the ventral (tongue) to dorsal (foot) extent of motor cortex. Maximal interhemispheric correlation is highest near the tongue representation and also peaks between the hand and foot representations, possibly corresponding to the trunk representation.

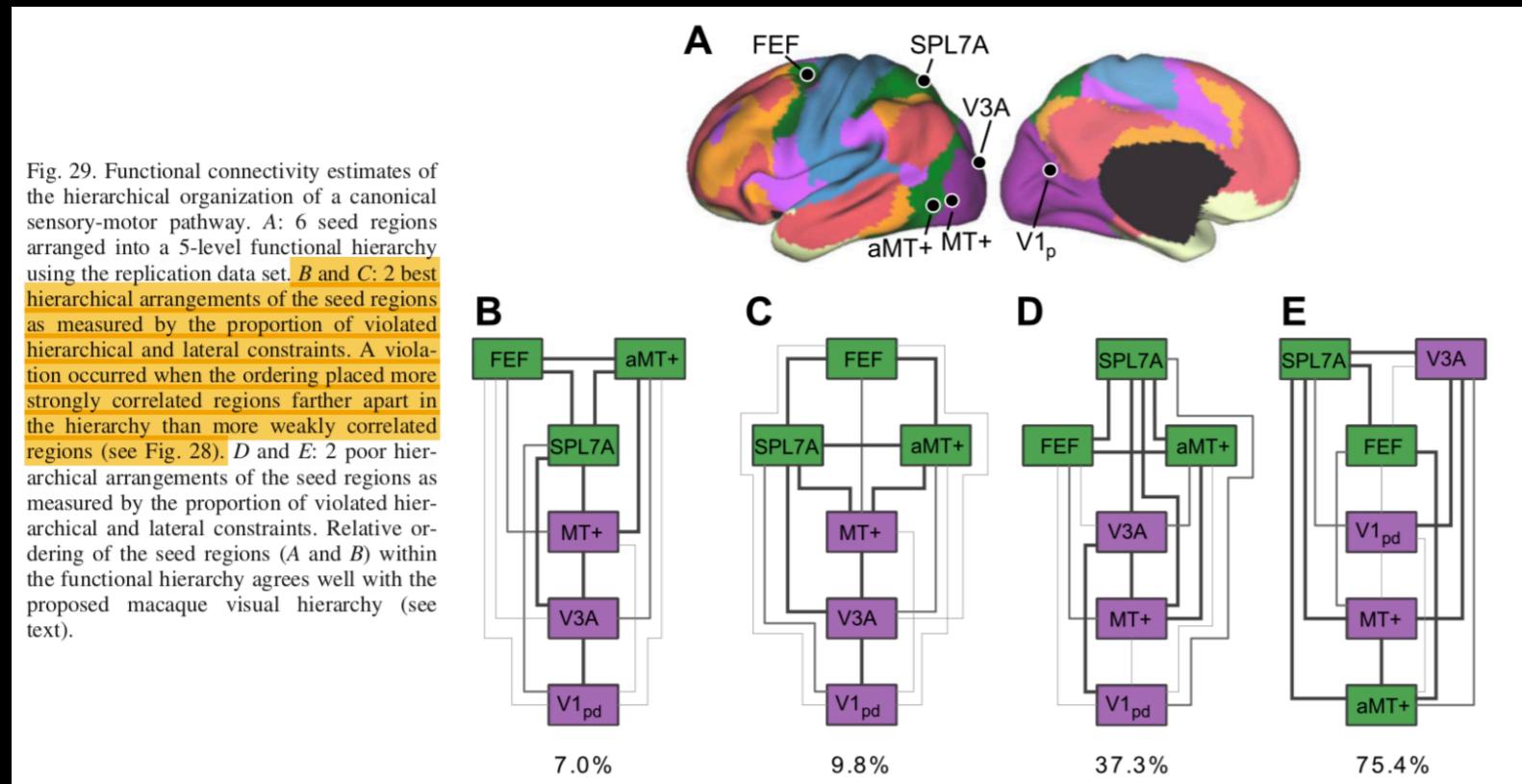
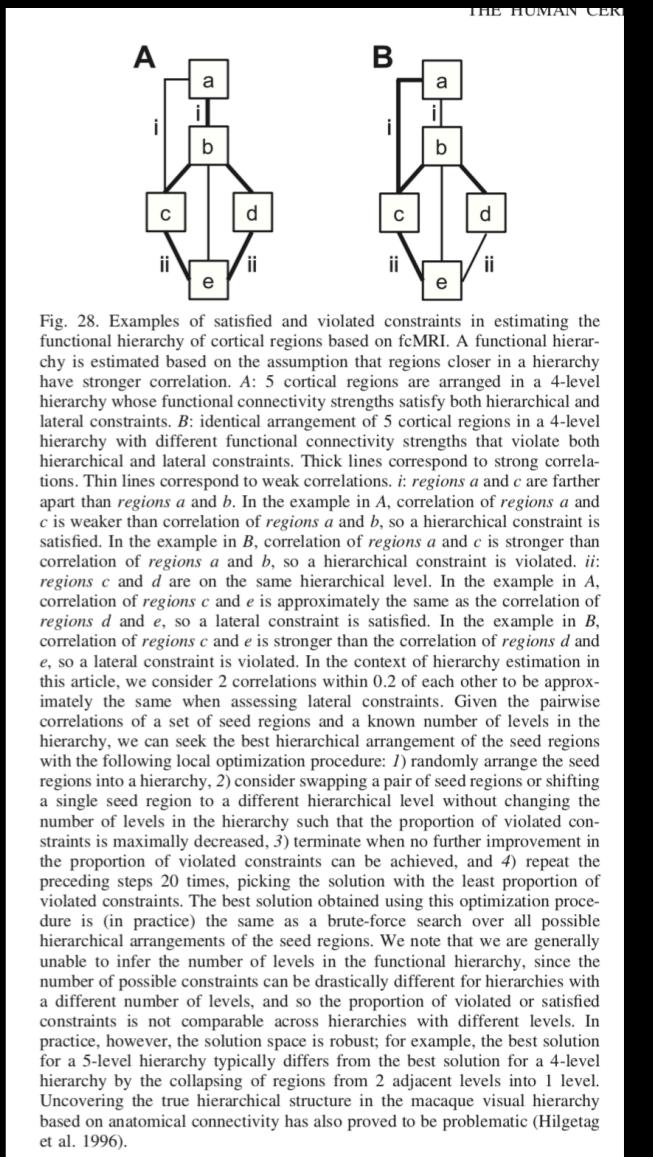
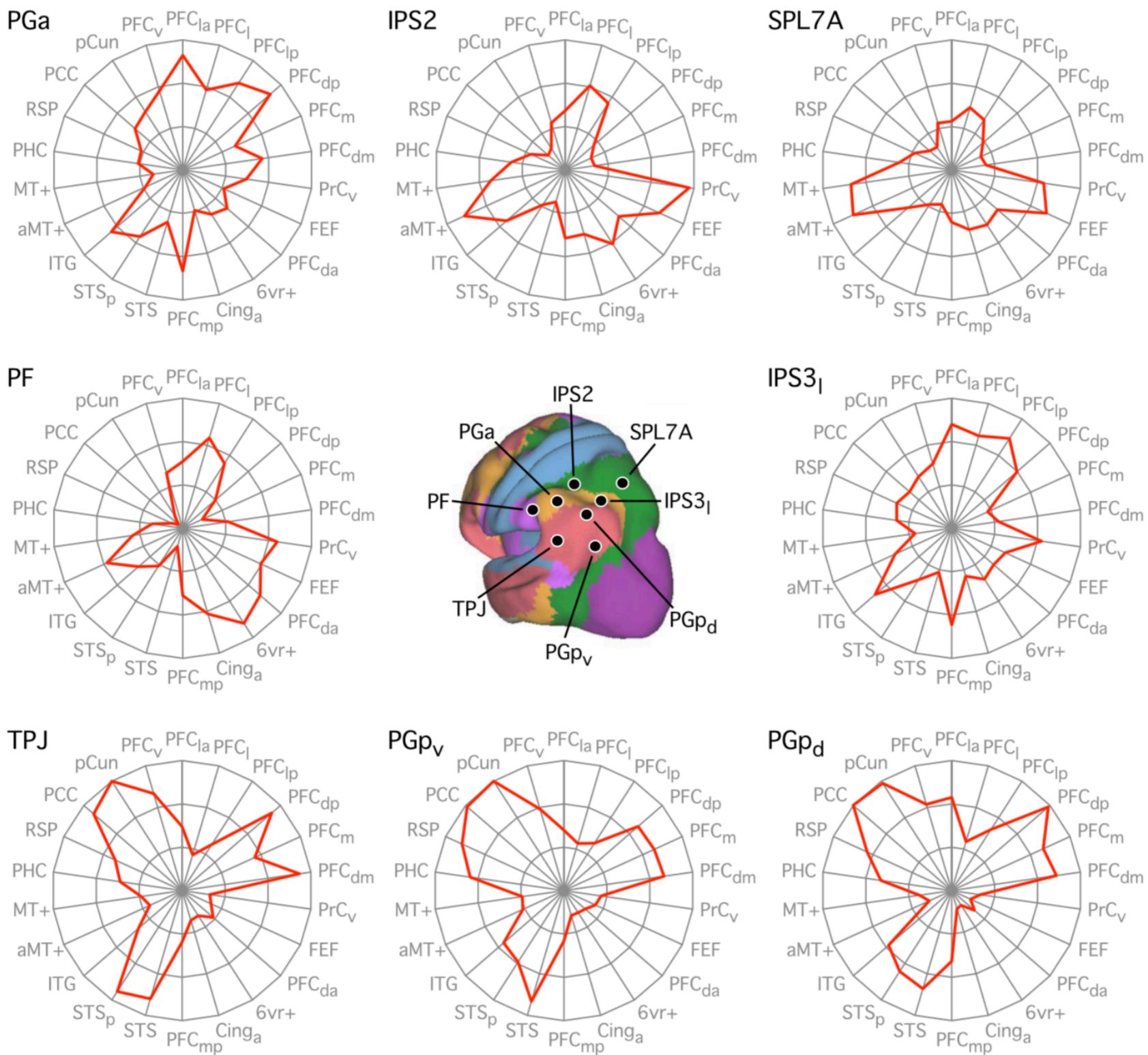
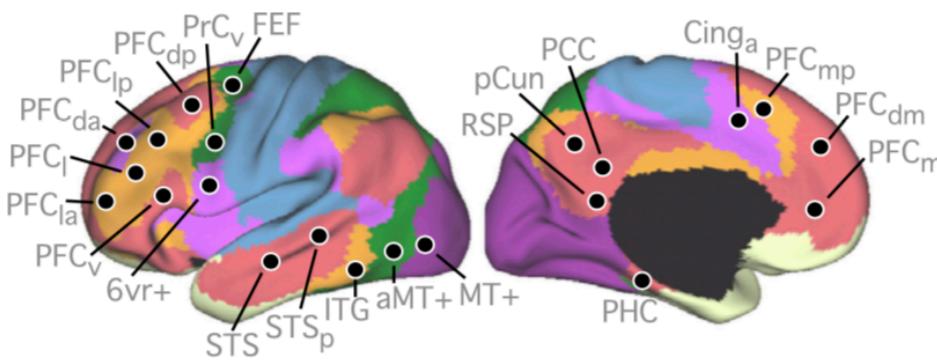


Fig. 29. Functional connectivity estimates of the hierarchical organization of a canonical sensory-motor pathway. **A:** 6 seed regions arranged into a 5-level functional hierarchy using the replication data set. **B** and **C:** 2 best hierarchical arrangements of the seed regions as measured by the proportion of violated hierarchical and lateral constraints. A violation occurred when the ordering placed more strongly correlated regions farther apart in the hierarchy than more weakly correlated regions (see Fig. 28). **D** and **E:** 2 poor hierarchical arrangements of the seed regions as measured by the proportion of violated hierarchical and lateral constraints. Relative ordering of the seed regions (**A** and **B**) within the functional hierarchy agrees well with the proposed macaque visual hierarchy (see text).





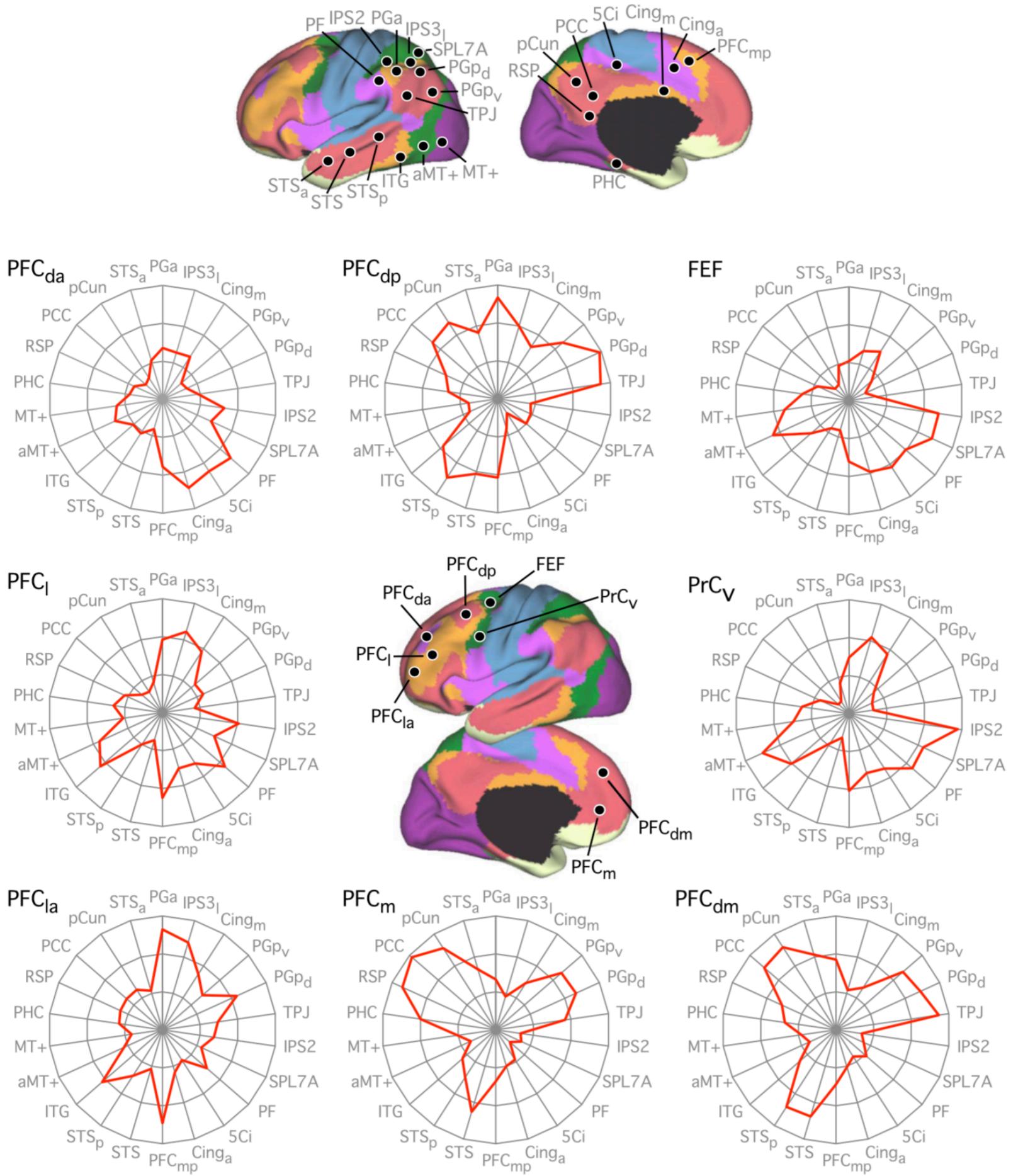


Fig. 31. Adjacent frontal regions exhibit distinct functional connectivity fingerprints. The format and plotting are the same as for Fig. 30 with regions tailored for exploration of frontal cortex. The coordinate locations are reported in Table 4. The polar scales range from  $r = -0.4$  (center) to  $r = 0.5$  (outer boundary) in 0.3-step increments.