Computational Methods in NeuroImage Analysis

Instructor: Moo K. Chung

mchung@wisc.edu

September3, 2010

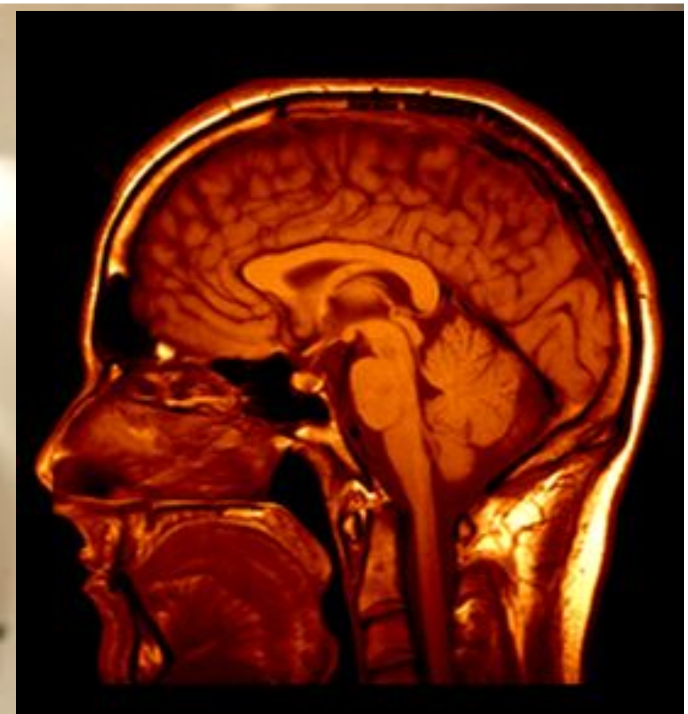
Instructor

Moo K. Chung 정무경
Associate Professor of Biostatistics and Medical Informatics
University of Wisconsin-Madison

Visiting Associate Professor of Brain and Cognitive Sciences
Seoul National University

Office Hour: Friday 12:00-1:00pm.
Email: mkchung@wisc.edu
Official Webpage: www.stat.wisc.edu/~mchung

http://brainimaging.waisman.wisc.edu



Your instructor is from

Waisman Laboratory for Brain Imaging and Behavior

3T MRI, PET, microPET, EEG, MEG, eye tracking, etc. everything under a single roof. Research only facility.

6 faculty + 10 PhD level scientists + 10 postdocs + 5 administrative staff + 50 graduate students + plus many undergraduate students + bunch of rodents and monkeys

Waisman laboratory for brain imaging

Active research areas: autism, depression, mood disorders, emotion related, meditation, DTI, MRI anatomical studies, developmental, animal studies

Dalai Lama & Richard J. Davidson

Course Aims

- To present **computational** and **statistical** techniques used in the field of brain image analysis, with an emphasis on actual **computer implementation**.
- MATLAB is the language of instruction but students can use any computer language to do a course project.

Target Audience

- This course is designed for researchers and students who wish to analyze and model brain images quantitatively beyond t-statistics and ANOVA.
- The course material is applicable to a wide variety of other medical and biological imaging problems.
- Course requirement: none.

Course Evaluation

- Submission of research proposal & preliminary analysis before the course drop deadline. 10%
- Give 20-minute oral presentation at the end of the semester. 20%
- Final exam at the end of November. 30%
- Submission of the final research report of about 15-25 pages excluding figures, tables and references. It also should contain more than 20 related references. 40%

Course Workload

- Approximately 10-20 hours/week depending on the qualification of students assuming you have 60-80 work hours/week.

- Read my lecture notes, textbook and about two assigned papers per week.

Course Topics

- Numerical techniques for (ordinary and partial) differential equations, FEM
- Spectral methods (Fourier analysis, PCA, sparse-PCA, functional-PCA, marching persuit)
- Optimization (least squares, multivariate general linear model (MGLM), L1-norm minimization, maximum likelihood)
- Discrimination and classification (linear, quadratic and logistic discrimination and SVM).
- Geometric and topological computation (curvatures, Euler characteristics, other topological invariants).
- Brain connectivity & network modeling

Course website

brainimaging.waisman.wisc.edu/~chung/neuro.processing/

Lecture notes will be uploaded 30mins before each lecture.
Feel free to bring laptops for note taking and web surfing.

Textbook

"Computational Neuroanatomy: The Methods" to be published in 2011. It can be downloaded from the webpage. It's huge at 60-100MB.

Sample data & Codes

Look for directory \data and \matlab few hours before each lecture starts.

Class discussion board

groups.google.com/group/brainimage

If you don't become a member, you won't receive any email from me.

Tips for students

1. Your best friend
www.google.com

2. Your second best friend
scholar.google.com

Occam's razor

- When given two equally valid explanations (model) for a phenomenon, one should embrace the less complicated formulation (model).

- All things being equal, the simplest solution tends to be the best one.

- If you want to try complicated modeling, do the simplest model first.

Do not try bang your head on the wall trying to do a complicated analysis when you can't even build a simpler model.

NOTES

1. **No plagiarism of any sort** will be allowed in the course.

2. Work alone for the project. But feel free to discuss all other matters with class mates and the instructor.

3. Office hour: talk to me after each class or send email to mkchung@wisc.edu to set up the appointment.

Lecture 1

Overview of Computational Methods

September 3, 2010

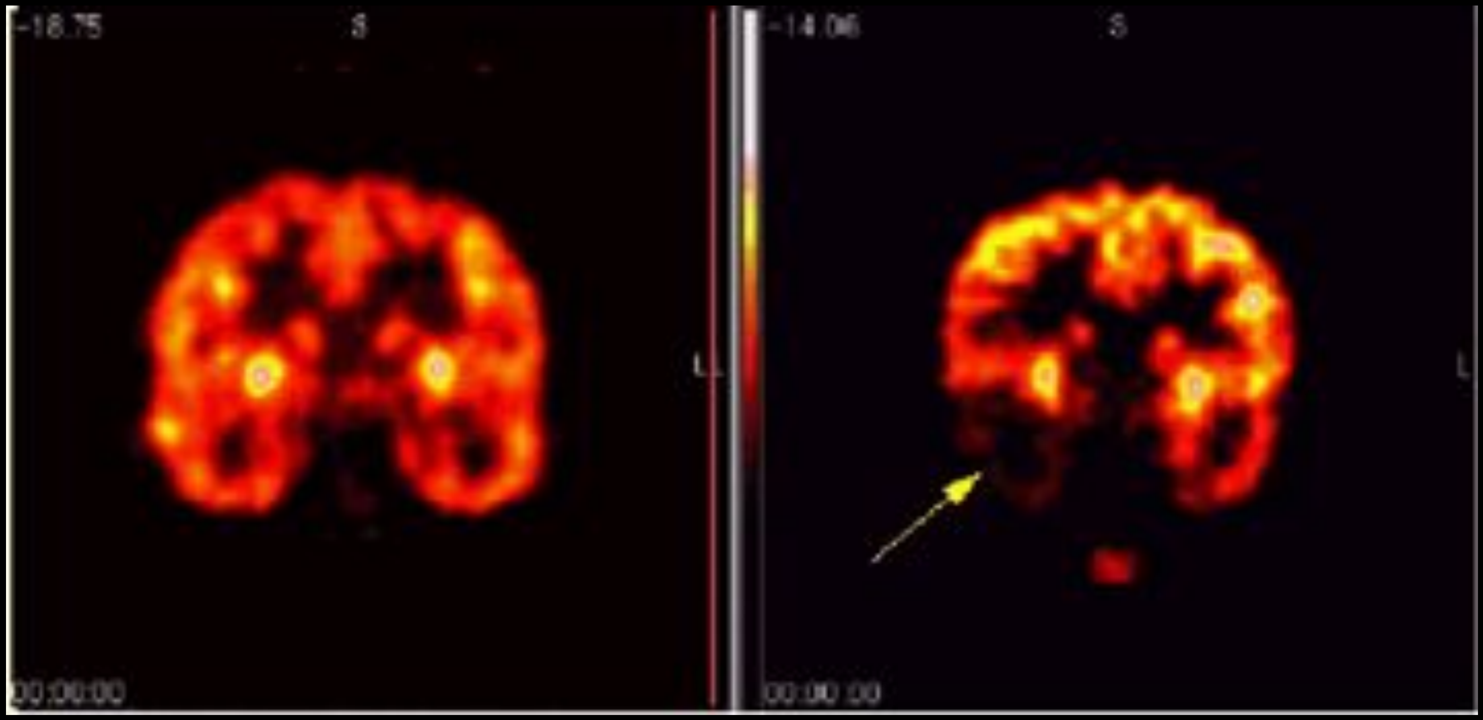
DATA

- Brain images: various imaging modalities can be modeled and analyzed in a similar mathematical fashion.

- Examples of brain images:

MRI, fMRI, PET, DTI

Positron Emission Tomography (PET)

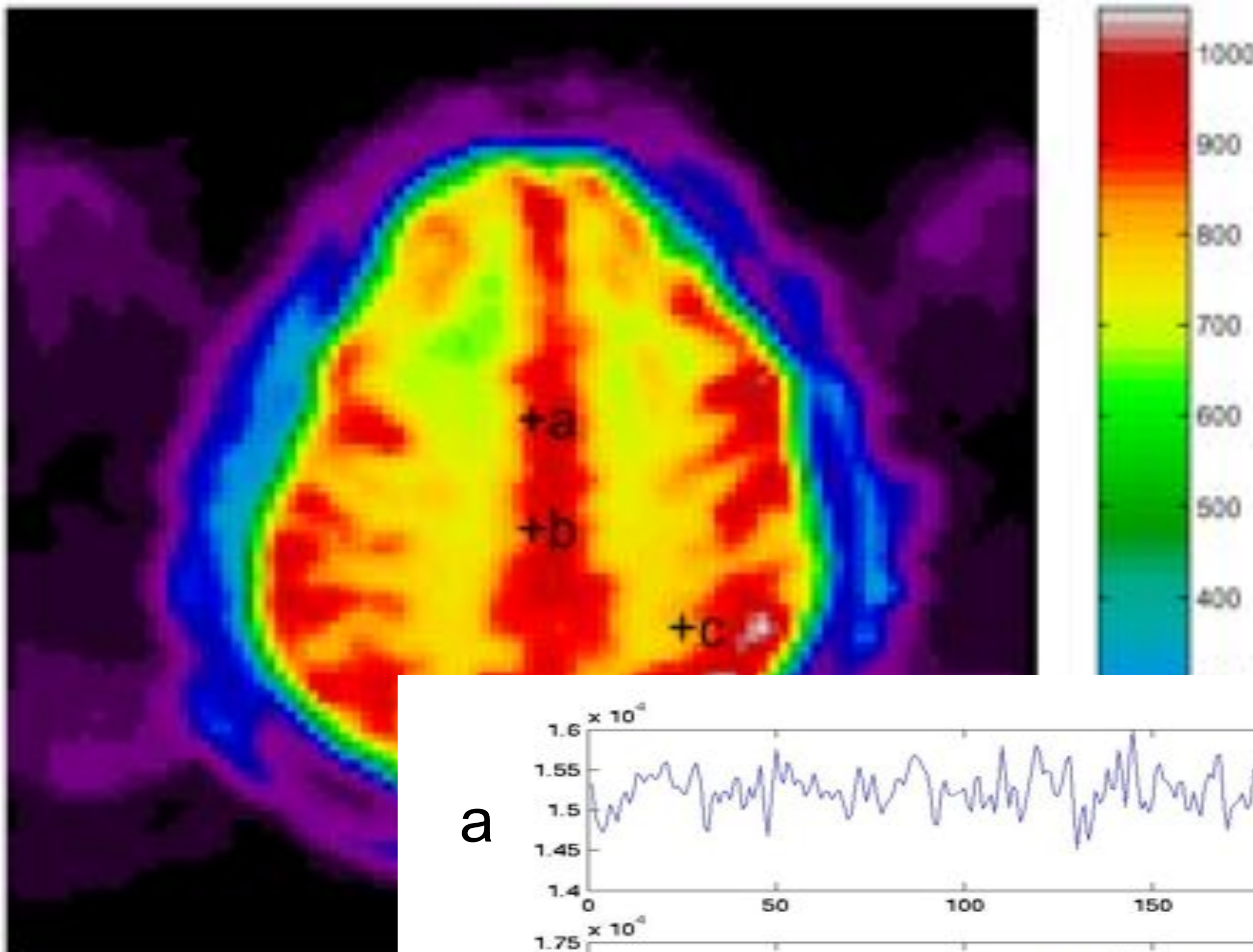


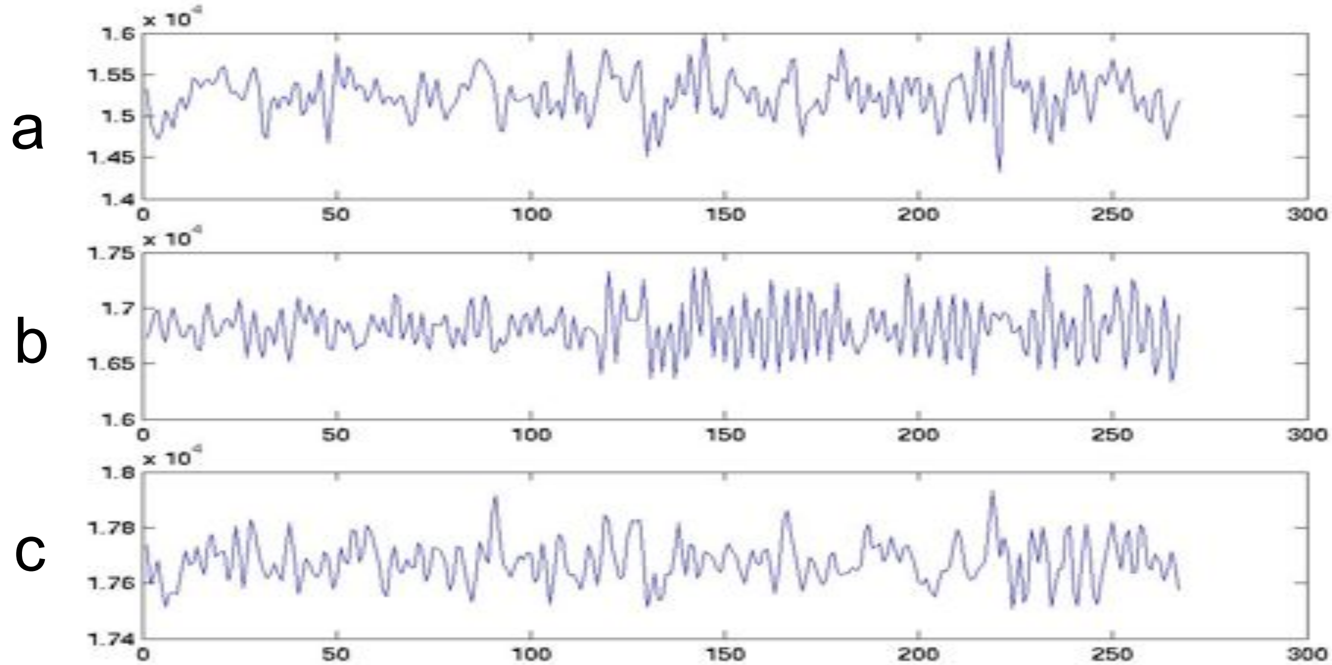
Normal Brain

Brain of 9 year old girl suffering from epilepsy.

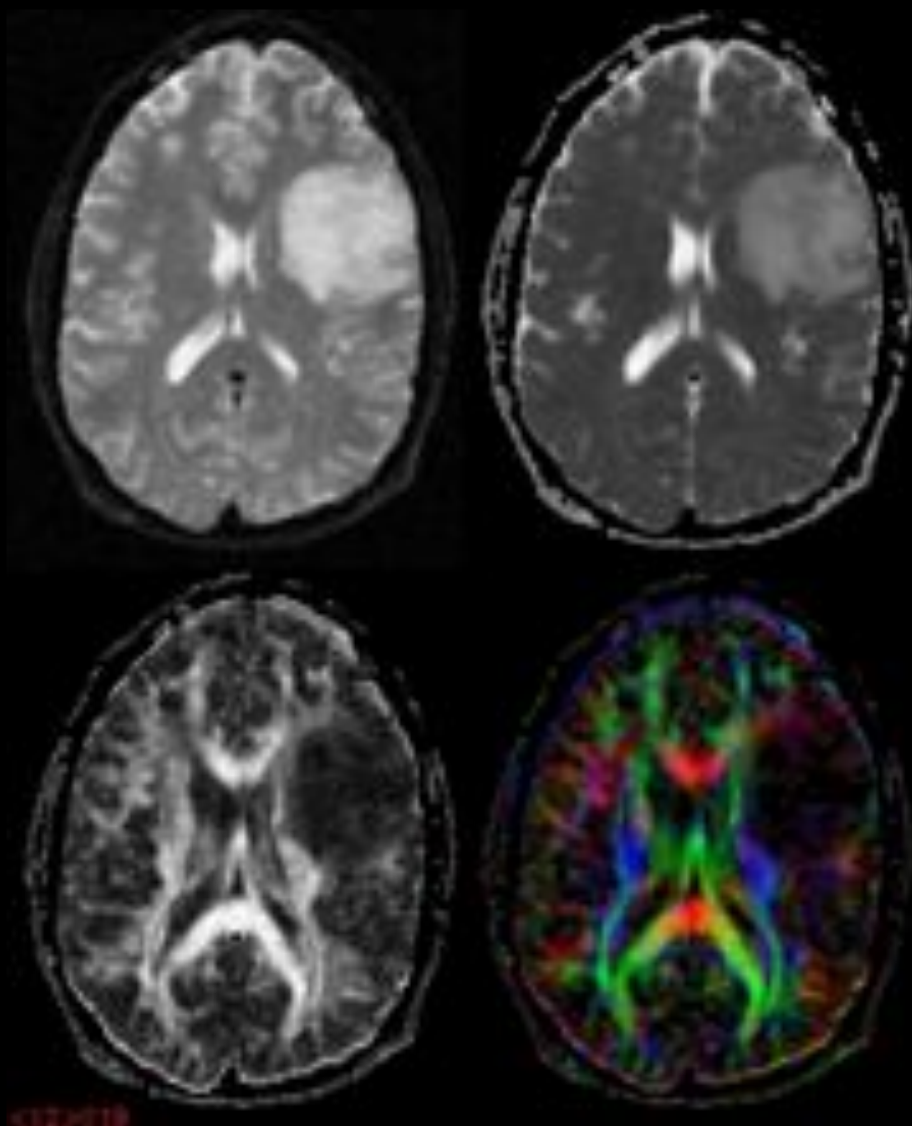
Montreal Neurological Institute

Functional MRI





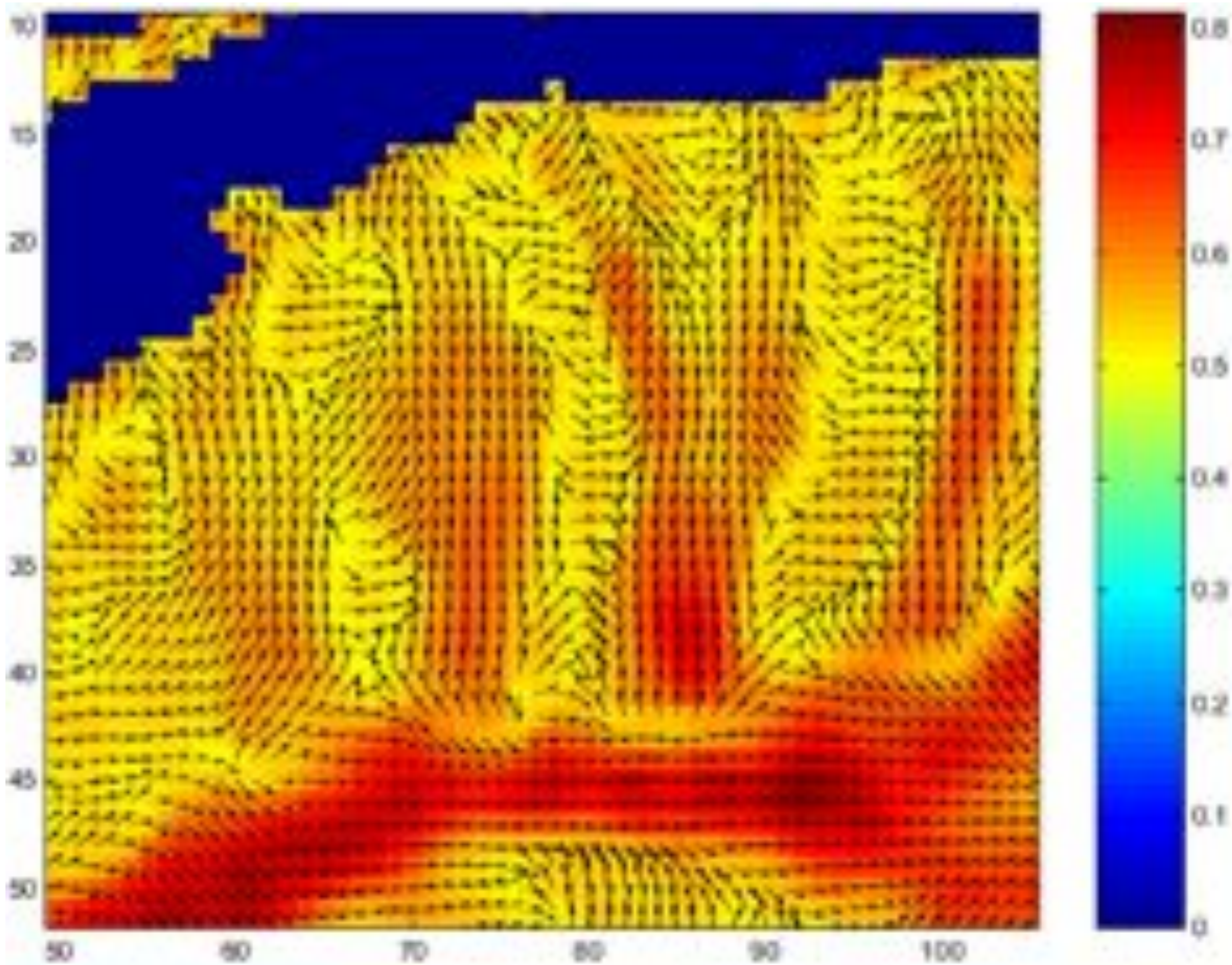
Diffusion tensor imaging (DTI)



Tensor data
= 3 by 3 matrix values
at each voxel are diffusion
coeffiients.

Andrew L. Alexander
University of Wisconsin-Madison

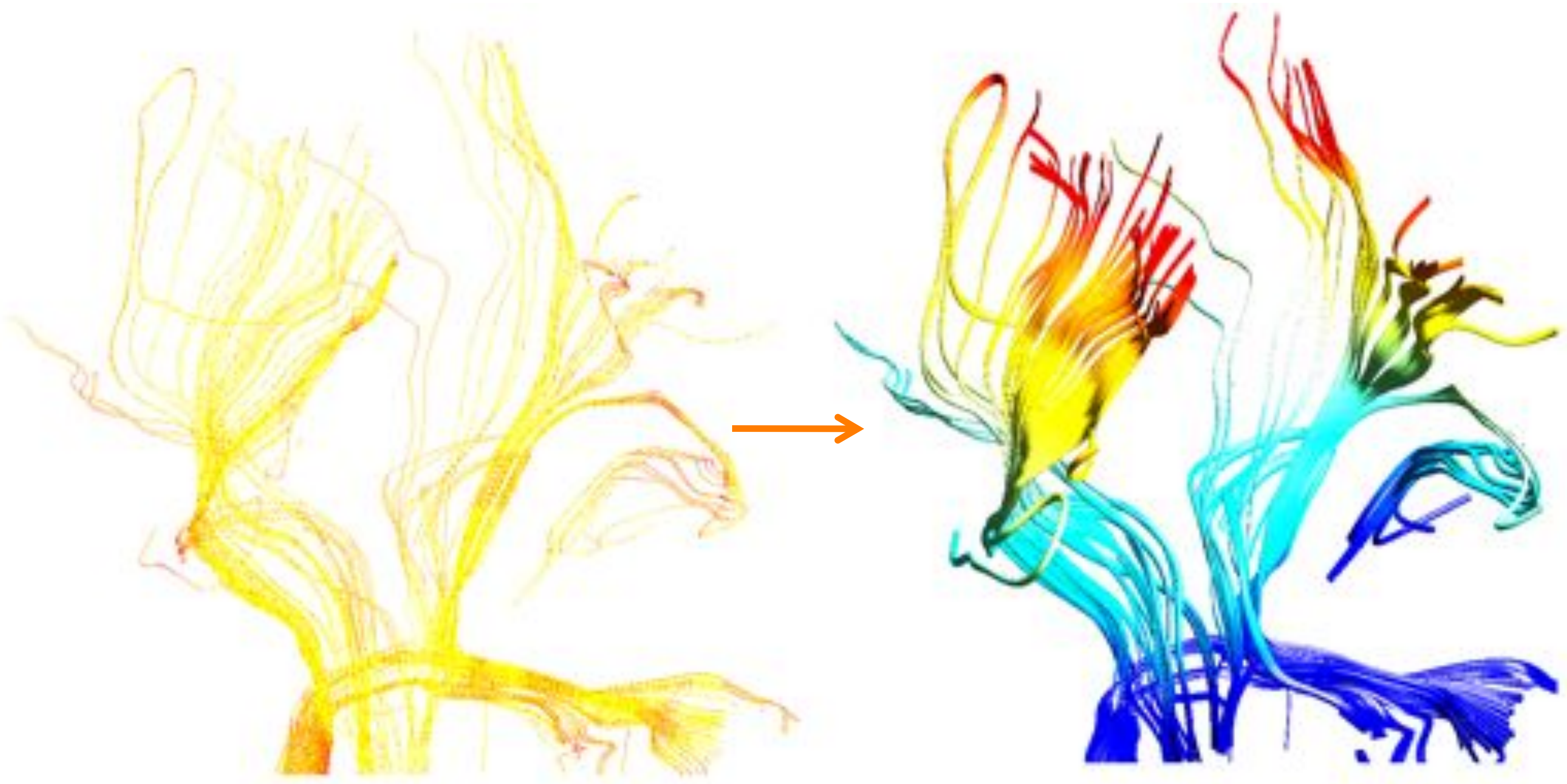
Principal eigenvectors of the diffusion coefficient matrix can be considered as the tangent vector of the stream lines that represents white fiber.

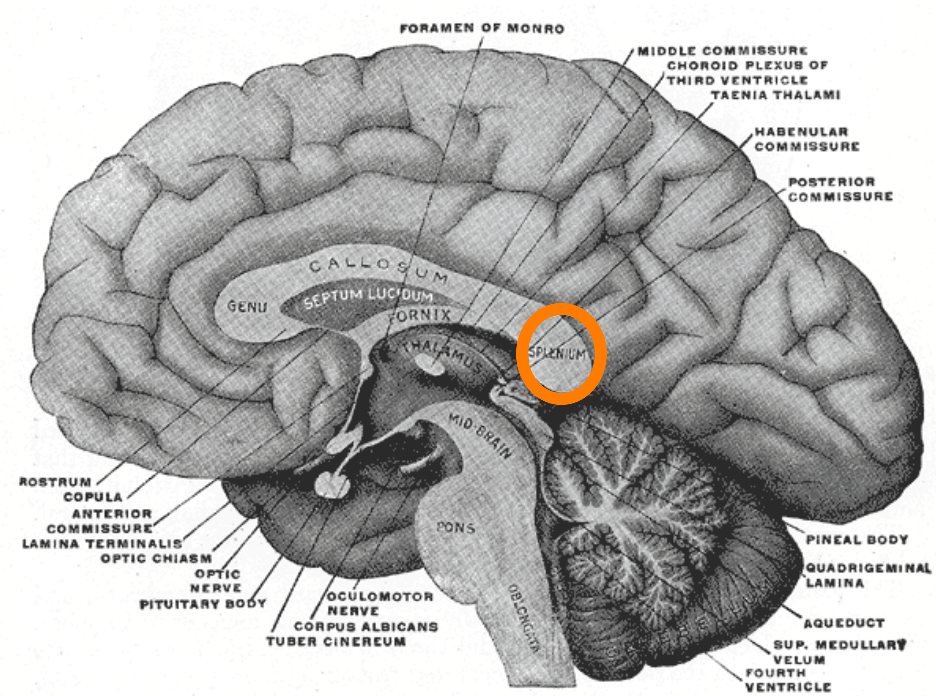


Intensity = eigenvalues

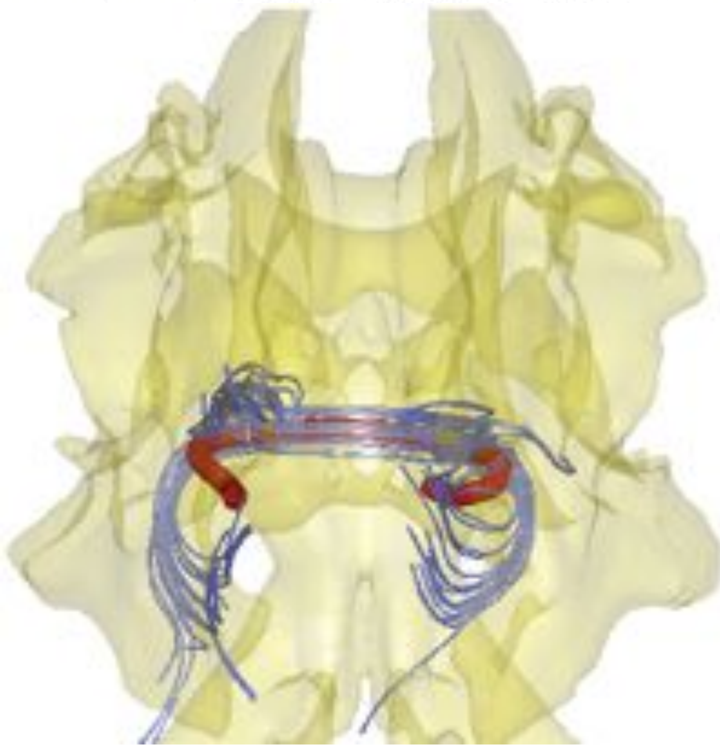
Streamline based tractography

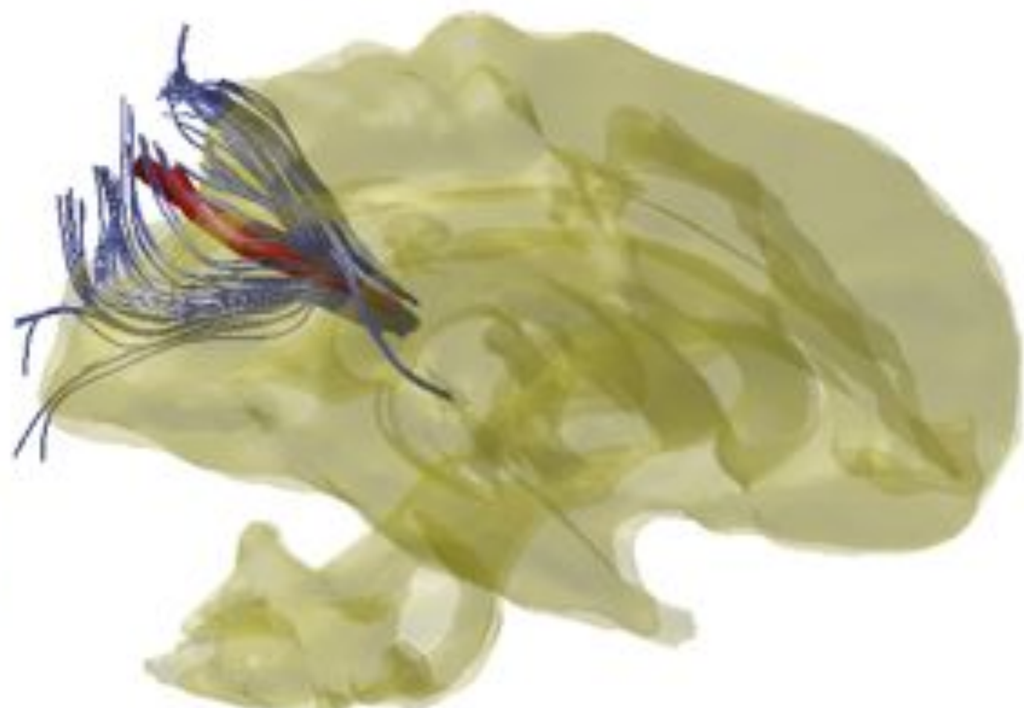
second order Runge-Kutta algorithm (Lazar et al., HBM. 2003).



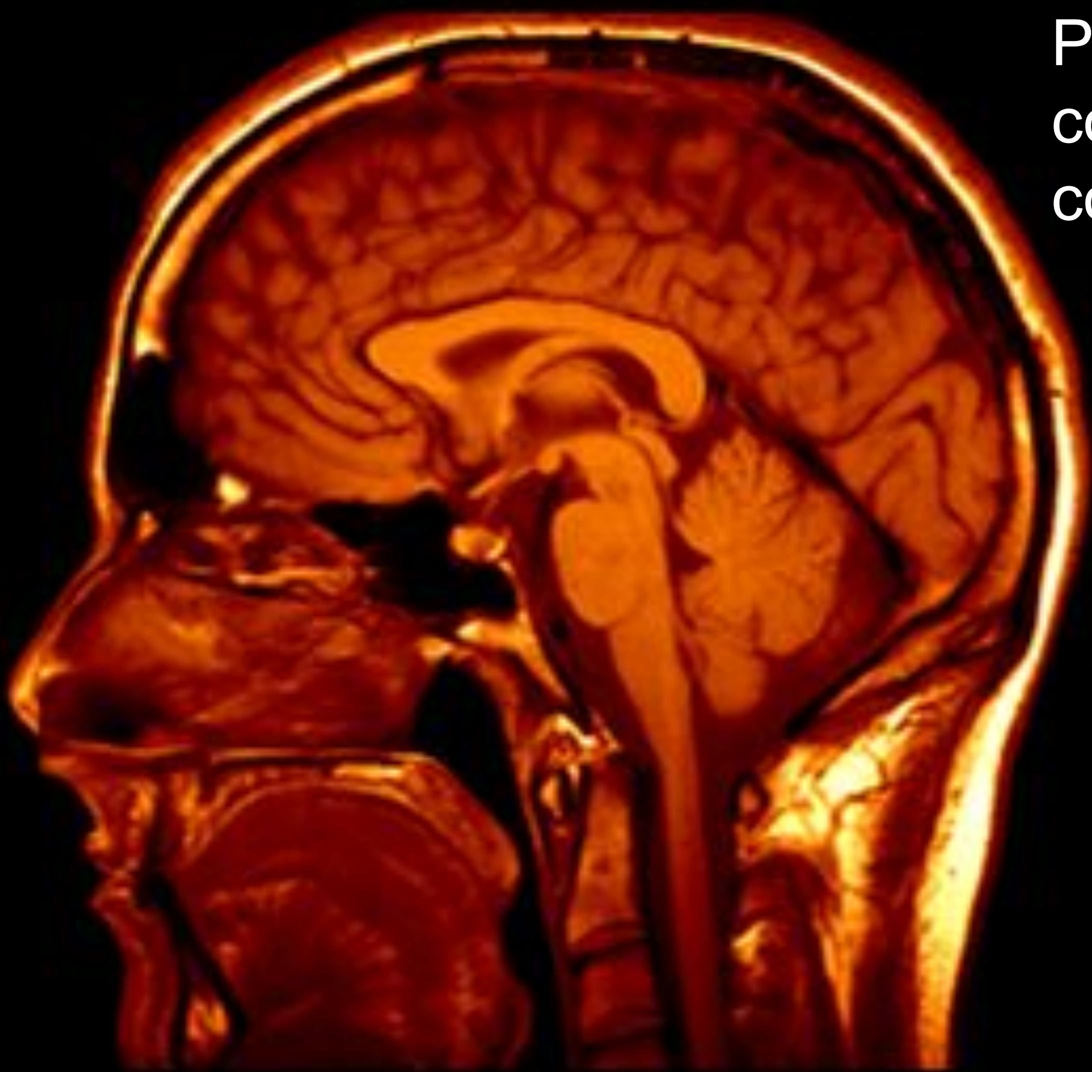


Tracts passing through the splenium of the corpus callosum





3T Magnetic resonance imaging (MRI)



Provide greater image contrast in soft tissues than computed tomography (CT)

Computational Issues in Brain Image Analysis

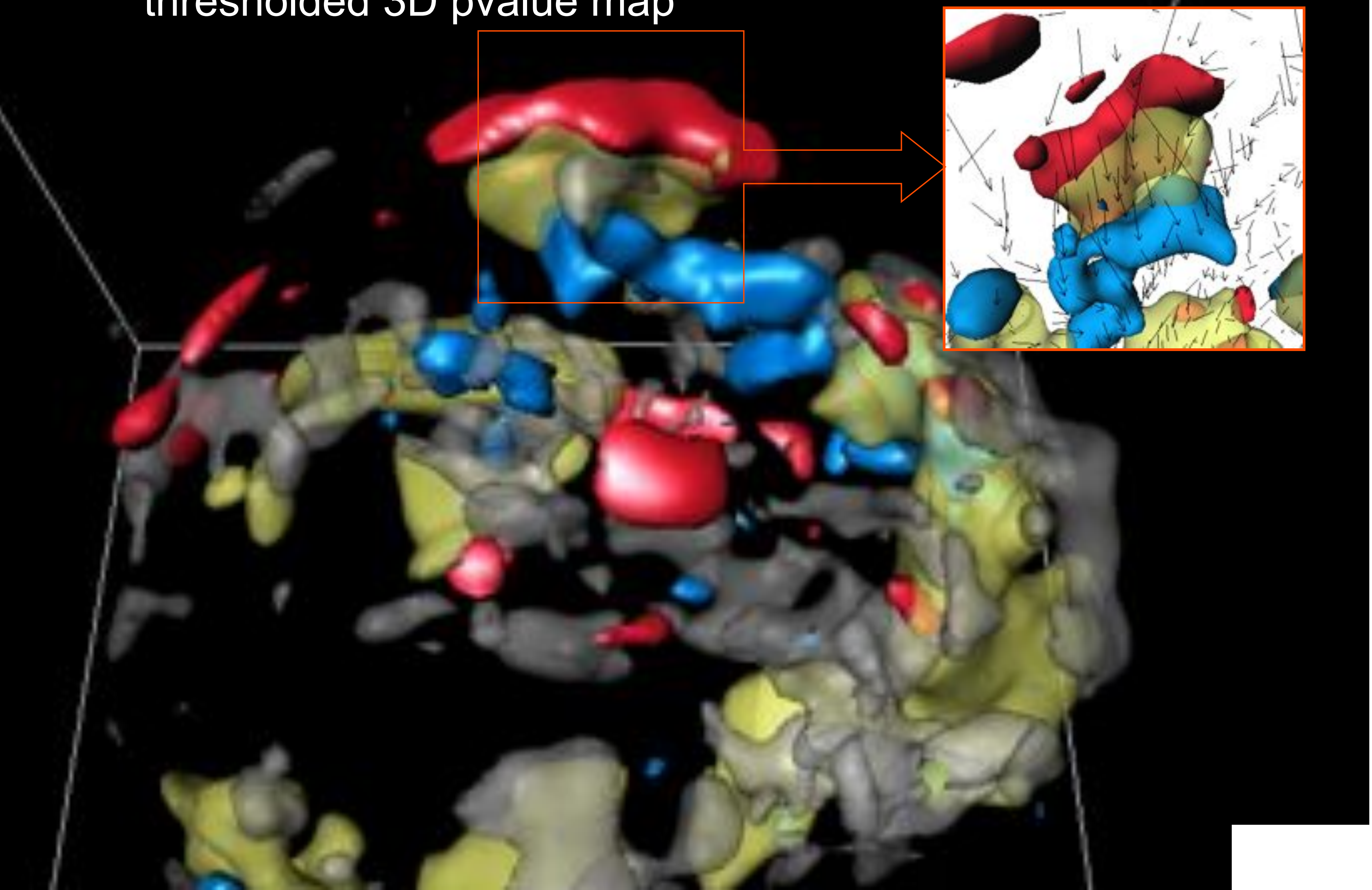
- Differential equations (ordinary & partial)
- Variation & optimization (least squares, L1 norm minimization)
- Spectral approaches (Fourier, PCA etc)
- Discrimination & classification
- Geometric & topological computation

Data & Image Visualization

Data & image visualization has to be your first step in analyzing images

Statistical visualization is an important issue

thresholded 3D pvalue map

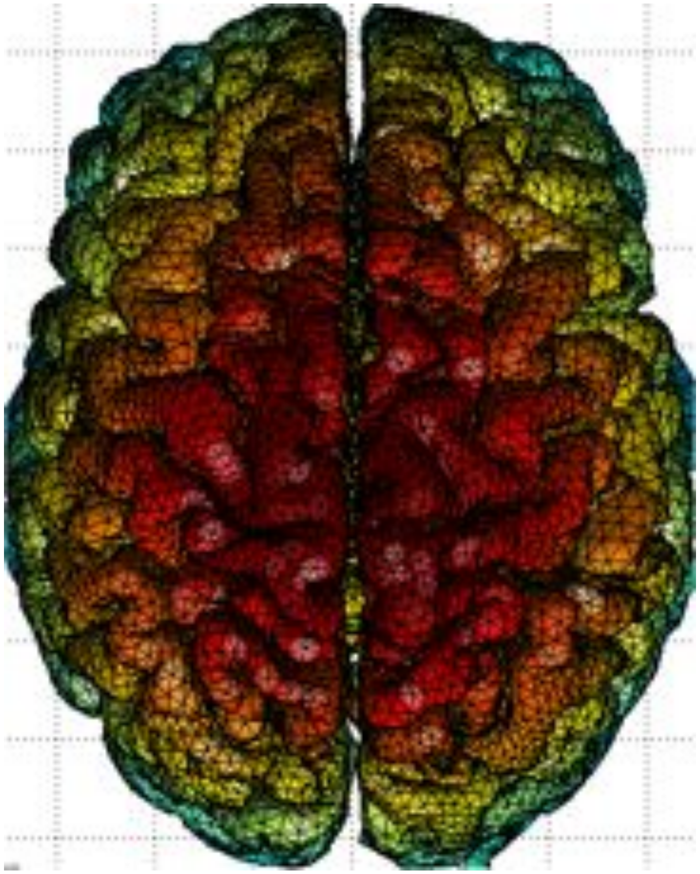


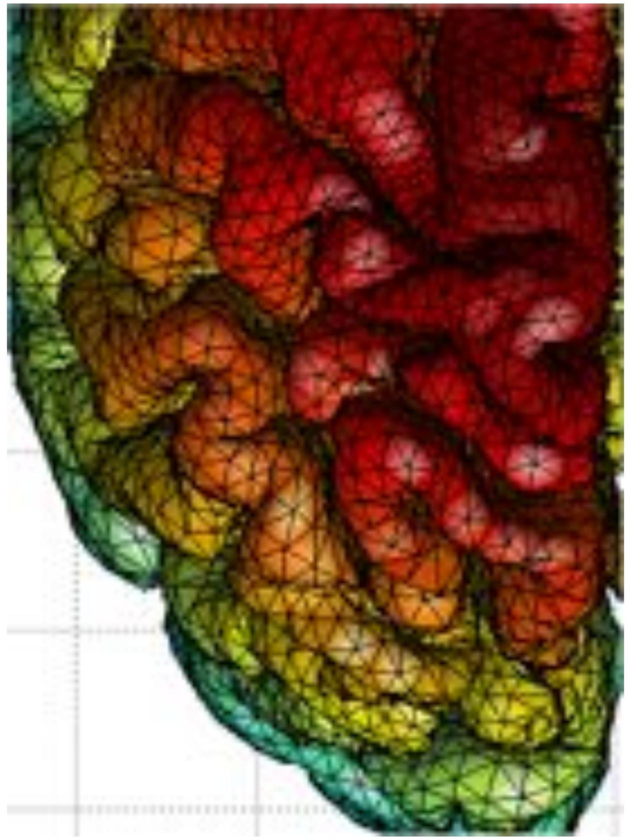
Least squares estimation

statistical parameter estimation technique
by the sum of squared residual

Cortical Surface
Polygonal mesh

Mesh resolution 3mm





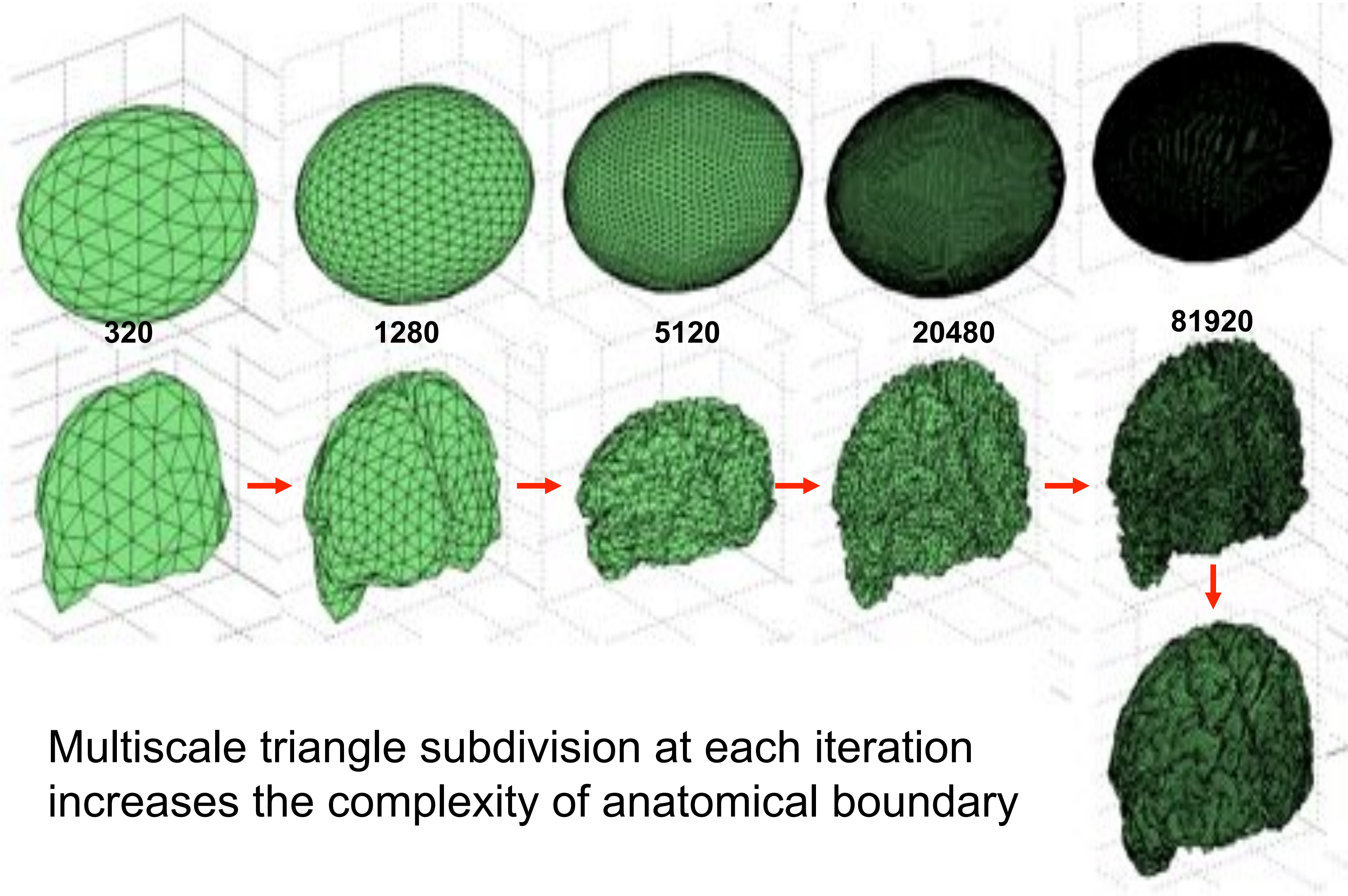
82,190 triangles

40,962 vertices

Spherical harmonic representation ↓

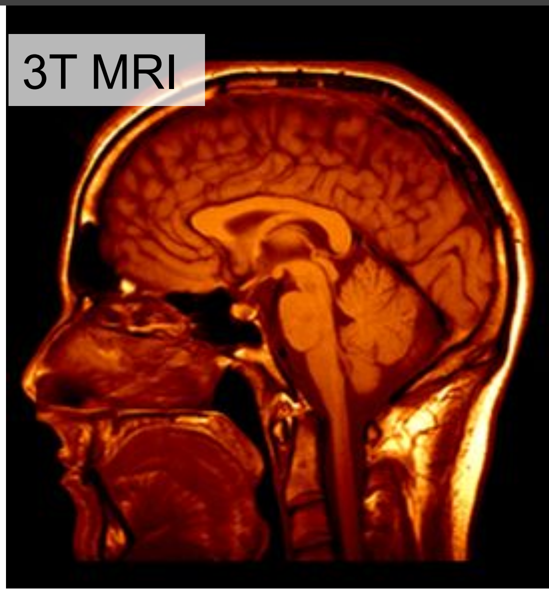
20,000 parameters per surface

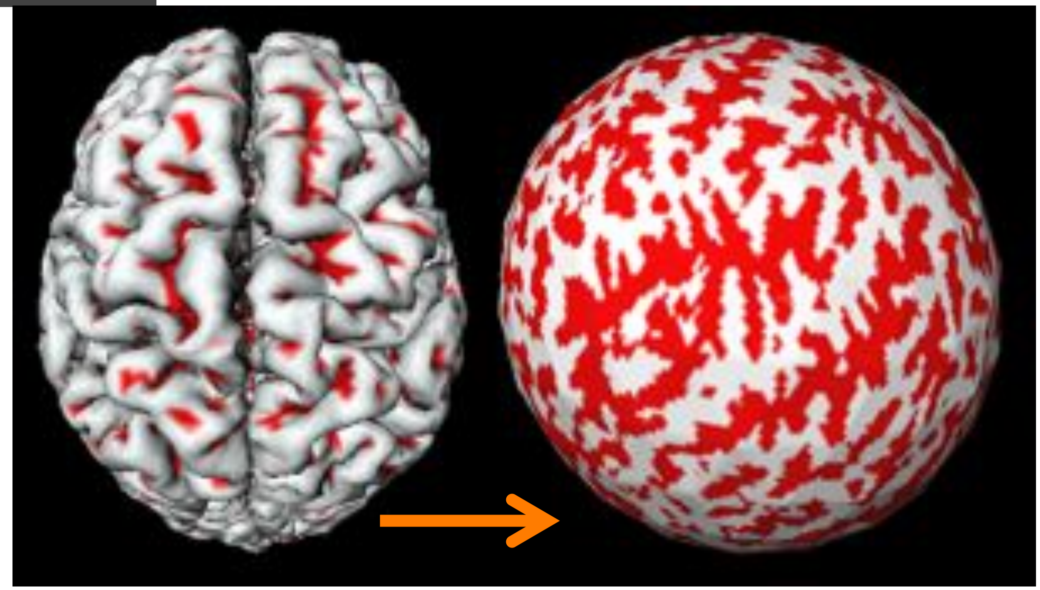
Deformable surface algorithm McDonalds *et al*. (2001) NeuroImage



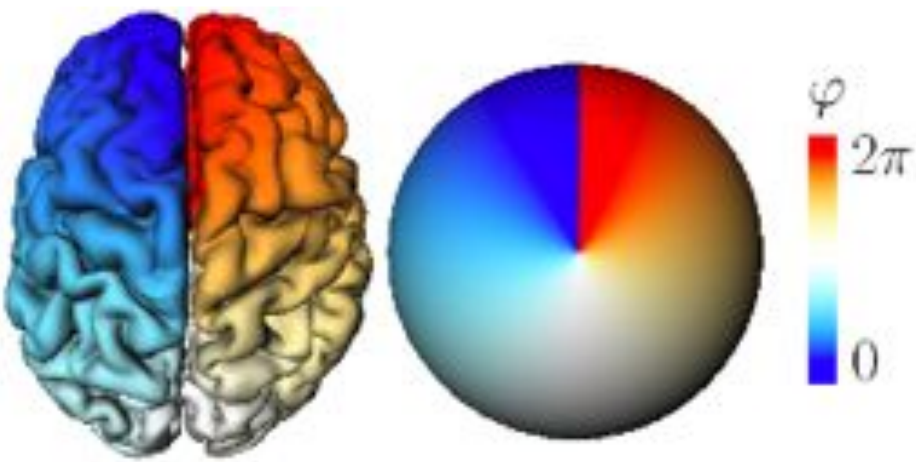
Multiscale triangle subdivision at each iteration increases the complexity of anatomical boundary

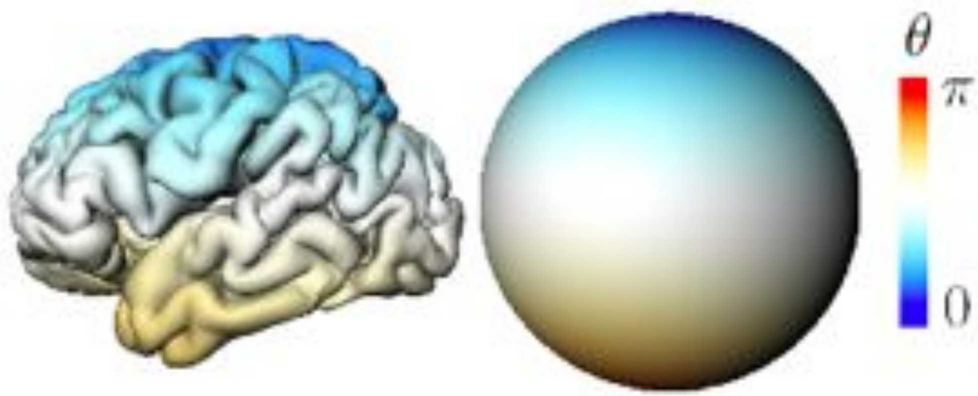
Parameterize mapping to a sphere





Deformable surface algorithm



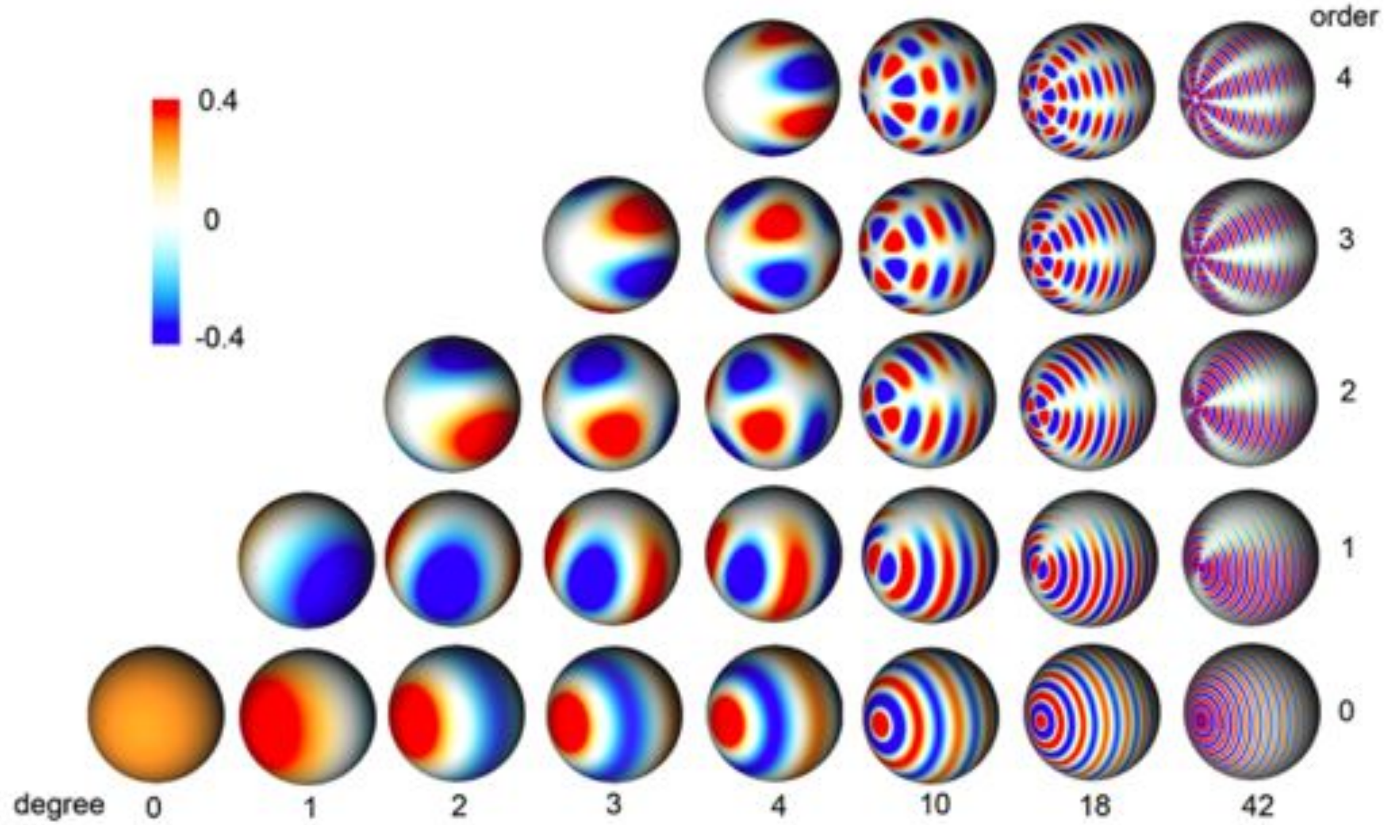


Spherical angle based coordinate system

Spherical harmonic of degree *l* and order *m*

$$Y_{lm} = \begin{cases} c_{lm} P_l^{|m|}(\cos\theta)\sin(|m|\varphi), & -l \le m \le -1, \\ \frac{c_{lm}}{\sqrt{2}} P_l^0(\cos\theta), & m = 0, \\ c_{lm} P_l^{|m|}(\cos\theta)\cos(|m|\varphi), & 1 \le m \le l, \end{cases}$$

Coarse anatomical detail



Fine anatomical detail

Coarse detail ⟷ Fine detail

SPHRM representation

•Given functional measurement *f(p)* on a unit sphere, we represent it as

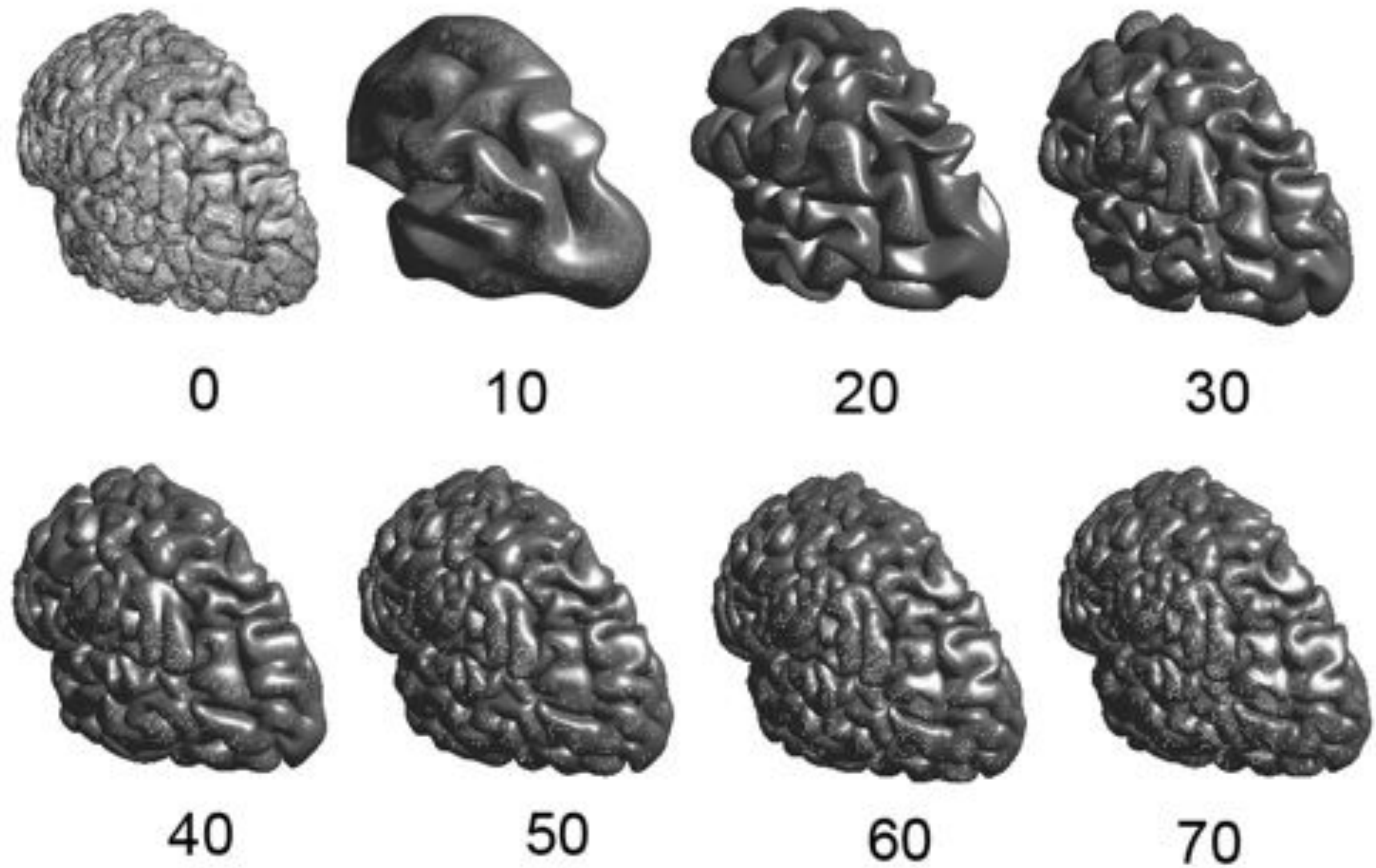
$$f(p) = \sum_{l=0}^{k} \sum_{m=-l}^{l} f_{lm} Y_{lm}(p) + e(p)$$

e: noise (image processing, numerical, biological)

f_{lm}: unknown Fourier coefficients

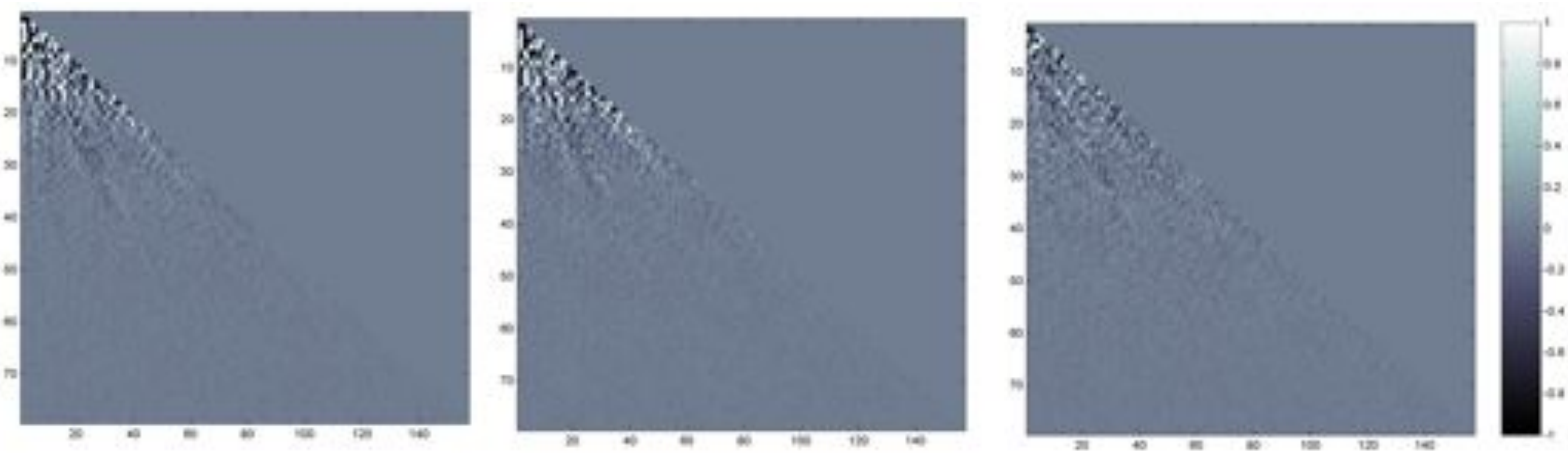
•The parameters are estimated in the least squares fashion.

FreeSurfer results



6241 Fourier coefficients can be used to quantify individual anatomical shape variations

Average SPHARM coefficients

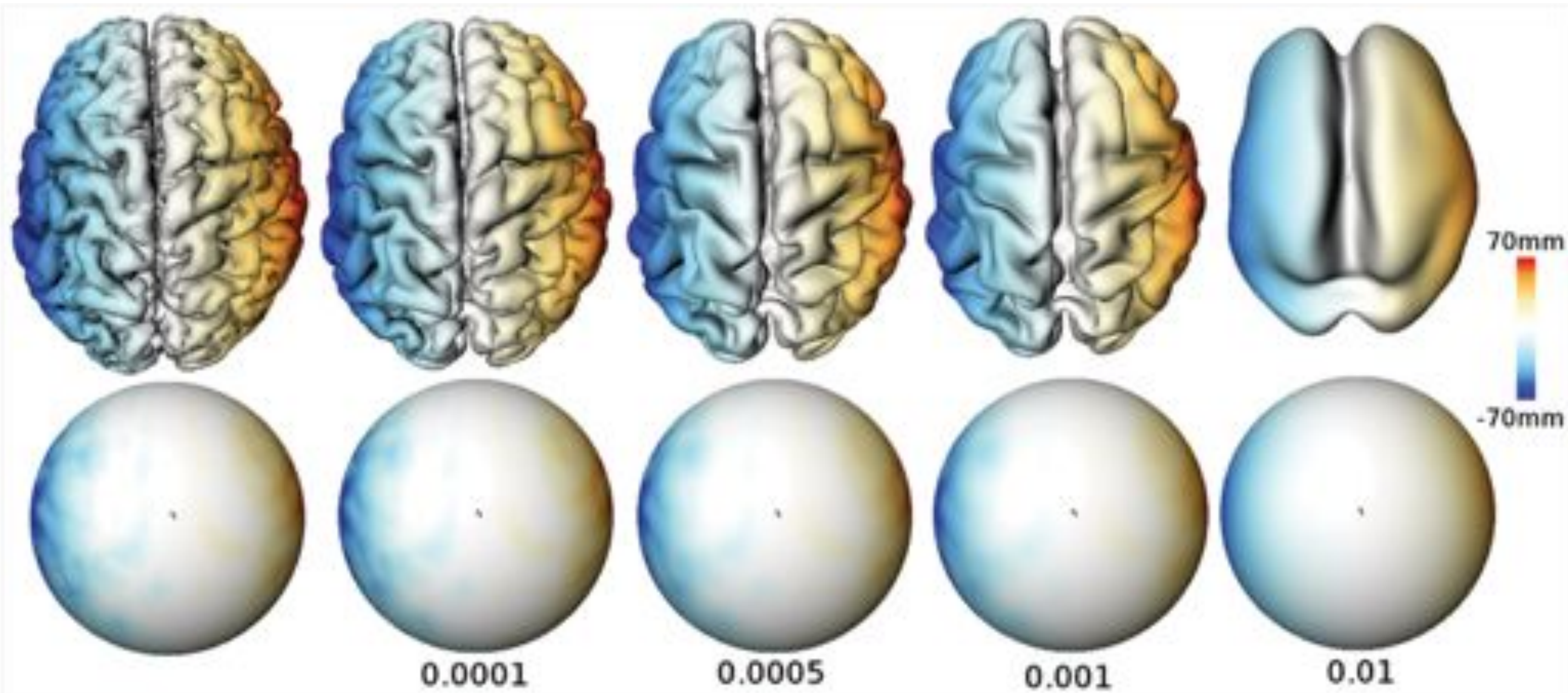


autistic control difference

78th degree Weighted-SPHARM representation

$$\sum_{l=0}^{k}\sum_{m=-l}^{l} e^{-l(l+1)t}\langle f, Y_{lm}\rangle Y_{lm}(\theta,\varphi)$$

Weighted-SPHARM



Color scale= x-coordinate

Maximum likelihood estimation

statistical parameter estimation technique
by maximizing a likelihood function

Image segmentation

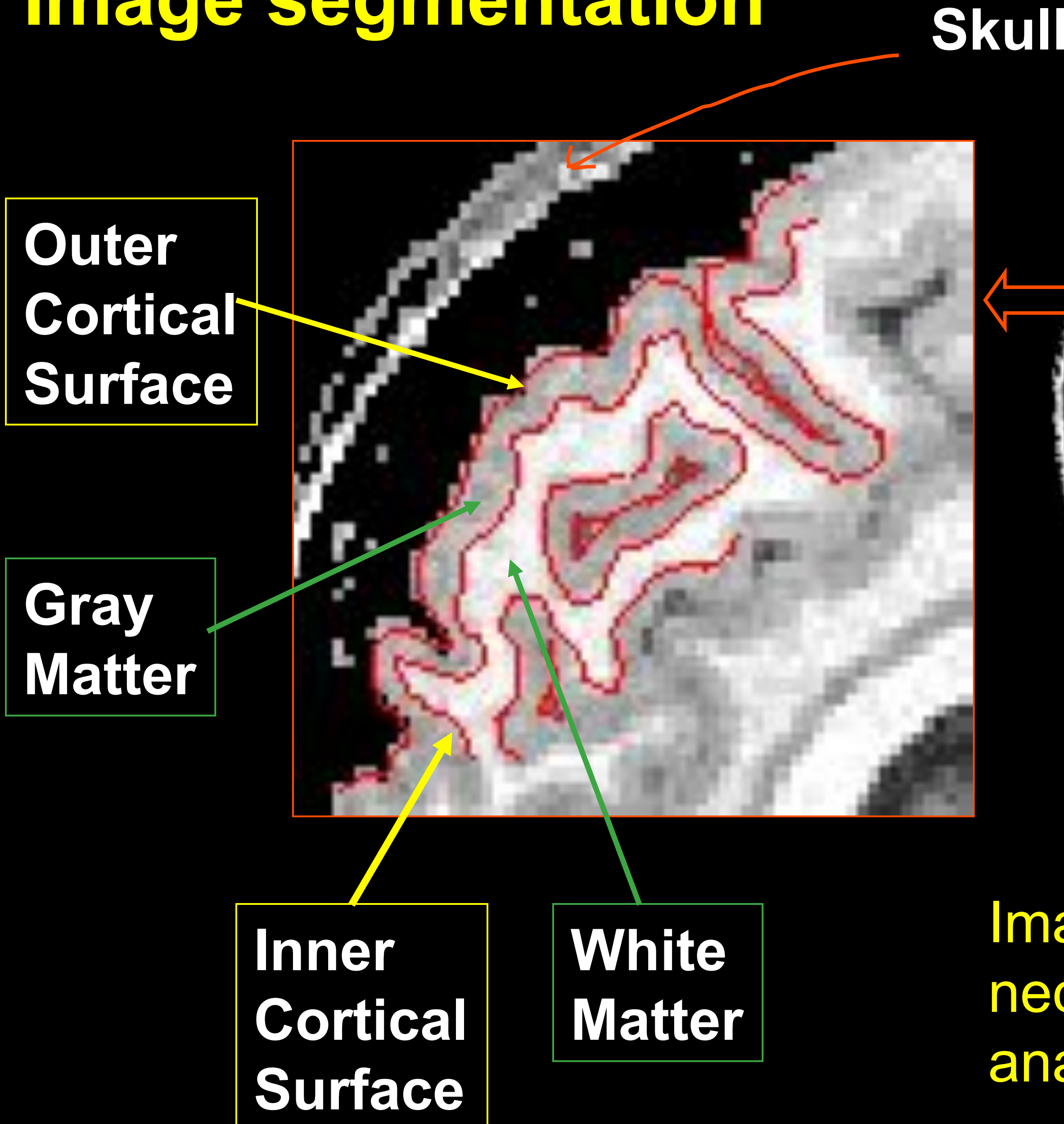
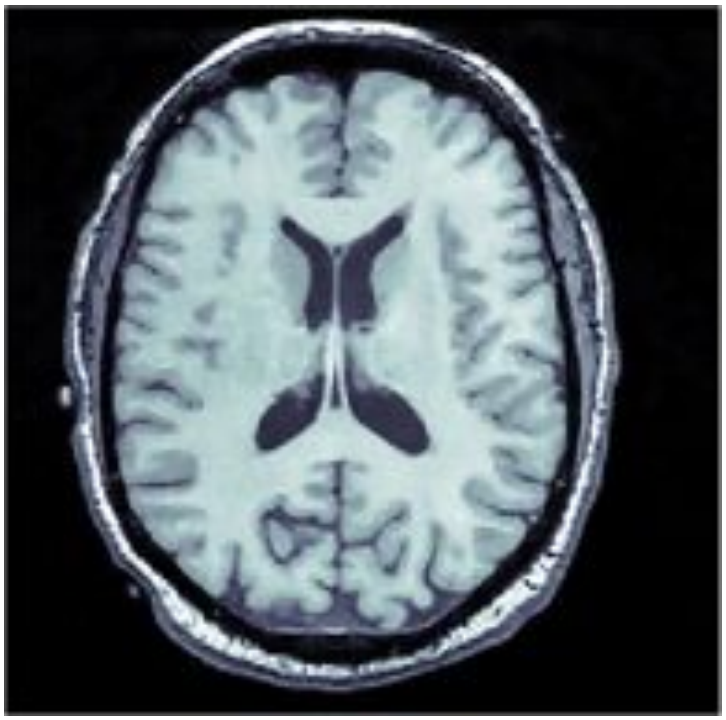
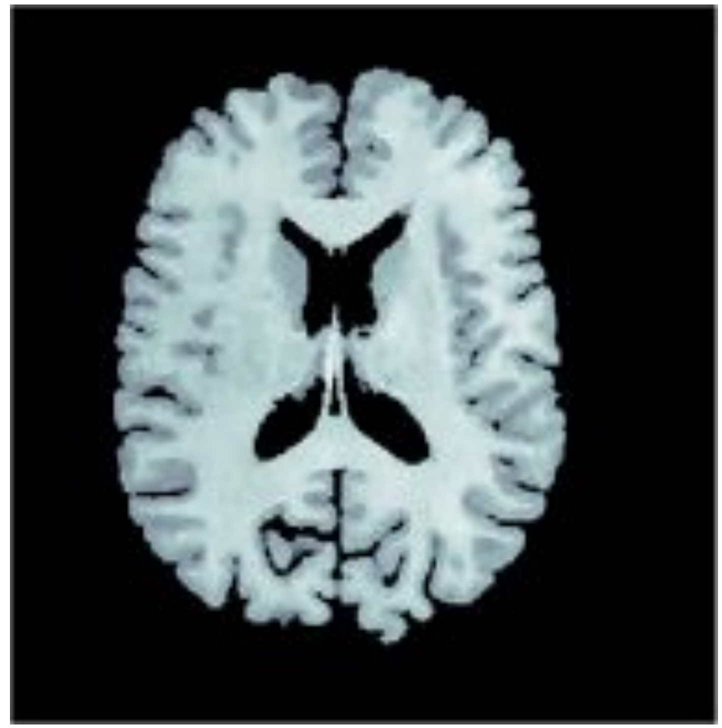


Image segmentation is necessary to quantify anatomical substructures

Segmentation based on Gaussian mixture model
SPM result

Automatic skull stripping can remove unwanted anatomical regions automatically.

Two-components Gaussian mixture model

$$f(y) = pf_1(y) + (1-p)f_2(y)$$

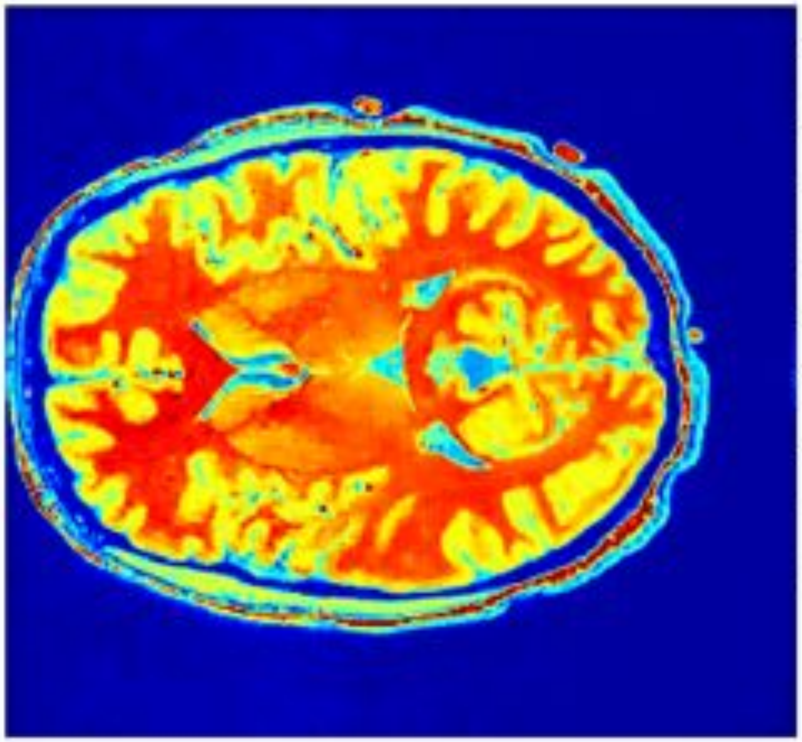
$$f_1(y) \approx N(\mu_1, {\sigma_1}^2)$$

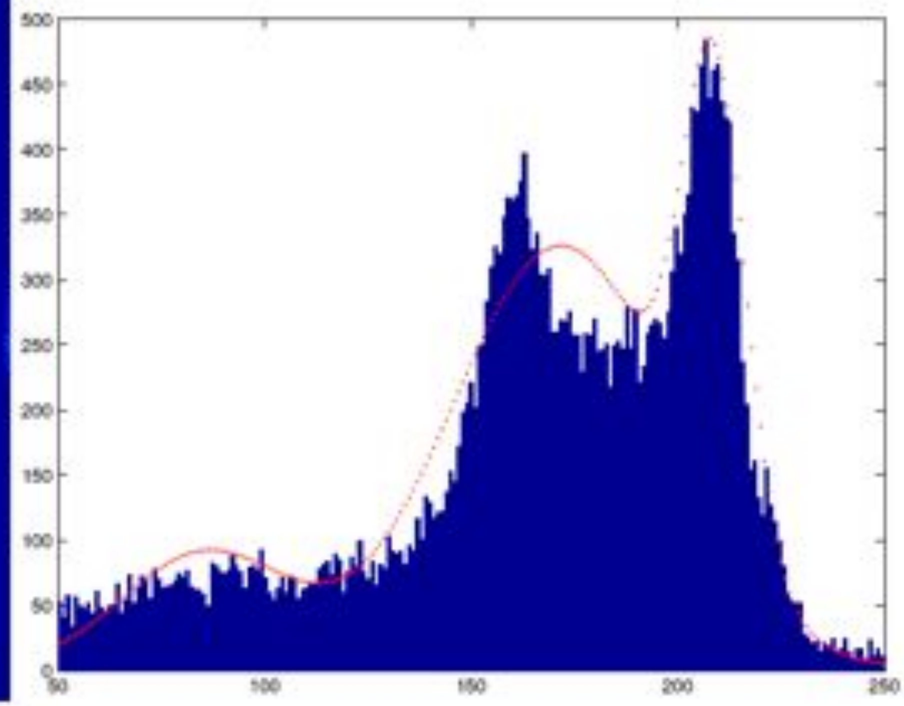
$$f_2(y) \approx N(\mu_2, {\sigma_2}^2)$$

p = mixing proportion → estimated tissue density

Parameters are estimated by the EM-algorithm

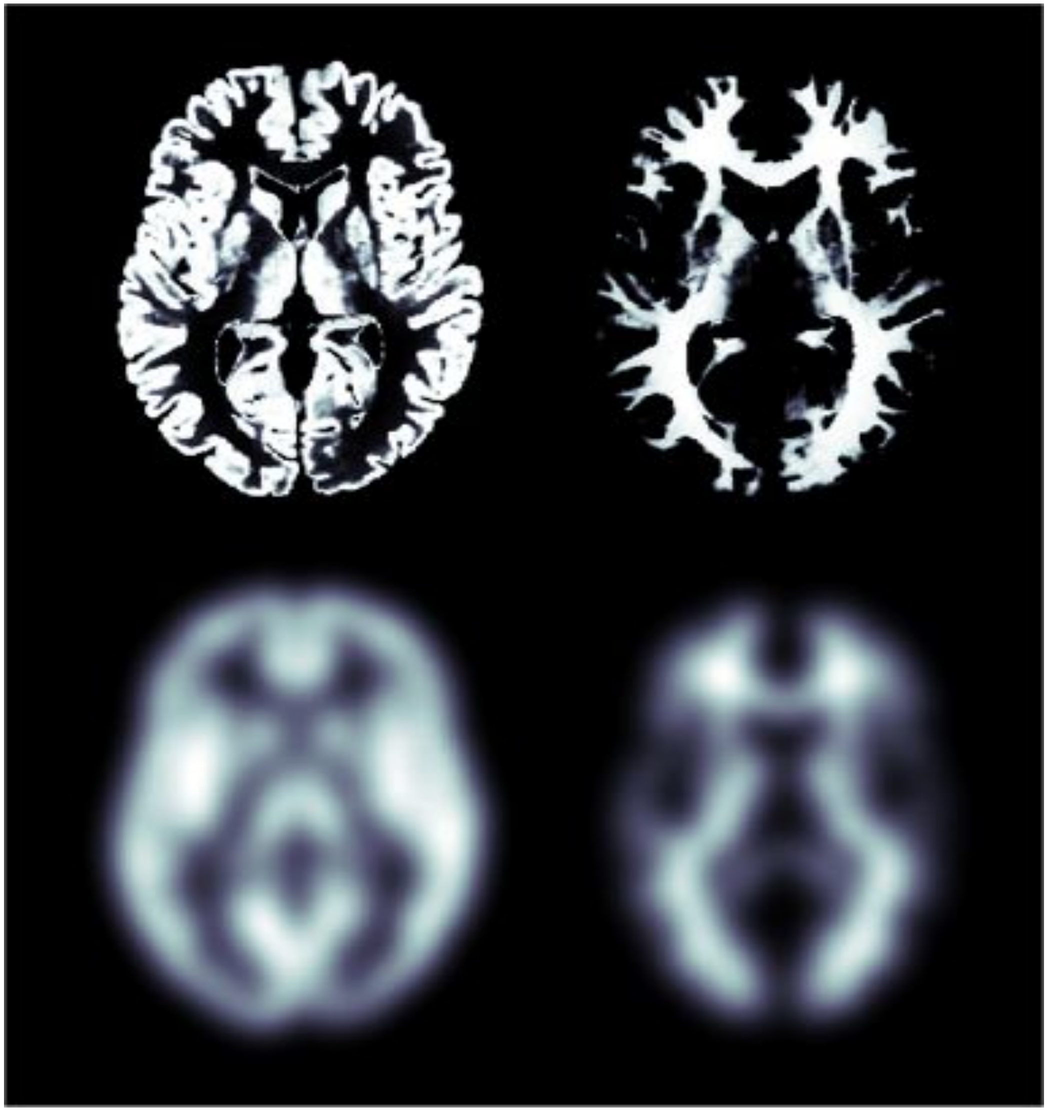
Gaussian mixture modeling - EM algorithm





Shubing Wang
Merck

Segmentation result



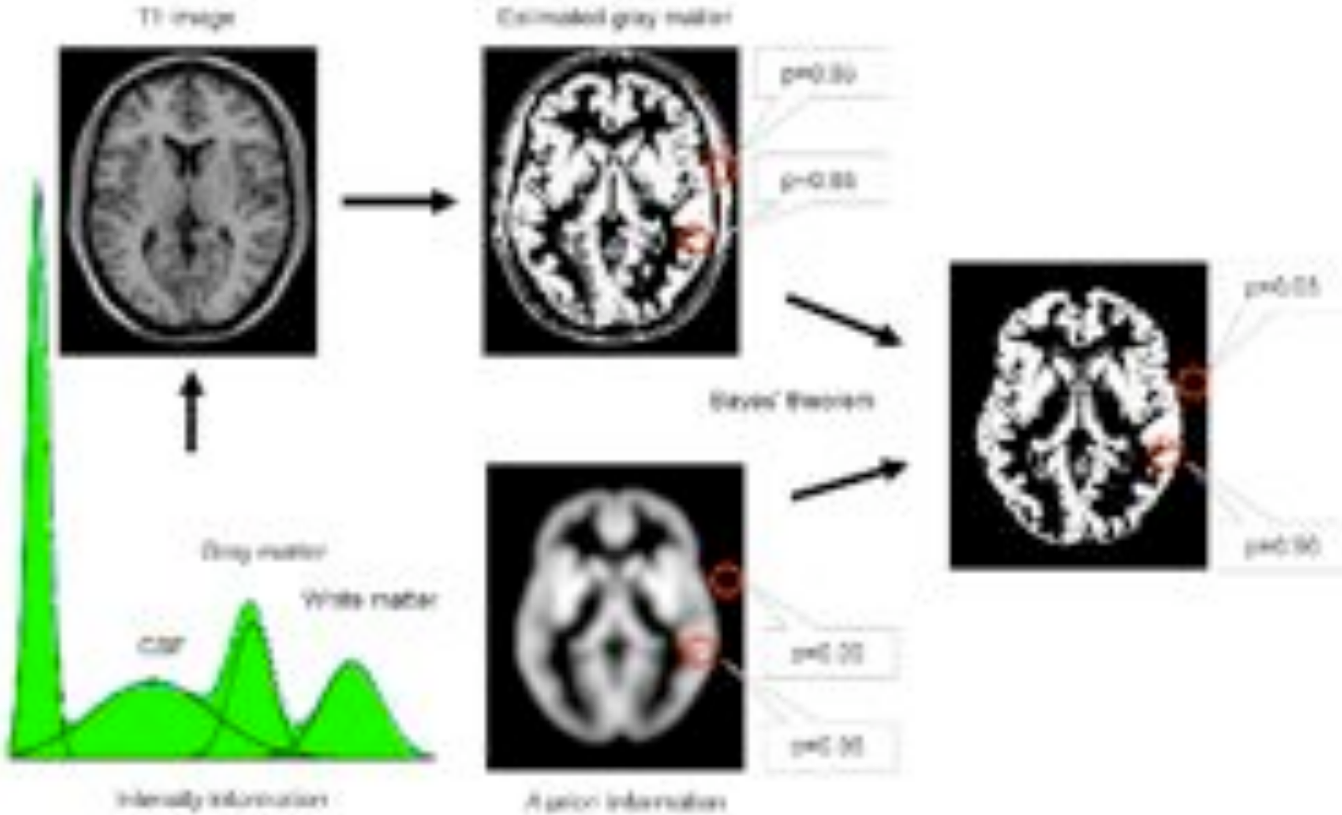
Binary masks: 0 or 1

Gaussian
kernel smoothing

Probability map [0 , 1]

Gray matter White matter

Probabilistic segmentation



Mietchen & Gaser, 2009

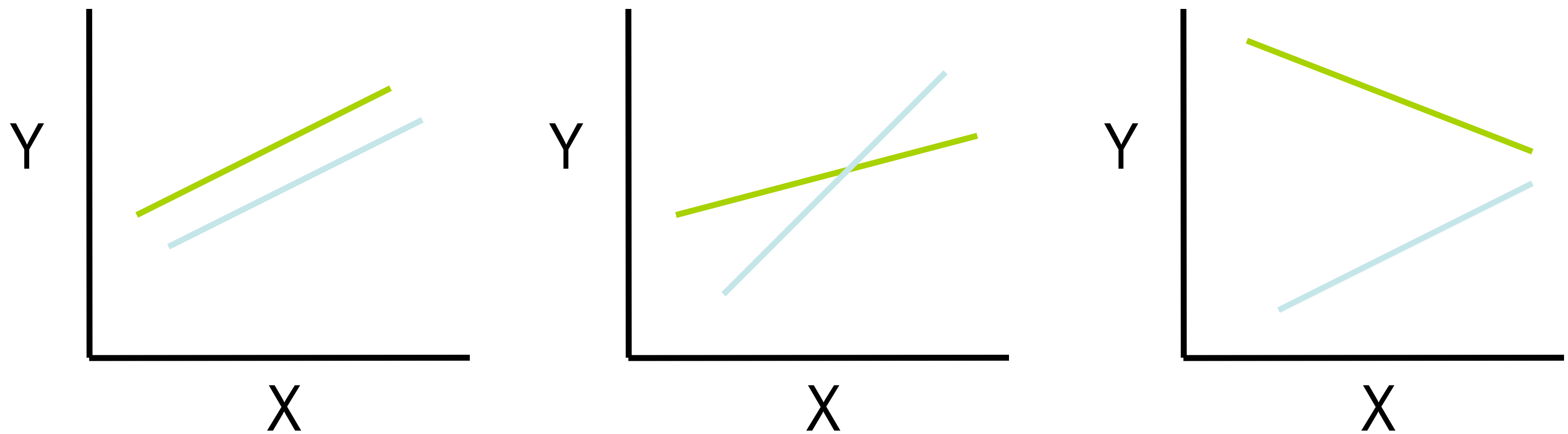
Multivariate General Linear Model

Multivariate version of general linear model.
SPM, AFNI do not have it. Only SurfStat has it.

Comparing rate of biological change between groups

Test null hypothesis

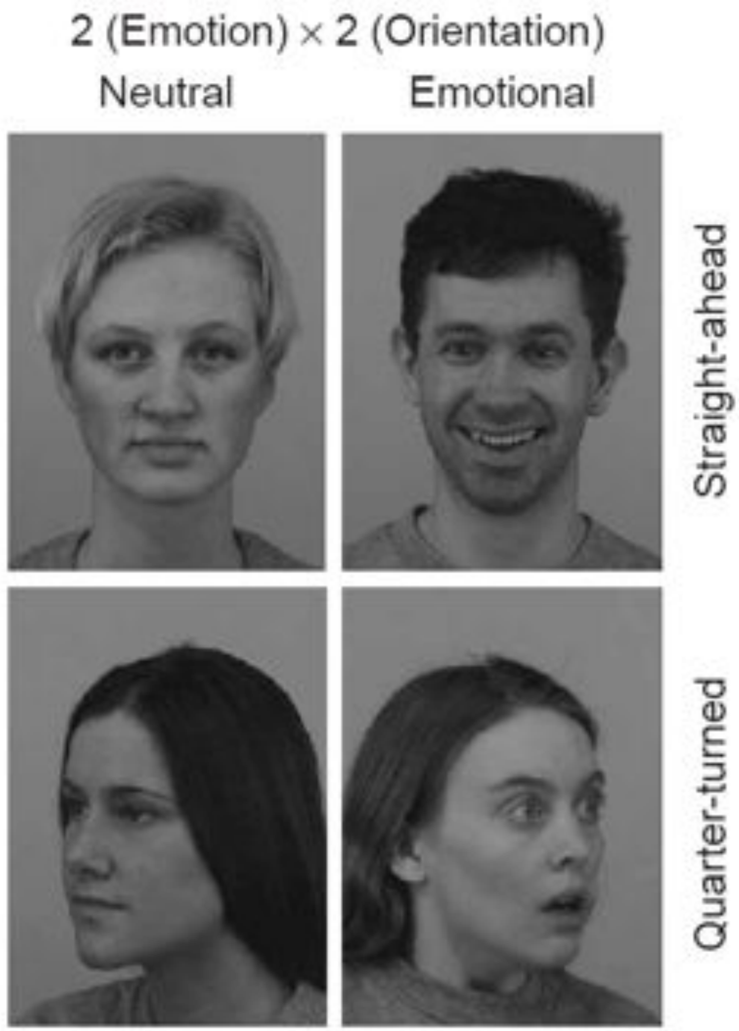
Ho: slopes in linear models are identical

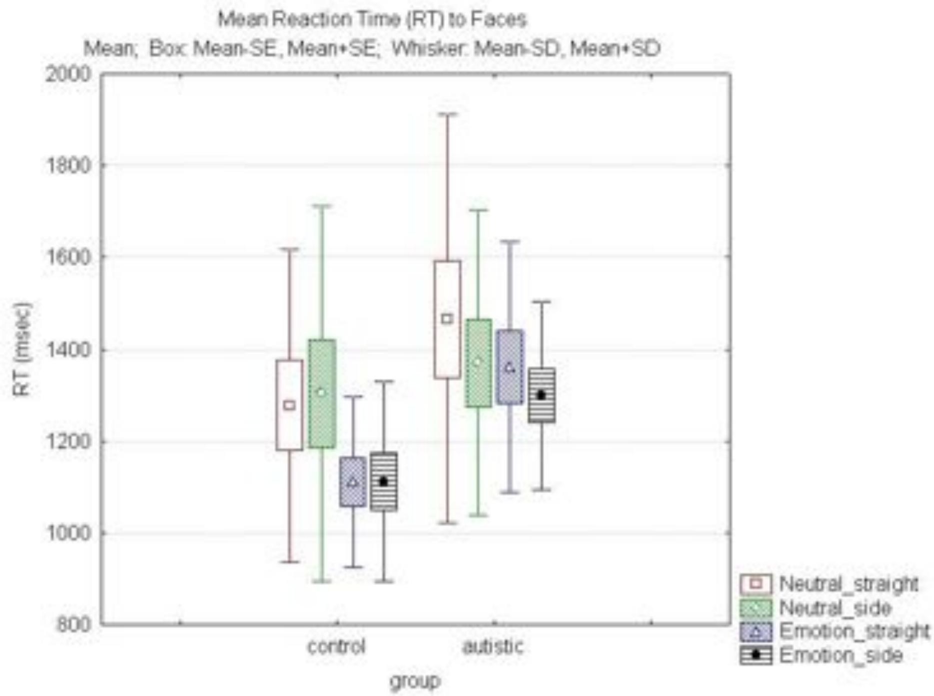


Question: what statistical procedure should we use?

Next question: what do you do with vector data?

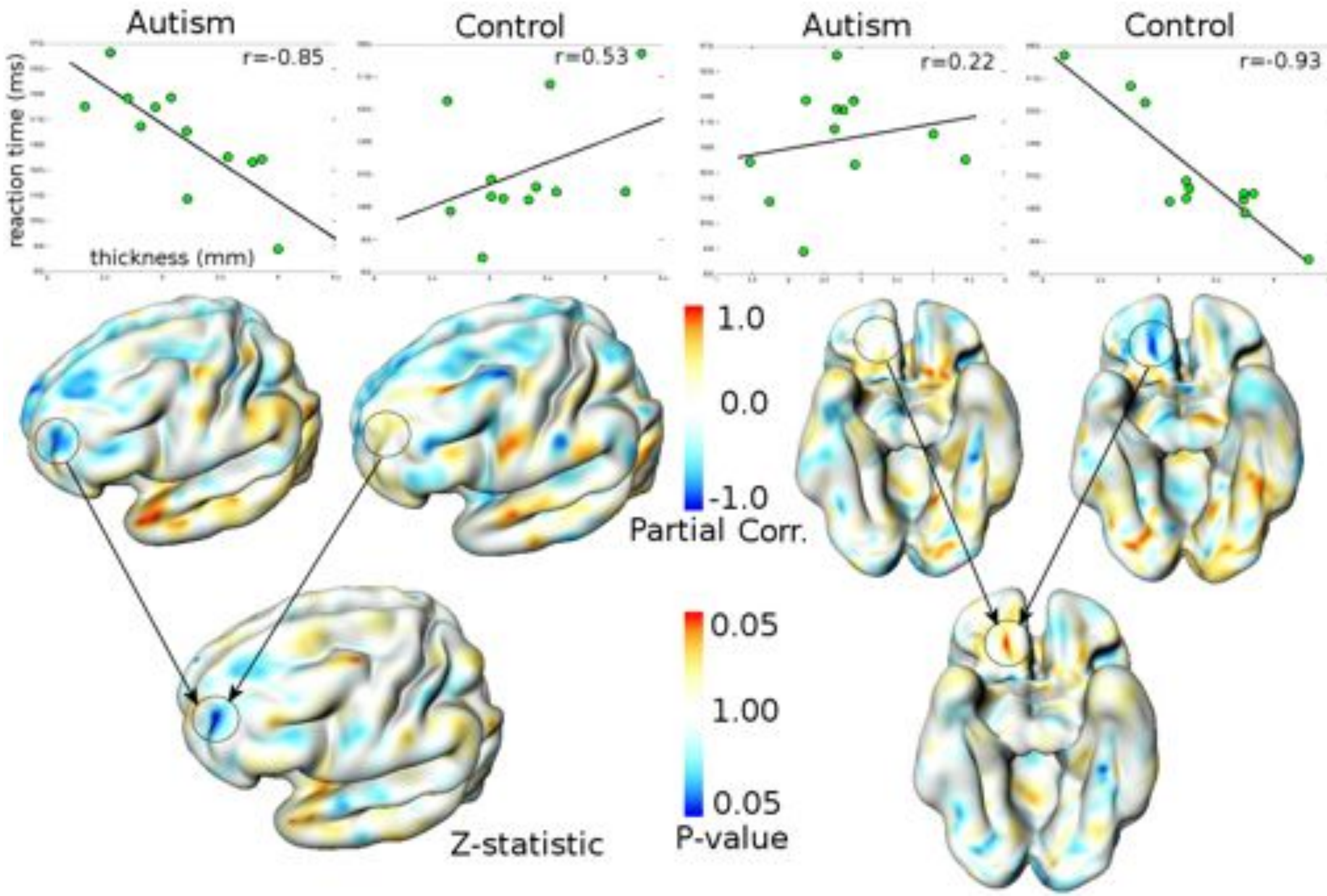
Facial emotion discrimination task response time
24 emotional faces, 16 neutral faces

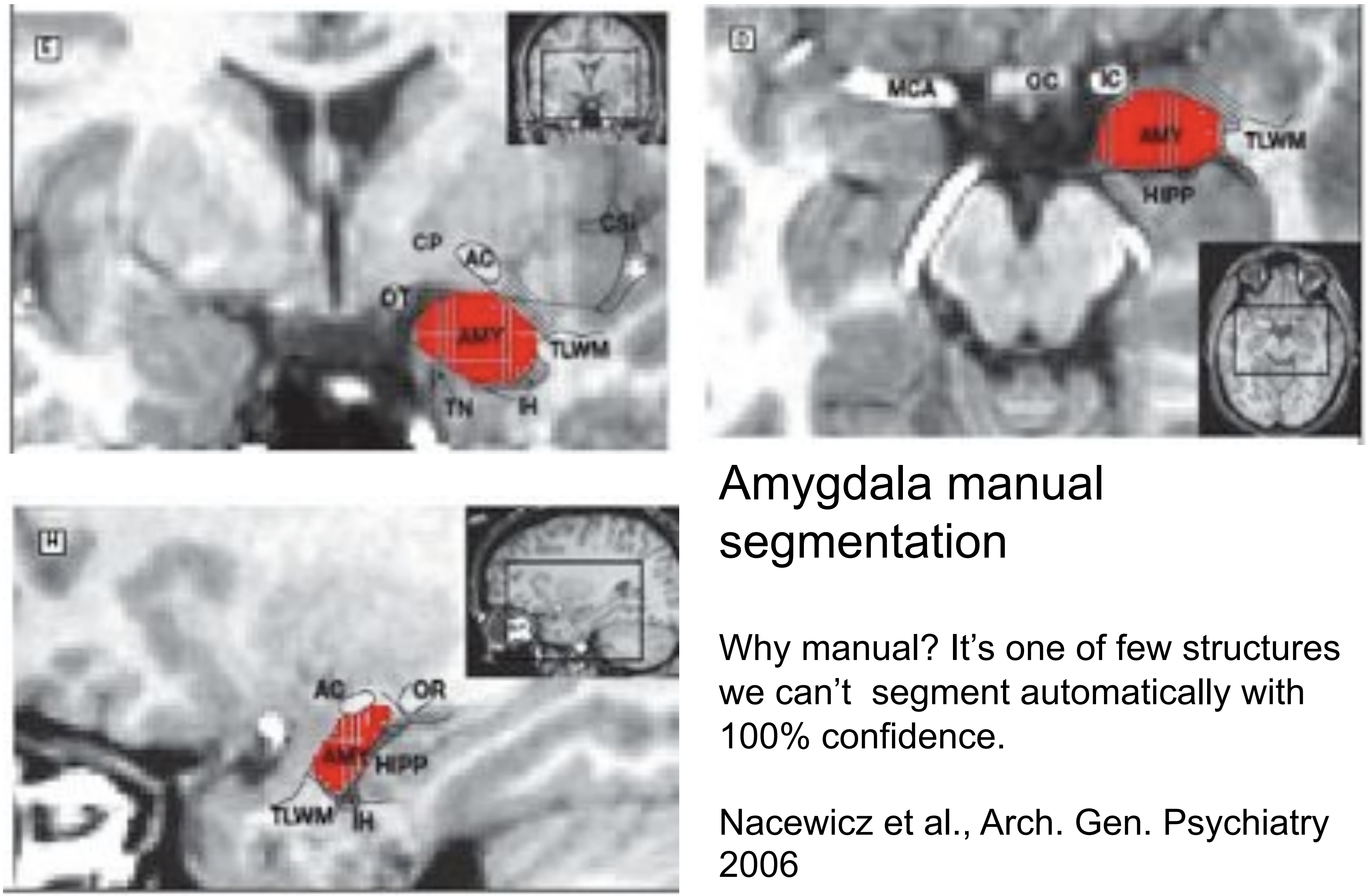




Dalton et al. (Nature Neuroscience 2005)

Correlating behavioral and imaging measures





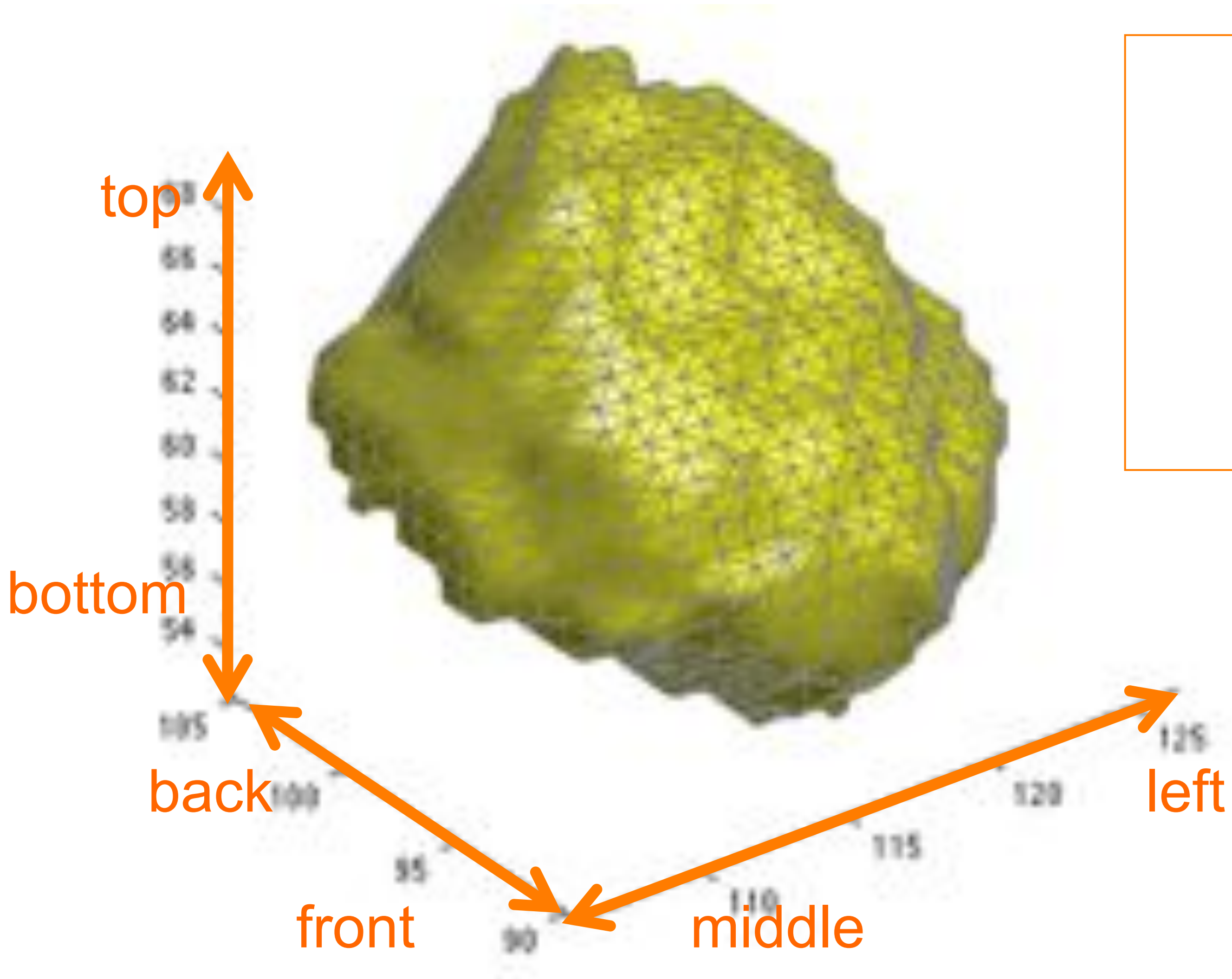
Amygdala manual segmentation

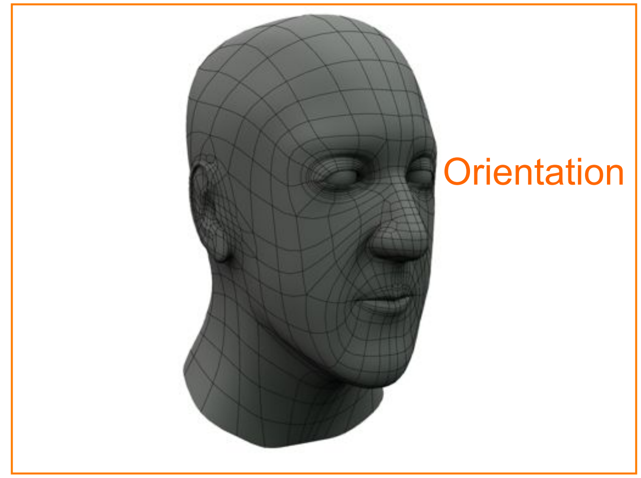
Why manual? It’s one of few structures we can’t segment automatically with 100% confidence.

Nacewicz et al., Arch. Gen. Psychiatry 2006

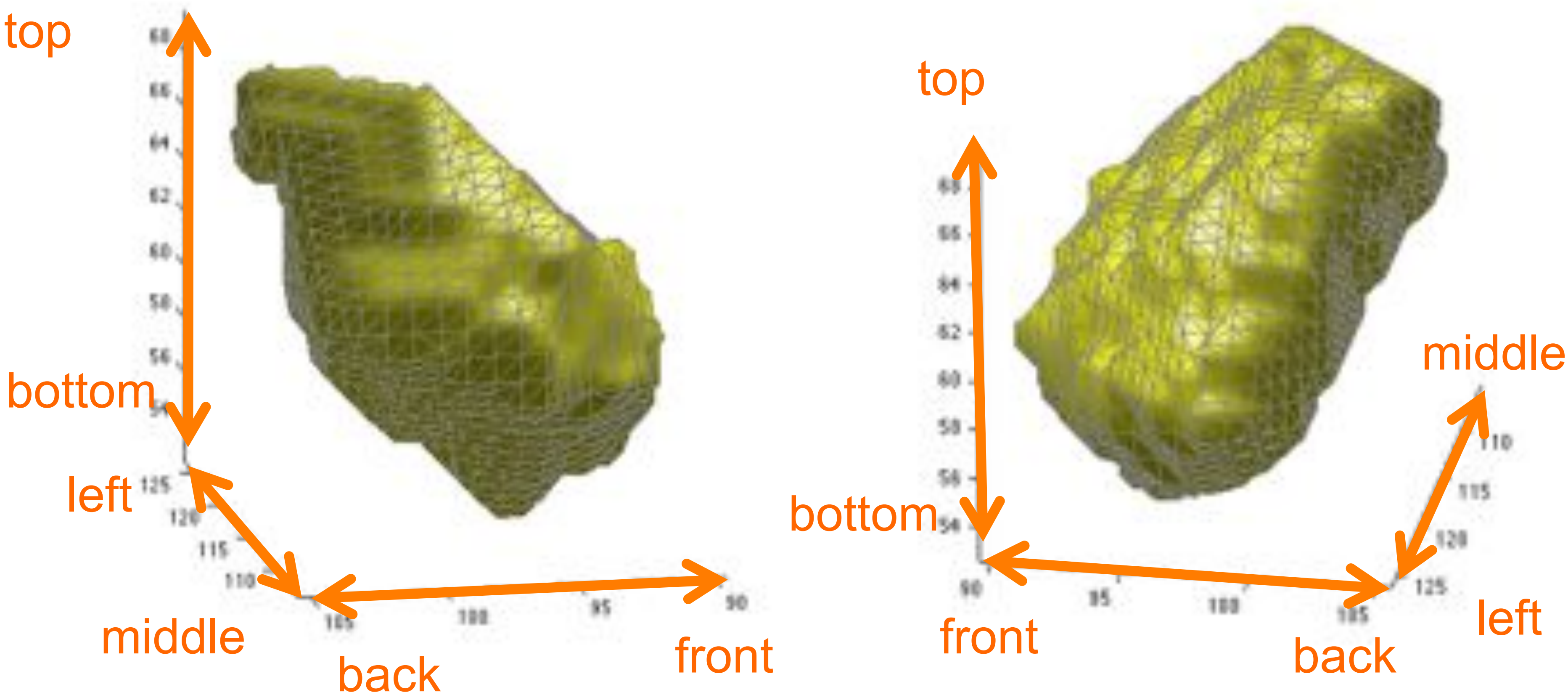
Chung *et al.,* NeuroImage 2010

2D surface model of left amygdala using marching cubes algorithm

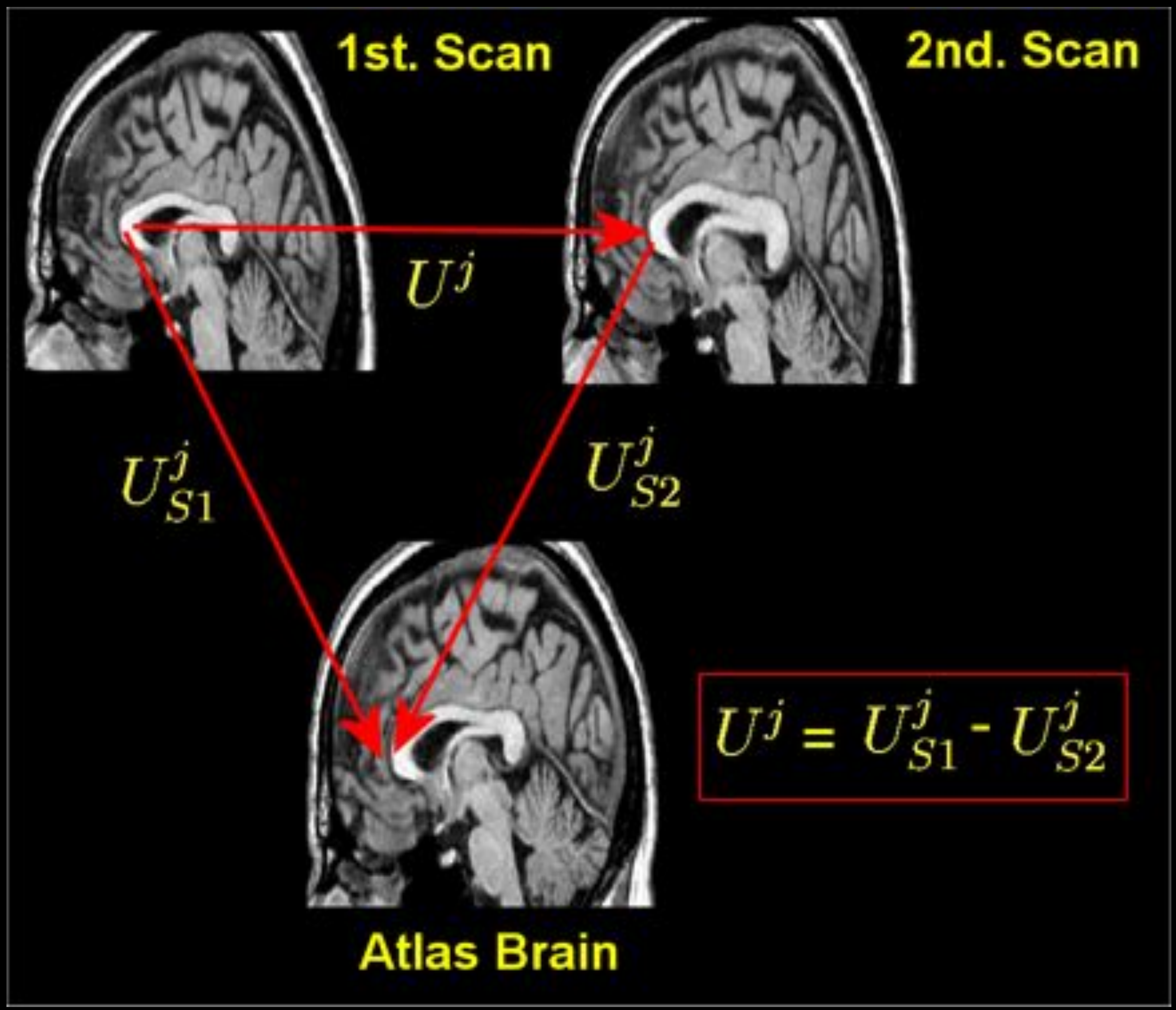




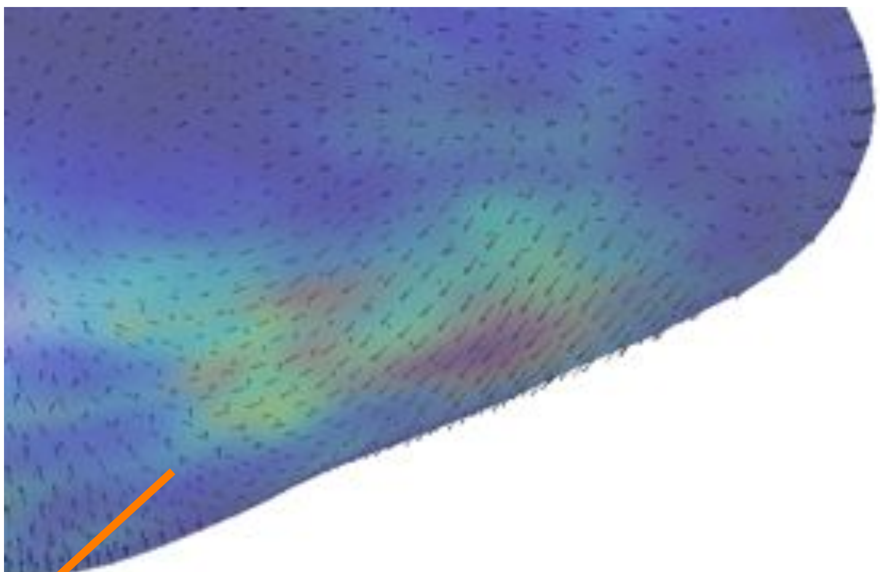
2D model of left amygdala of subject 001



Displacement vector fields in multiple images



Displacement vector field on a template



covariance
matrix

$$P_{n\times 3} = X_{n\times p}B_{p\times 3} + Z_{n\times r}G_{r\times 3} + U_{n\times 3}\Sigma_{3\times 3},$$

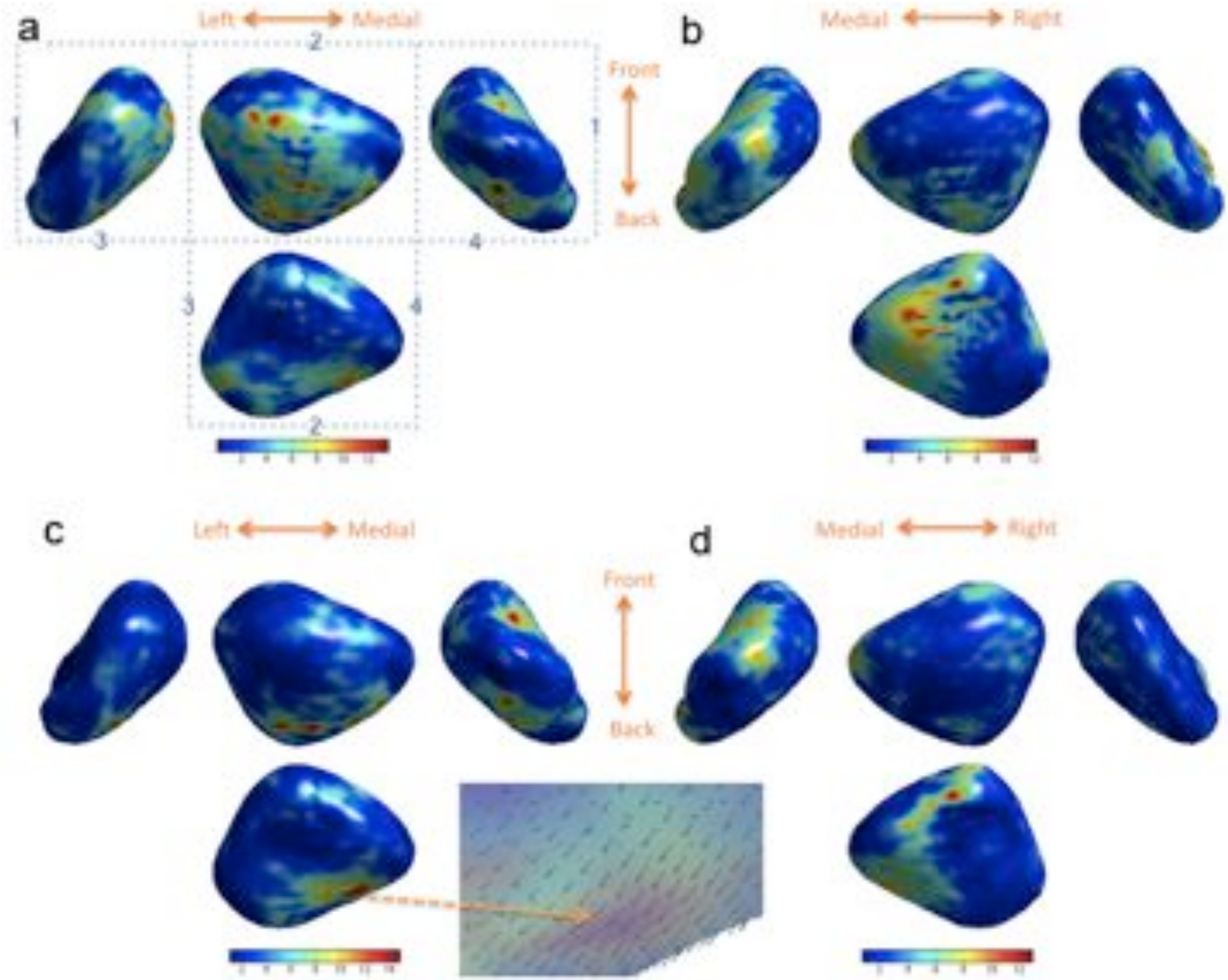
displacement
vector

variable of
interest

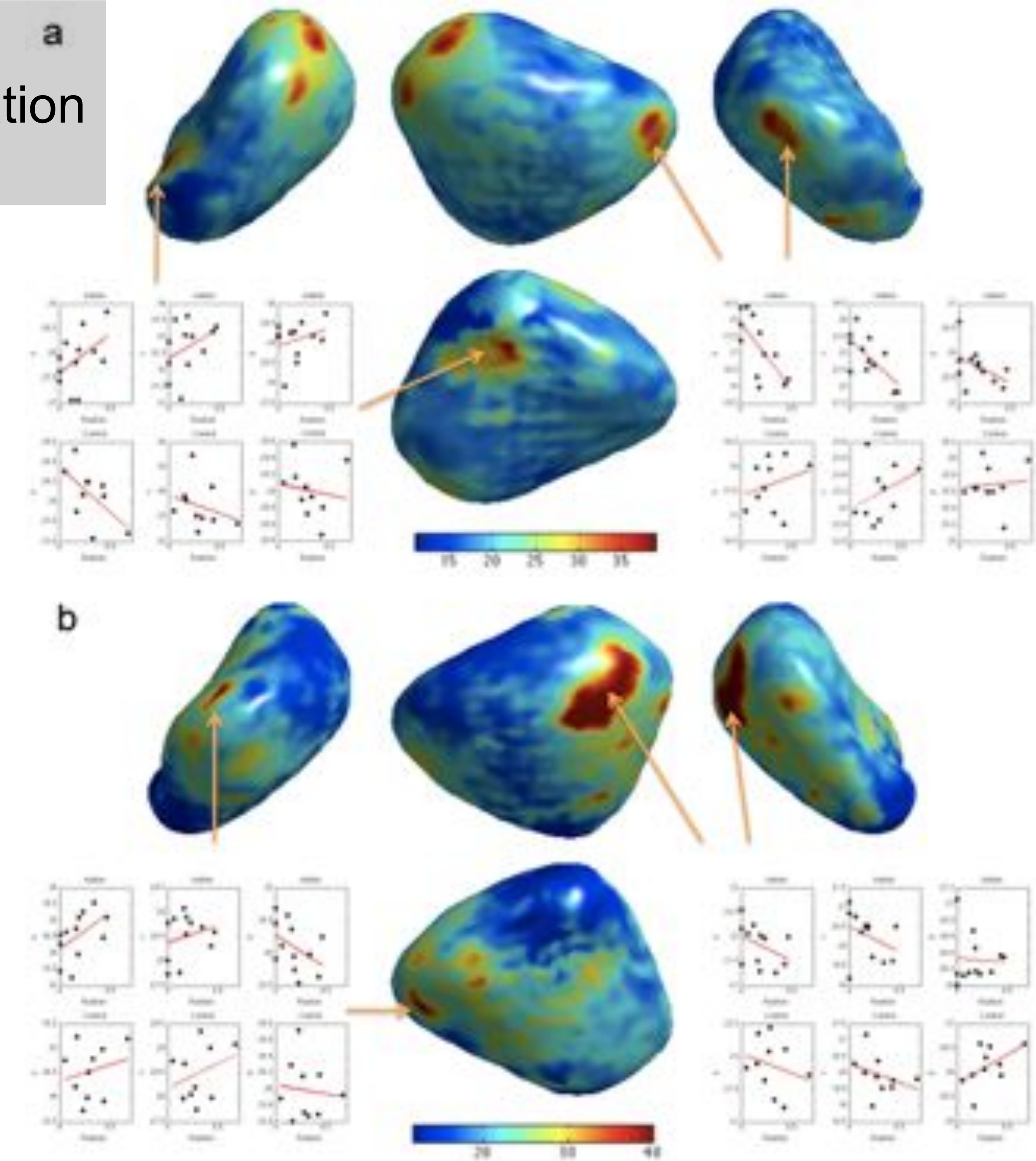
nuisance
covariates

noise

Difference between autism and controls



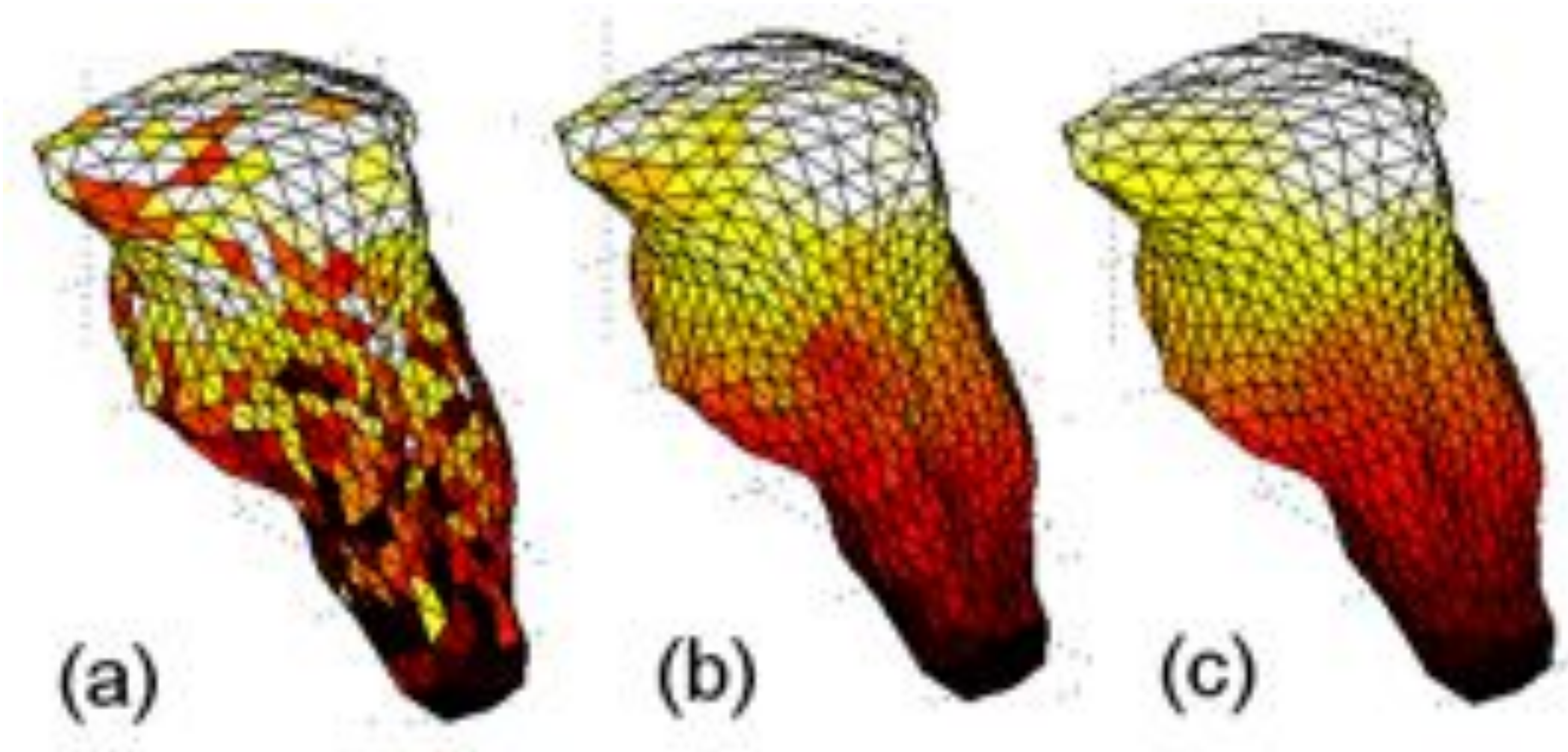
Interaction between shape and gaze fixation duration in autism



Finite Element Method

A numerical technique for discretizing a partial differential equation (PDE). PDE simplifies to a system of linear equations which can be solved by the usual least squares estimation.

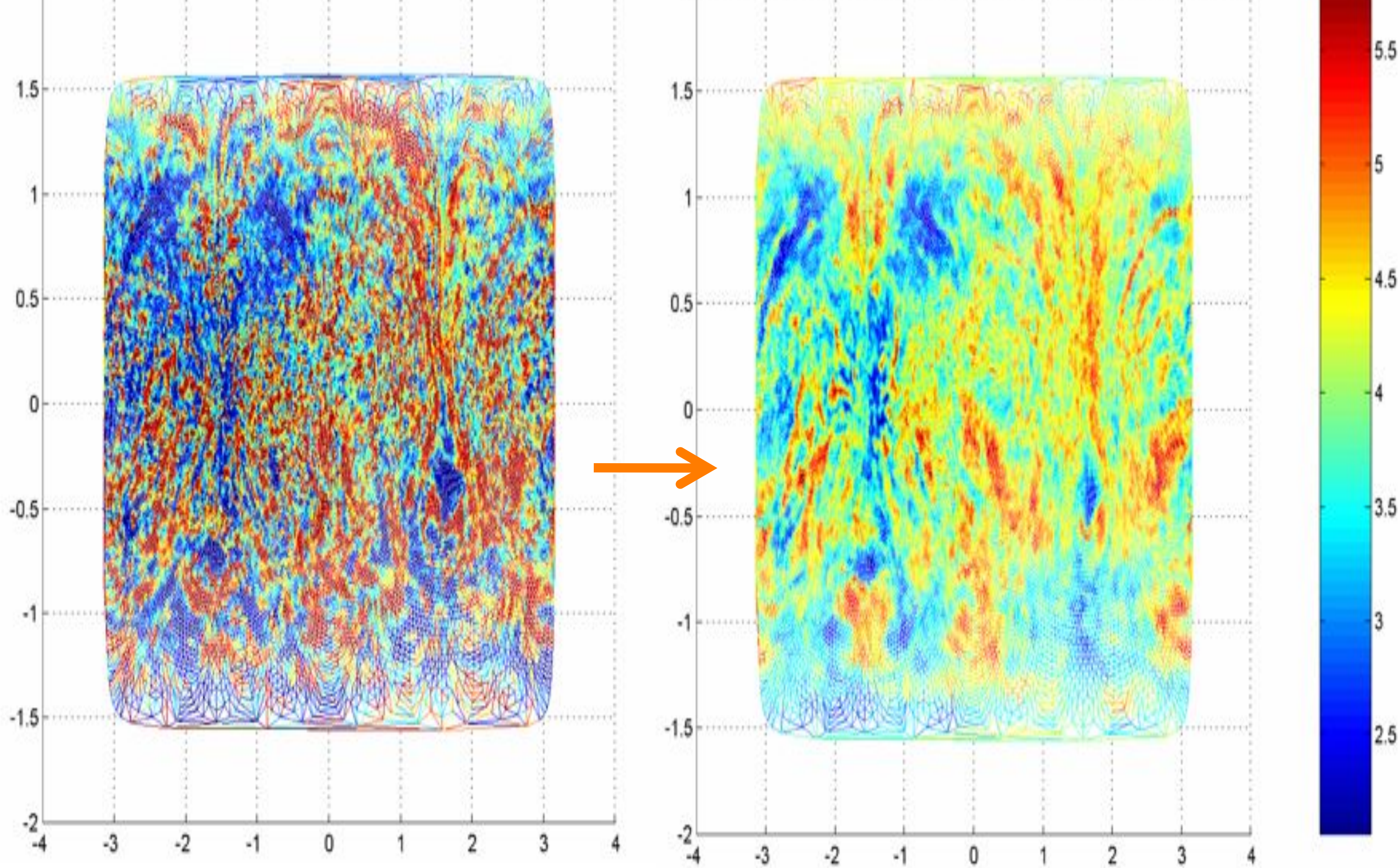
Smoothing along anatomical tissue boundary to increase SNR



Initial Signal **After 10 iterations** **After 20 iterations**

Diffusion smoothing: isotropic diffusion equation

Smoothed cortical thickness

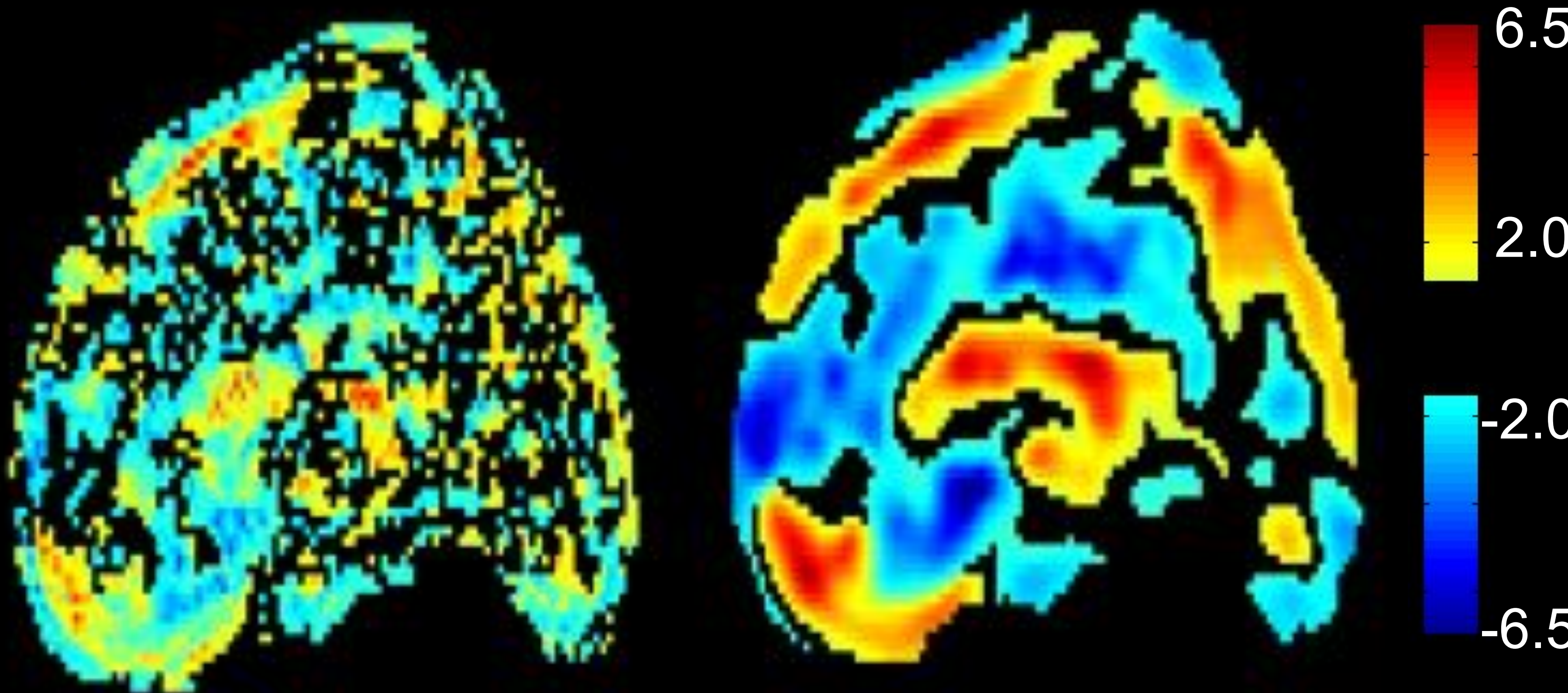


Thickness measure

Signal enhancement

Why do we smooth?

t-statistic map of Jacobian determinant change (volume change) for 28 normal subjects from age 12 to age 16.



10mm FWHM Gaussian kernel smoothing

Diffusion Smoothing

It can be shown that the convoluted signal $F(\mathbf{x},t) = F^*(\mathbf{x},\sqrt{2t})$ is the solution of a diffusion equation

$$\frac{\partial F}{\partial t} = \Delta F,\ F(\mathbf{x},0) = f(\mathbf{x})$$

where the n-dimensional Laplacian is given by $\Delta = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}$. The amount of smoothing is determined by the full width at the half maximum (FWHM) of Gaussian kernel

$$\mathrm{FWHM} = 4(\ln 2)^{1/2}\sqrt{t} = 2(2\ln 2)^{1/2}h$$

Since the cortical surface in non-Euclidean, the above Laplacian is not well defined on the cortical surface. The generalization of the Laplacian to an arbitrary curved surface is called the *Laplace-Beltrami operator* and it is defined in terms of the Riemannian metric tensors. For the Riemannian metric $ds^2 = \sum_{i,j=1}^n g_{ij}\ du^i du^j$, the Laplace-Beltrami operator is given by

$$\Delta F = \frac{1}{|g|^{1/2}} \sum_{i,j=1}^{n} \frac{\partial}{\partial u^i}\Big(|g|^{1/2} g^{ij} \frac{\partial F}{\partial u^j}\Big)$$

where $g^{-1} = (g^{ij})$ and $|g| = \det(g_{ij})$.

Finite Element Method

The ASP algorithm (MacDonald, *et al.*, 2001) is used to extract the outer cortical surfaces each consisting of 81,920 triangles from MR scans. At this surface sampling rate, the average intervertex distance is 3-4 mm. In order to estimate the Laplace-Beltrami operator on a triangulated cortical surface, we use the *finite element method* (FEM) (Chung, 2001). Let $F(\mathbf{p}_i)$ be the signal on the i-th node $\mathbf{p}_i$ in the triangulation. If $\mathbf{p}_1,\ldots,\mathbf{p}_m$ are m-neighboring nodes around $\mathbf{p}=\mathbf{p}_0$, the Laplace-Beltrami operator at $\mathbf{p}$ is estimated by

$$\widehat{\Delta F}(\mathbf{p}) = \sum_{i=1}^{m} w_i(F(\mathbf{p}_i) - F(\mathbf{p}))$$

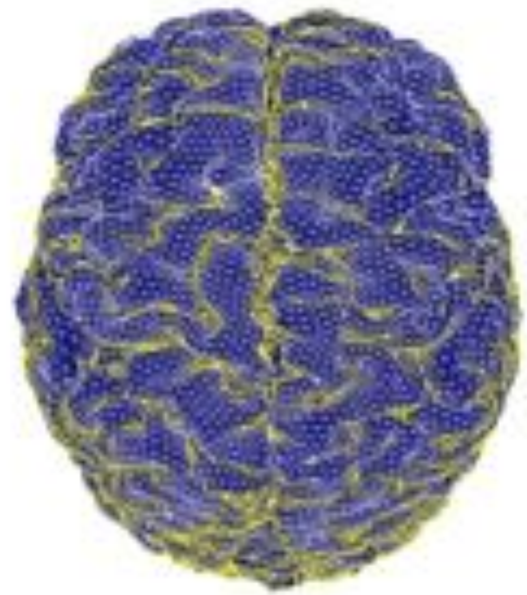
with the weights

$$w_i = (\cot\theta_i + \cot\phi_i)/|T|$$

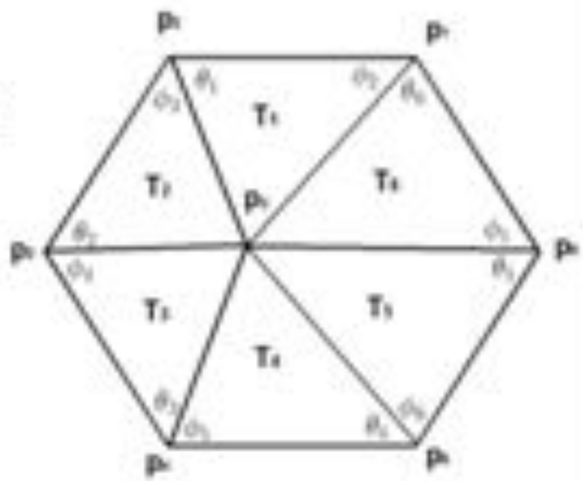
where θ_i and ϕ_i are the two angles opposite to the edge $\mathbf{p}_i$ - $\mathbf{p}$ in triangles and $|T|$ is the sum of the areas of m-incident triangles at $\mathbf{p}$. Then the diffusion equation is solved via the *finite difference scheme*:

$$F(\mathbf{p}, t_{n+1}) = F(\mathbf{p}, t_n) + (t_{n+1} - t_n)\widehat{\Delta F}(\mathbf{p}, t_n)$$

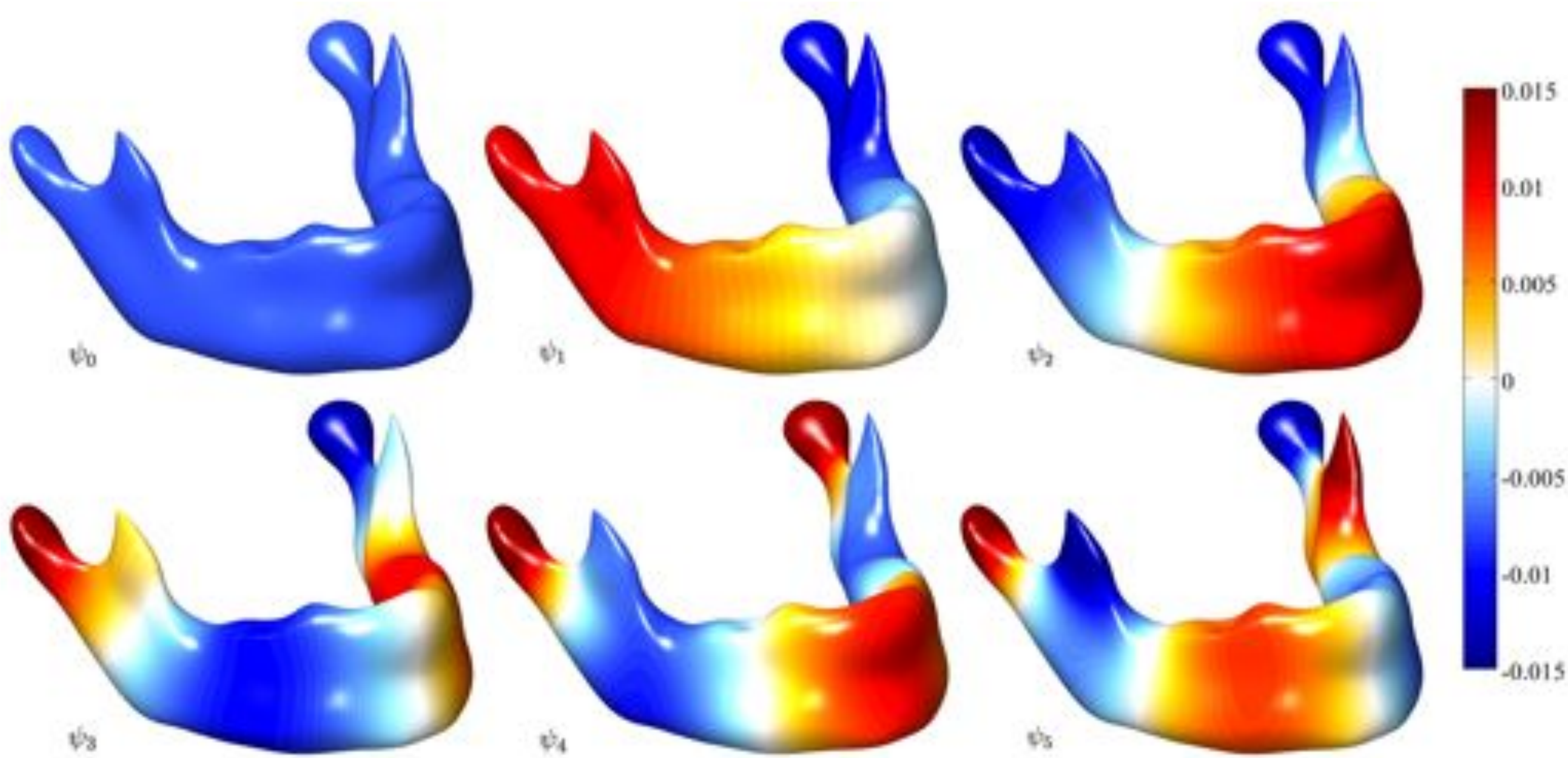
with the initial condition $F(\mathbf{p}_i, t_0)=f(\mathbf{p}_i)$. After N-iterations, the diffused signal is locally equivalent to Gaussian kernel smoothing with FWHM = $4(\ln 2)^{1/2}N^{1/2}(t_N-t_0)^{1/2}$.



A typical triangular mesh of the outer cortical surface consisting of 81,920 triangles and 40,962 vertices.

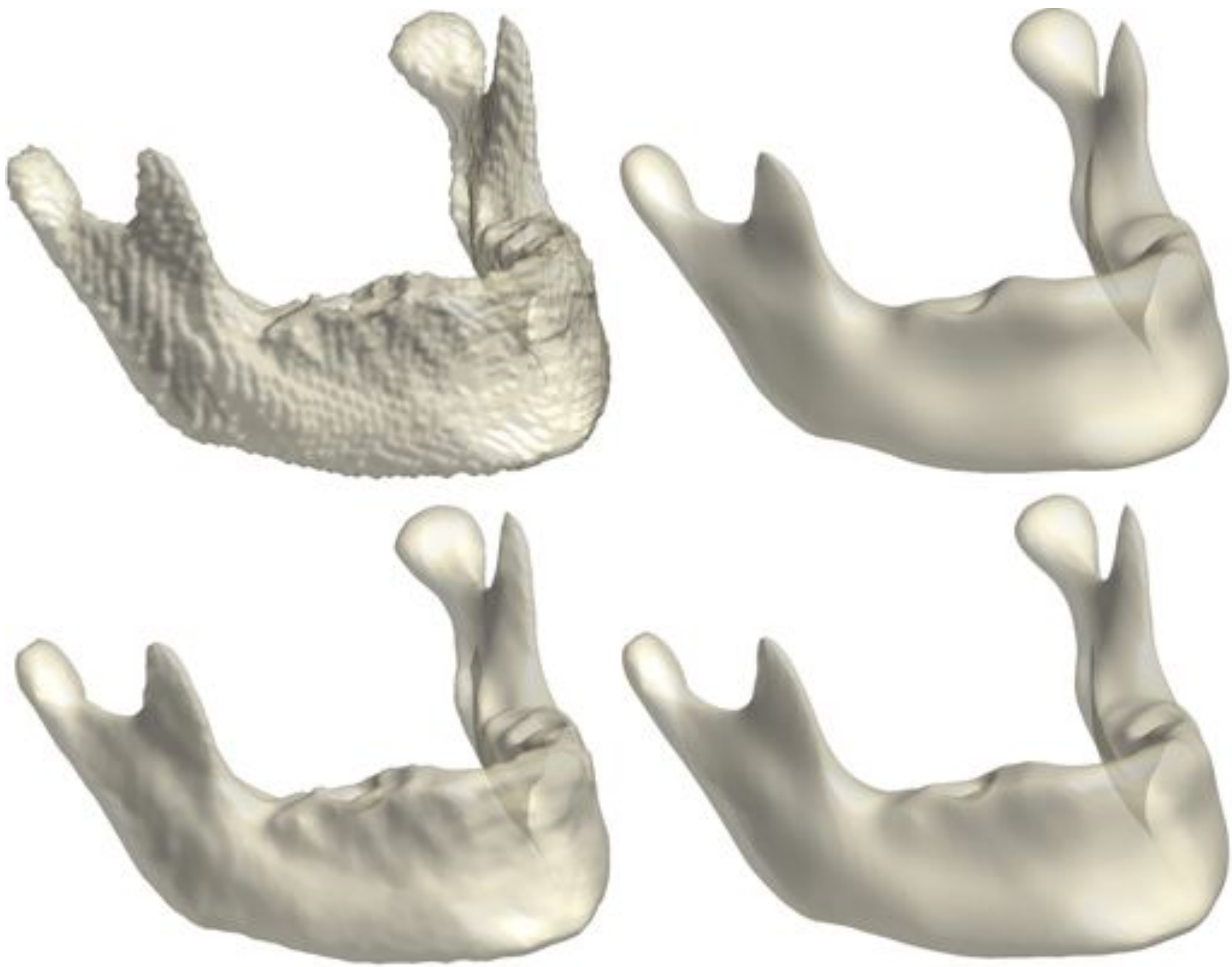


Eigenvalues of Laplace-Beltrami operator $$\Delta f = \lambda f$$

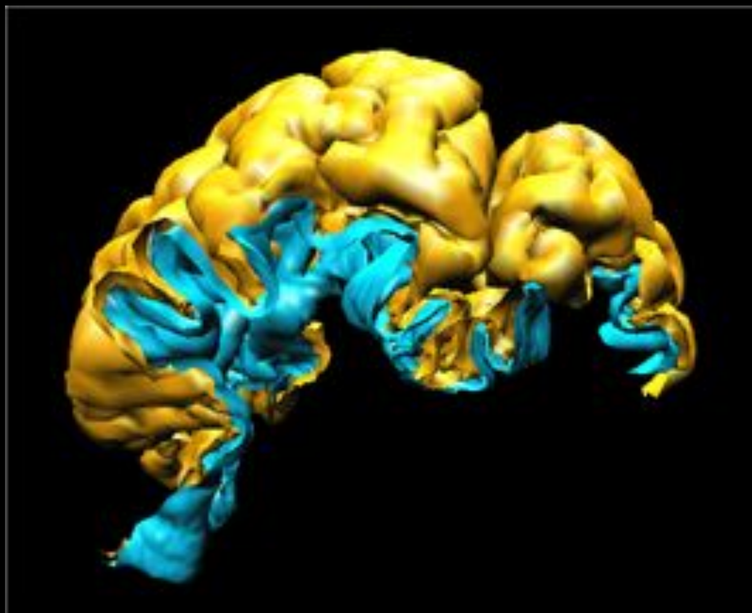


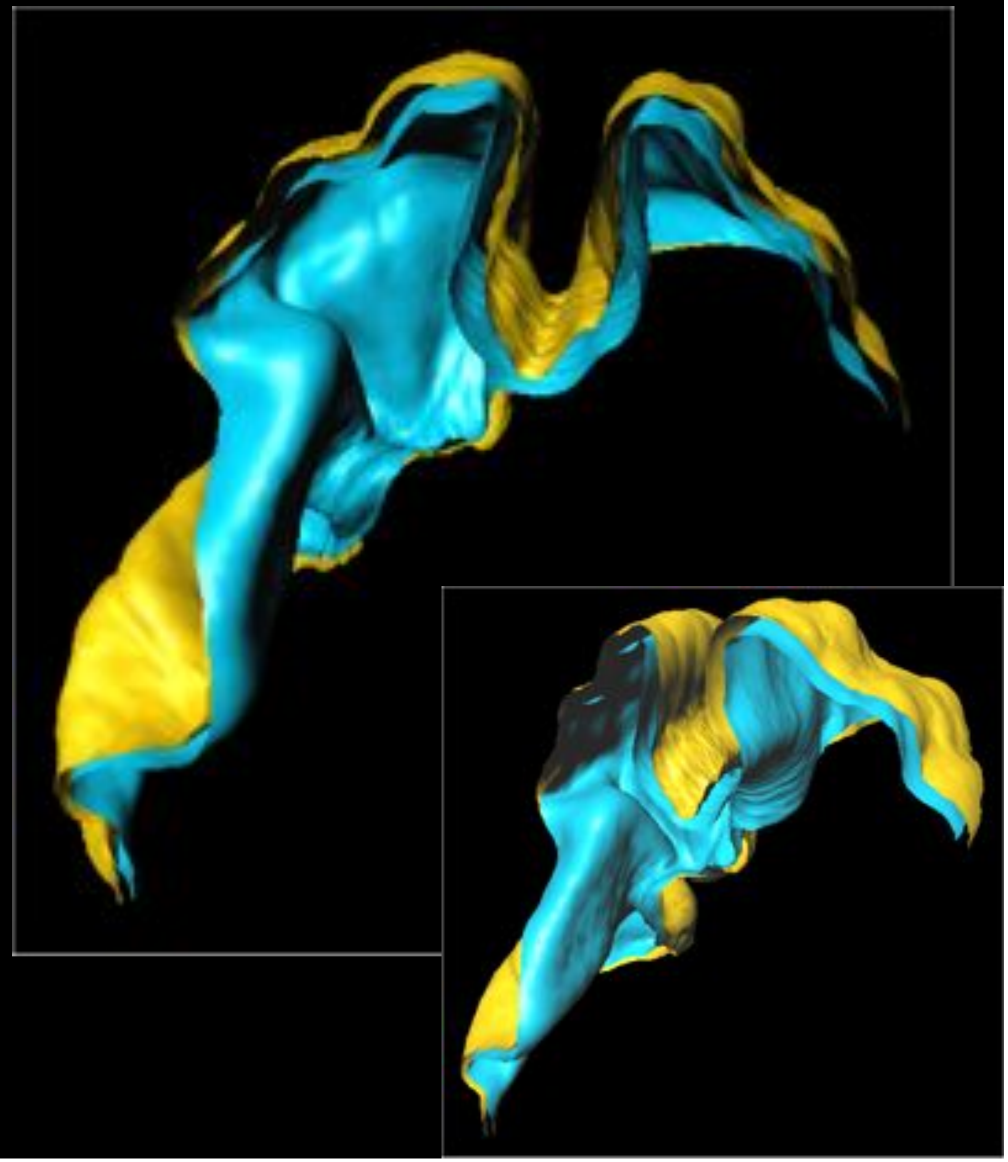
Seo *et al.*, MICCAI 2010

Heat kernel smoothing

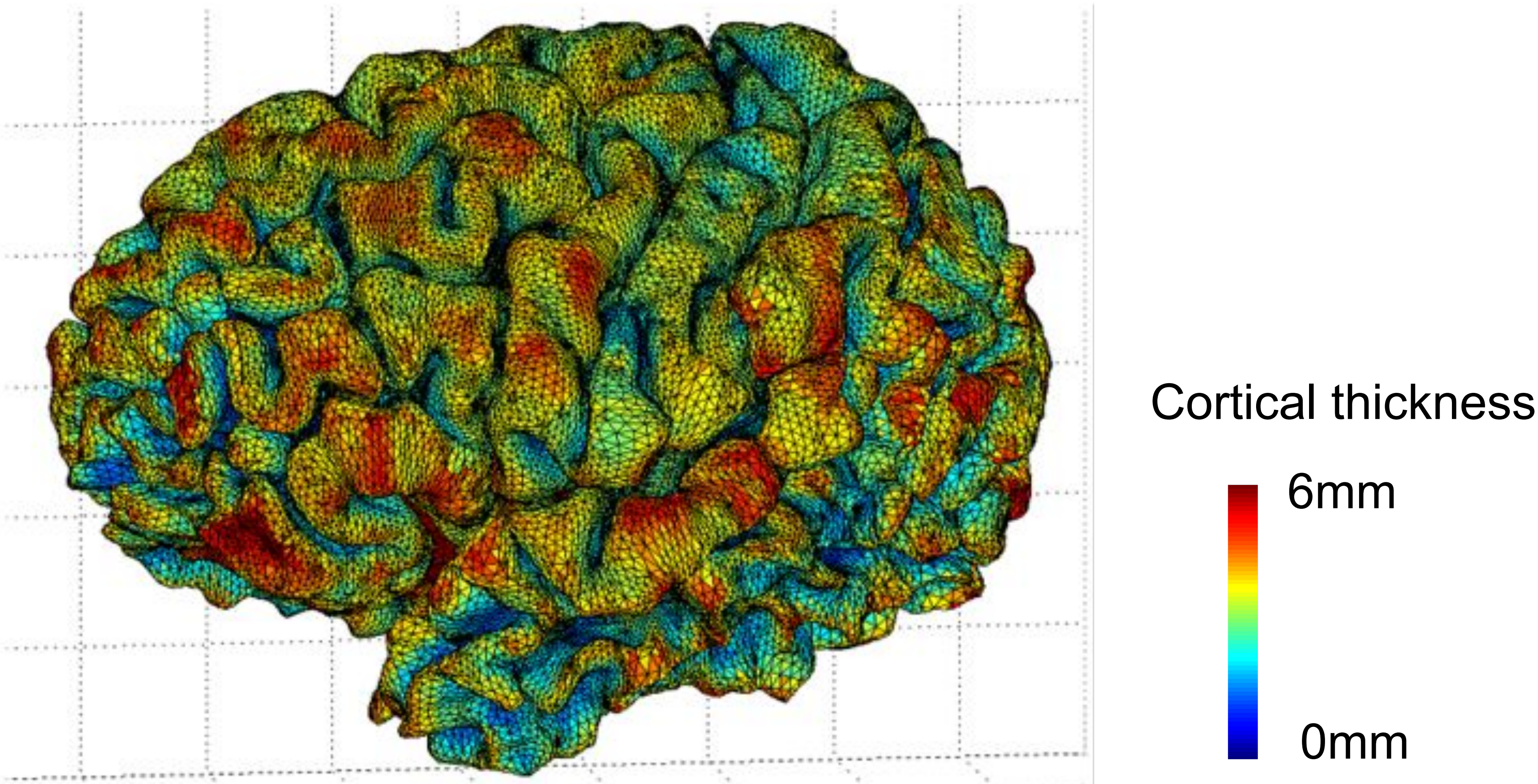


cortical thickness





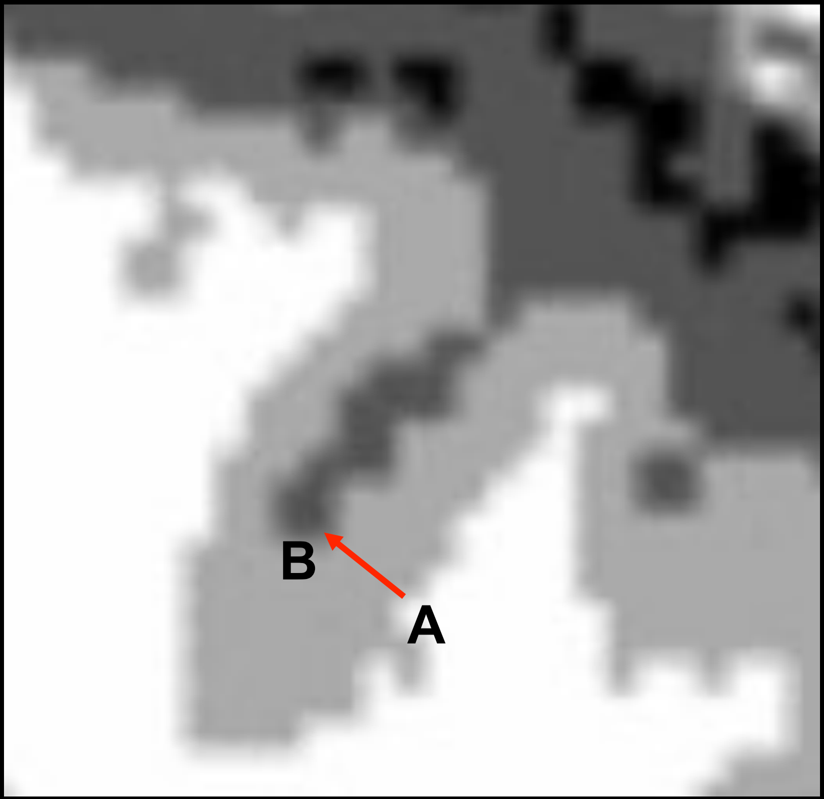
Cortical thickness = most widely used cortical measure



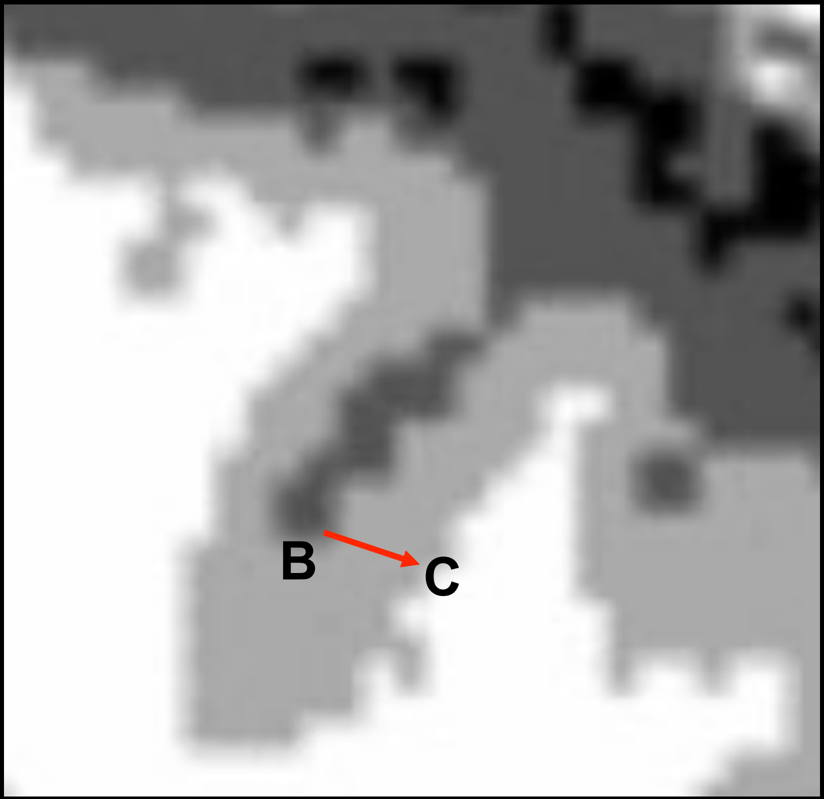
Chung *et al.*, NeuroImage 2003

Cortical thickness

Inconsistent mathematical definitions. There are at least five different methods of measuring distance between tissue boundaries.



orthogonal projection from A to B



orthogonal projection from B to C

Laplace equation method for defining cortical thickness

$$\Delta f = 0$$

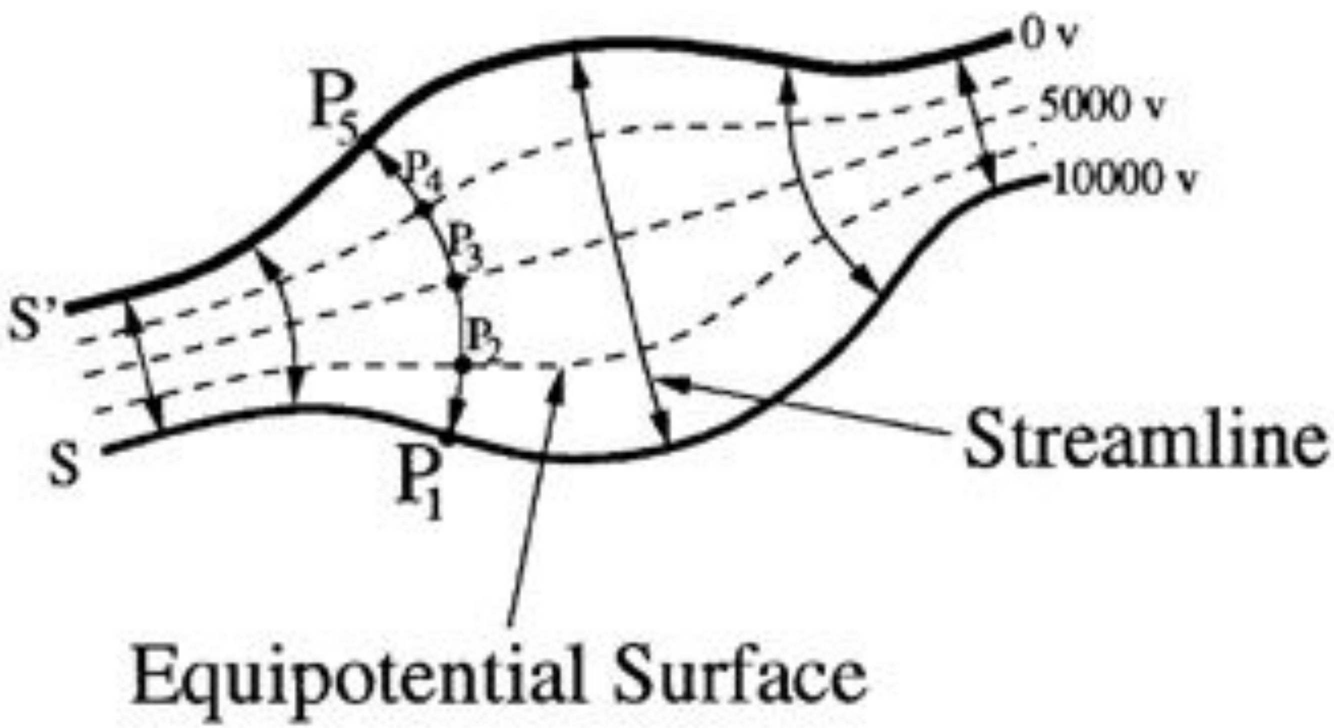


Figure 4.
Two-dimensional example of Laplace's method. Laplace's equation is solved between S and S′, which have predetermined boundary conditions of 10,000 V and 0 V, respectively. Three examples of resulting intermediate equipotential surfaces are indicated for 2,500 V, and 5,000 V, and 7,500 V. Field lines connecting S to S′ are defined as being everywhere orthogonal to all equipotential surfaces, as exemplified by the line connecting P to P′.

Jones *et al*. HBM. 2000

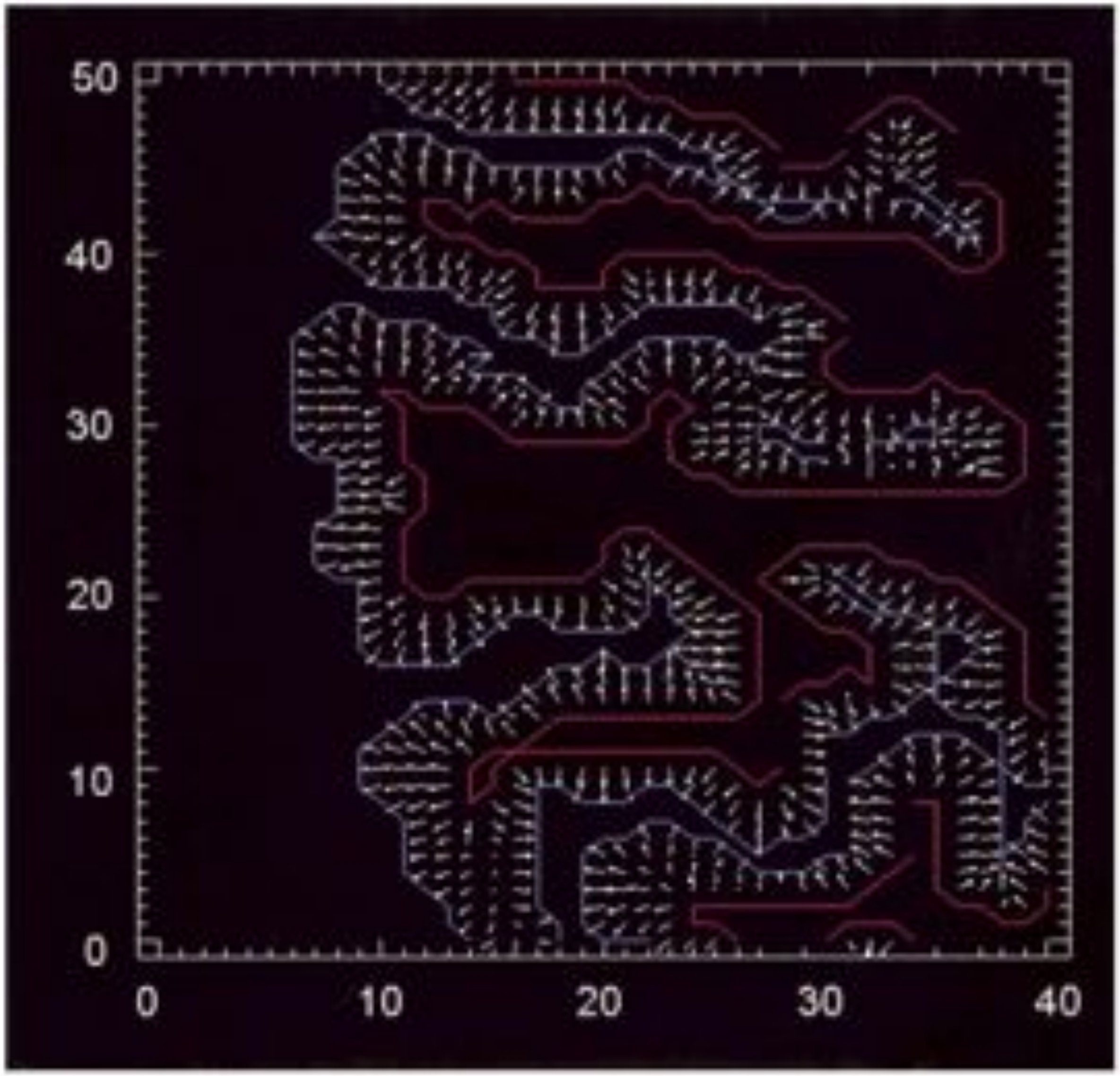
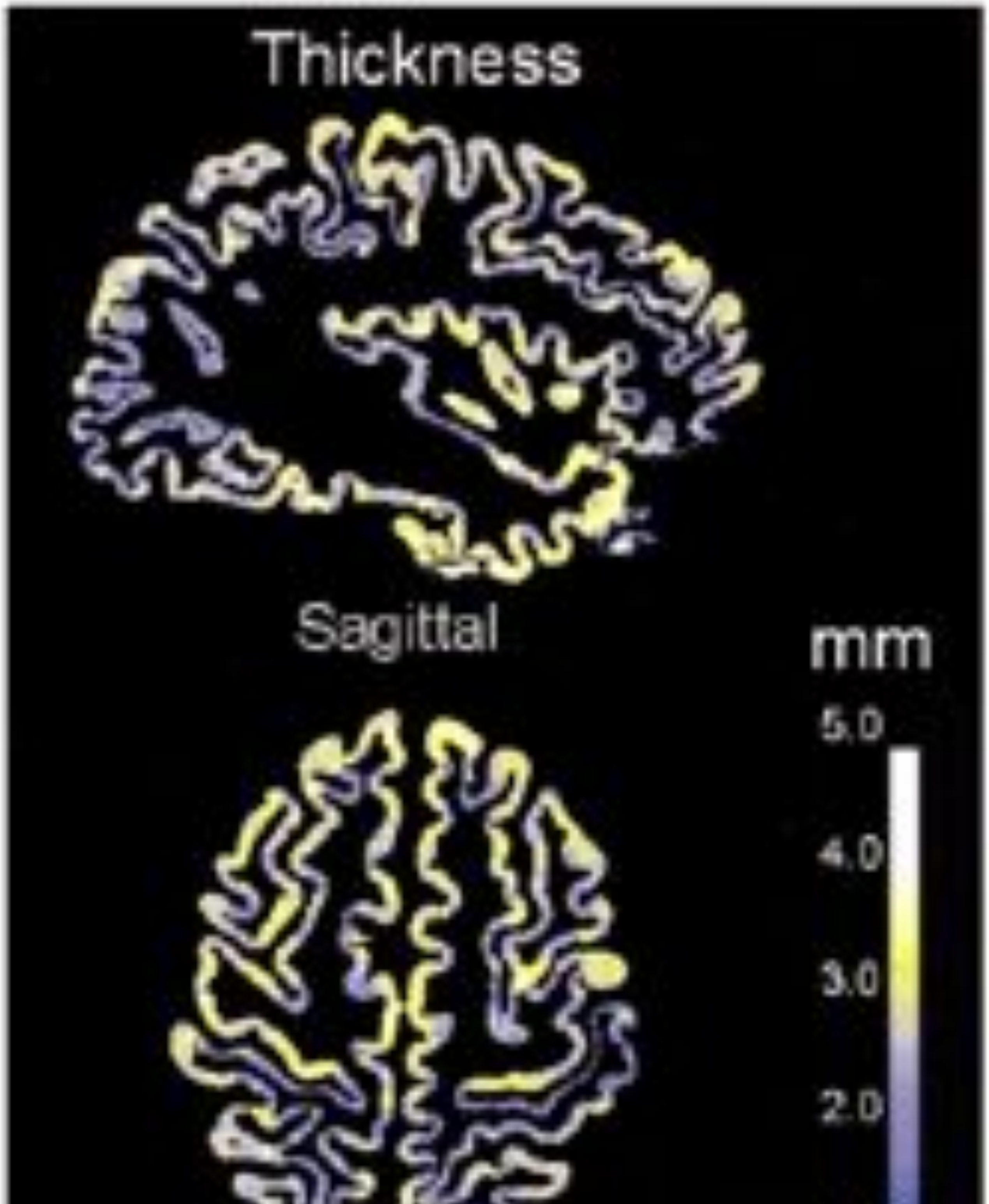
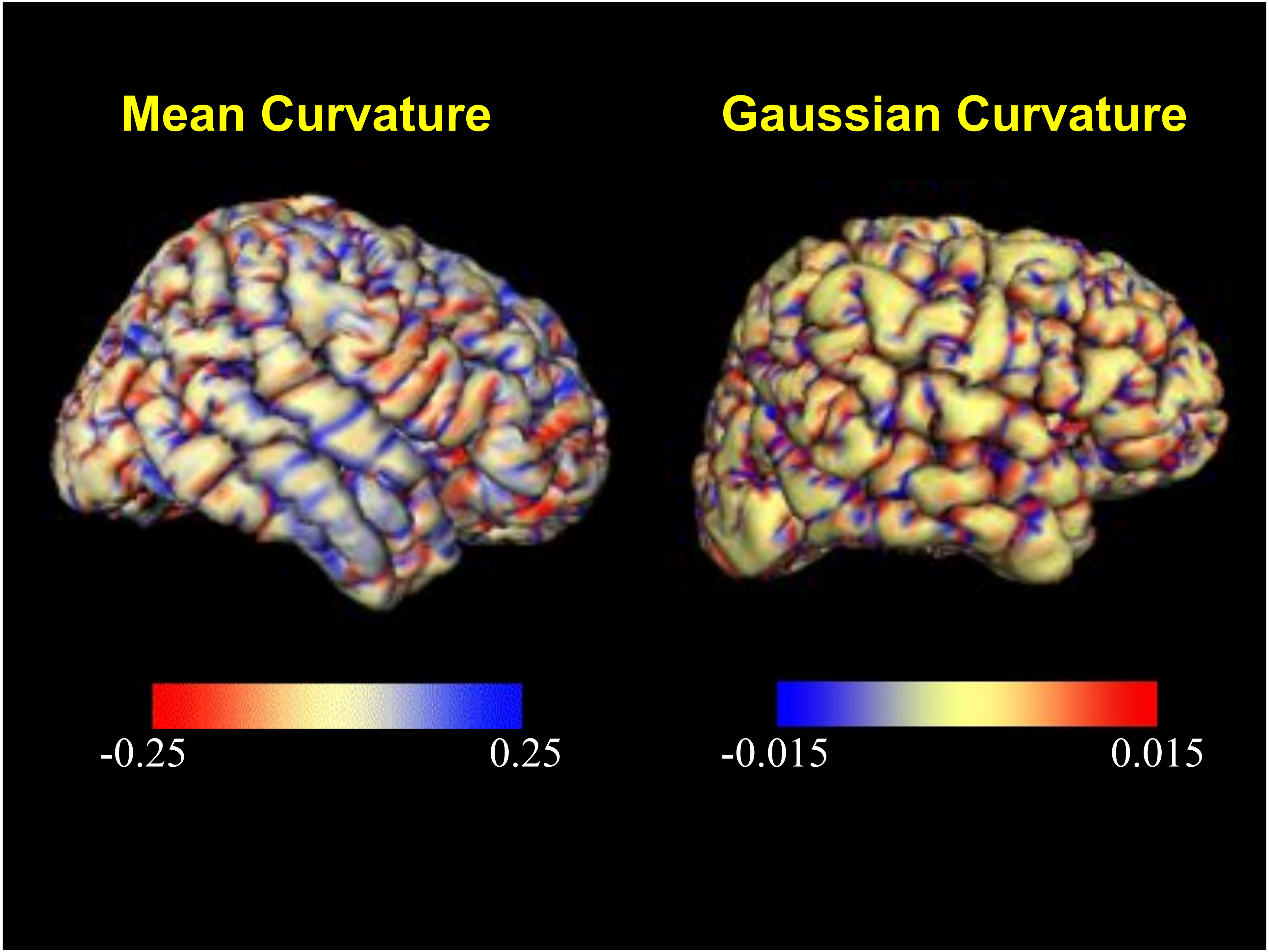


Figure 7.
Close-up example of gradients of Laplace's solution in an axial plane from real data. The blue line represents the gray-white junction, and the red line represents the gray-CSF junction. The small arrows are projections of the gradient vectors in the axial plane. These arrows are tangent to the streamlines connecting the two surfaces. Arrows appear short when they are projecting predominantly out of the axial plane [e.g., at position (15,5)]. The gradients are insensitive to small segmentation errors as seen but the sulcal discontinuity at position (30,29).

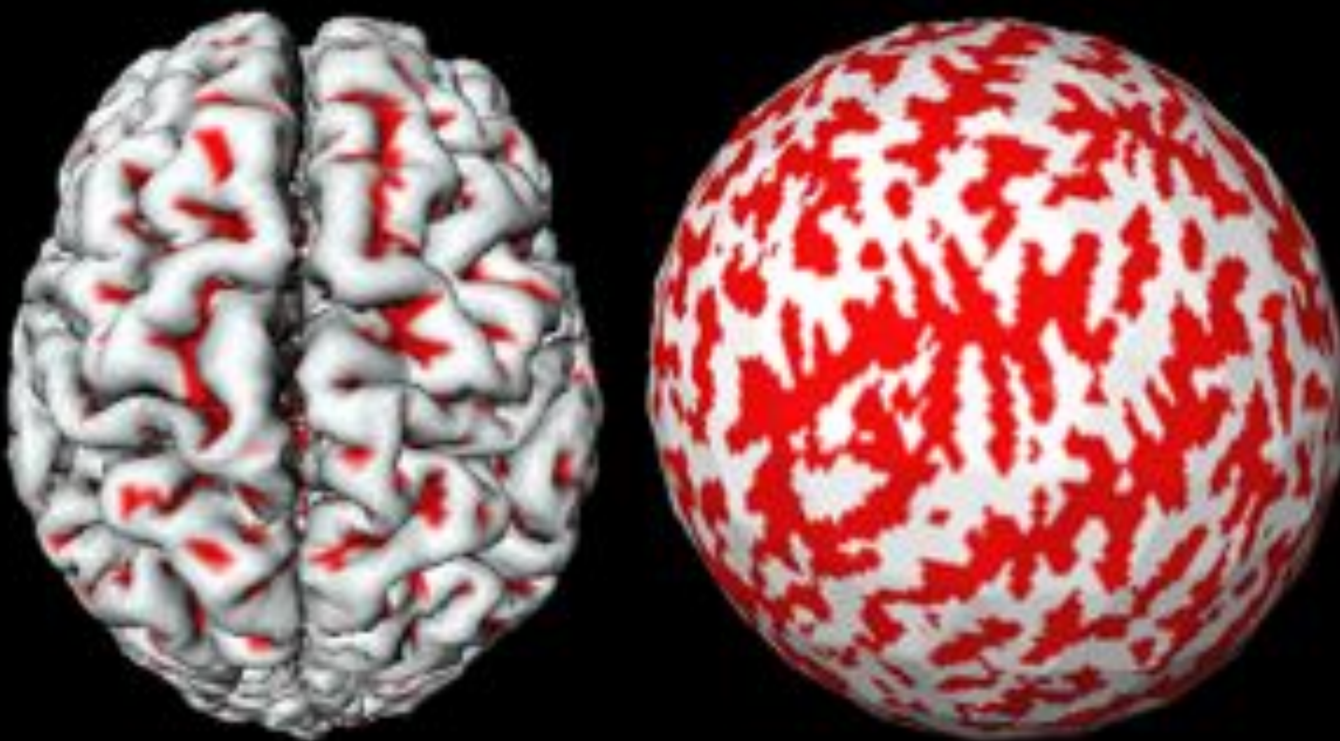

Thickness
Sagittal
mm
5.0
4.0
3.0
2.0

Geometric computation

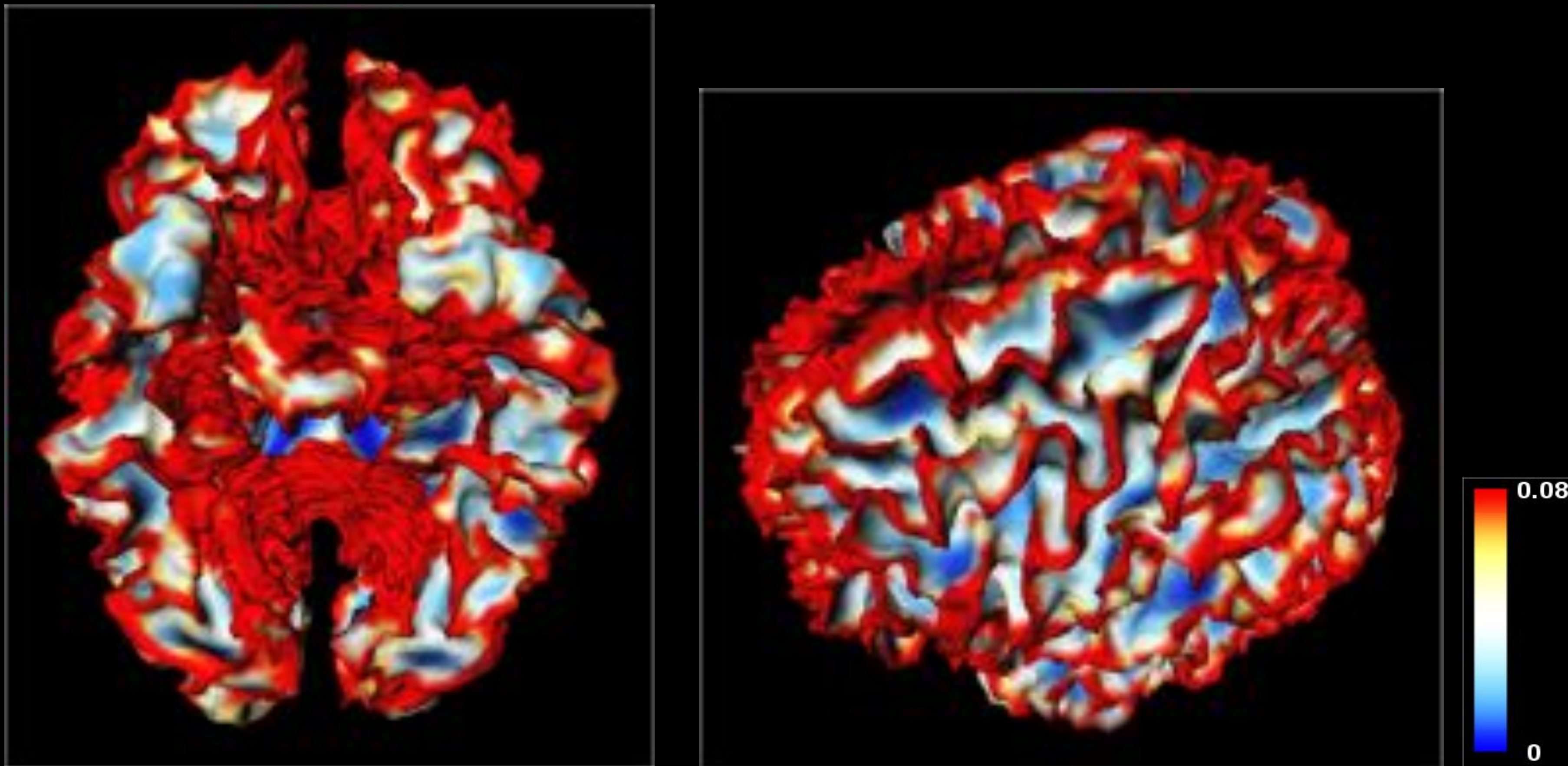
Geometric quantities such as curvatures, length, area, volume have been often used in characterizing brain shape.


Mean Curvature
Gaussian Curvature
-0.25
0.25
-0.015
0.015

Mean curvature can be used to quantify sulcal pattern



Application of curvature measure: Tensor-based morphometry (TBM)



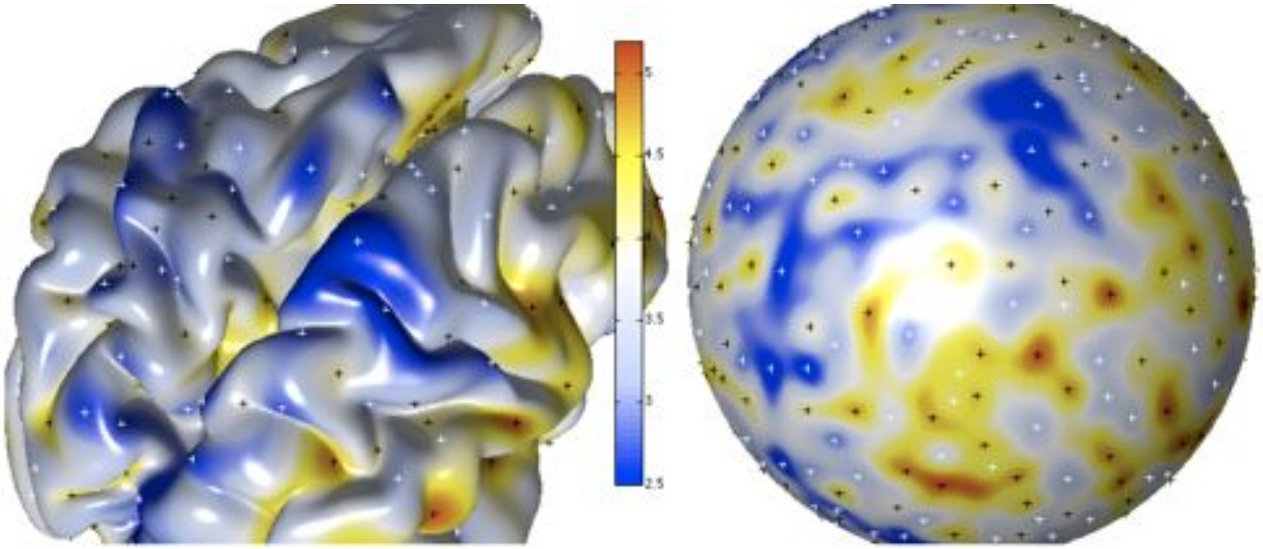
Thin-plate spline energy can be used to measure the curvature of the surface. Between ages 12 and 16, it increases both locally and globally.

Chung *et al.*, CVPR 2003

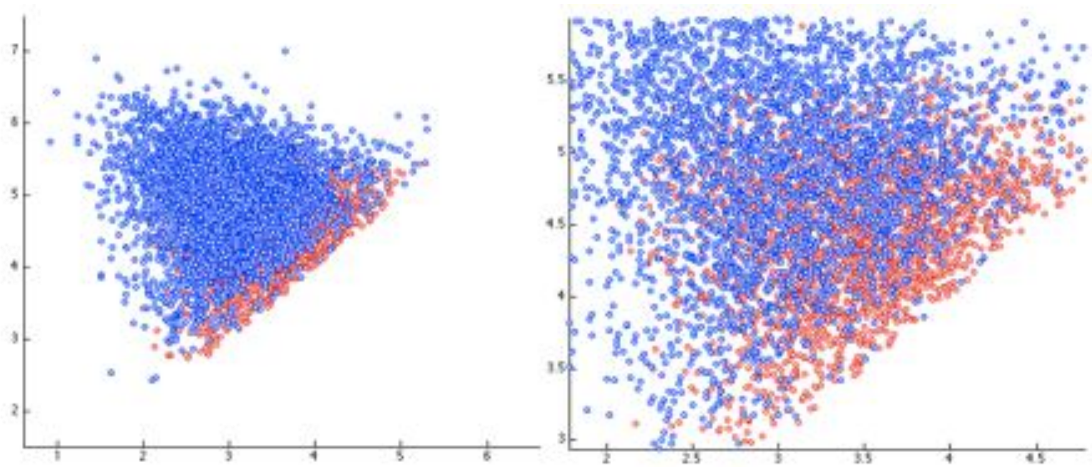
Topological computation

Topological properties are invariant under shape deformation. So topological invariants can be used to characterize an object of interest.

Topological metric obtained from cortical thickness



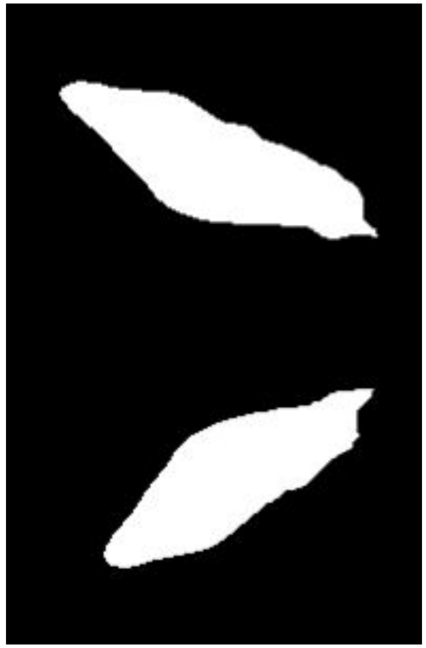
Topological data reduction



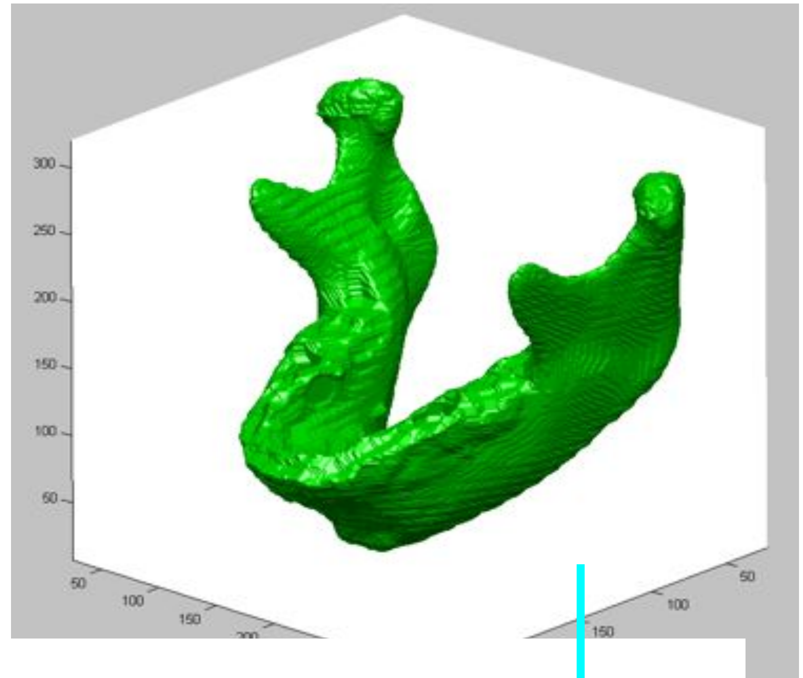
Chung et al. IPMI 2009

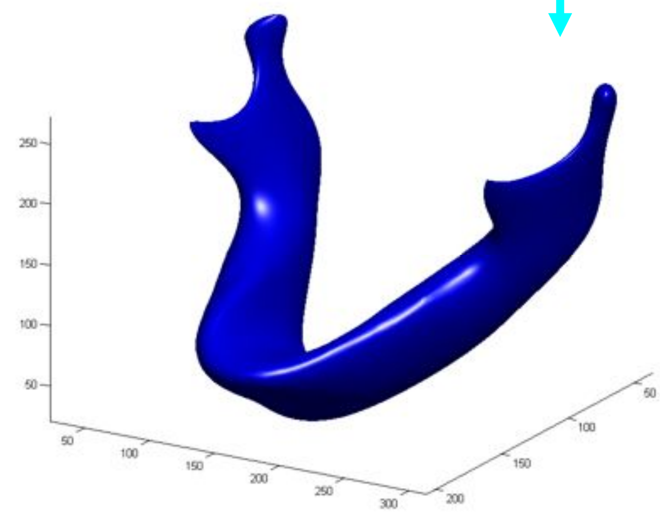
Topology correction in images

Histogram thresholding



Hole patching

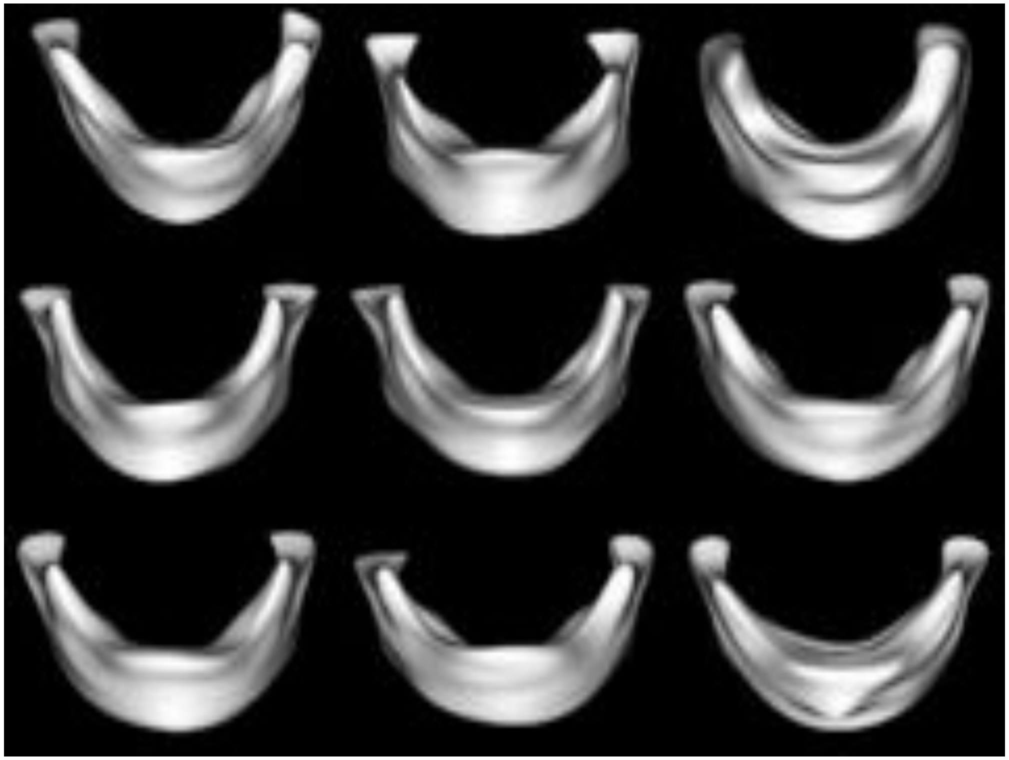


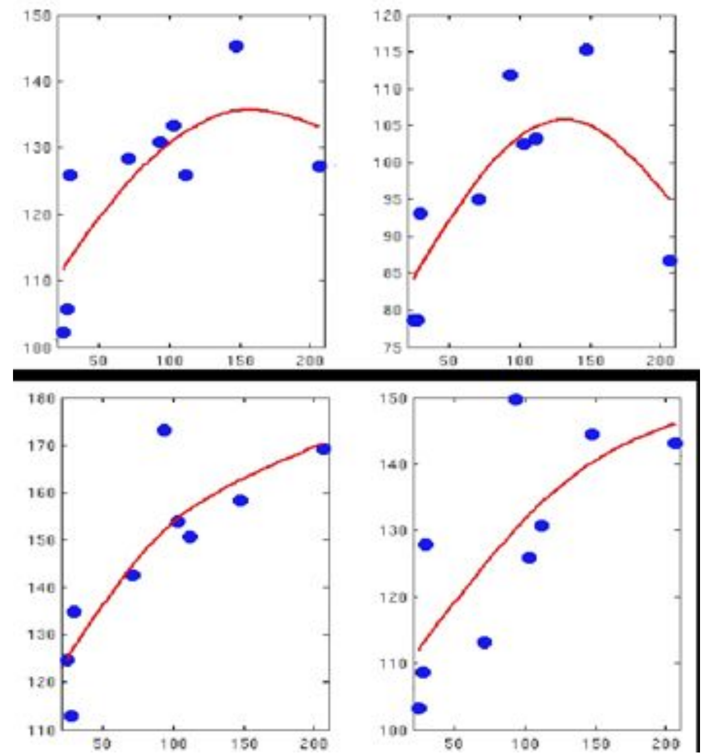


Automatic hole patching is necessary to construct surface topologically equivalent to sphere.

Approximately 20,000 triangle elements

Mandible surface growth modeling





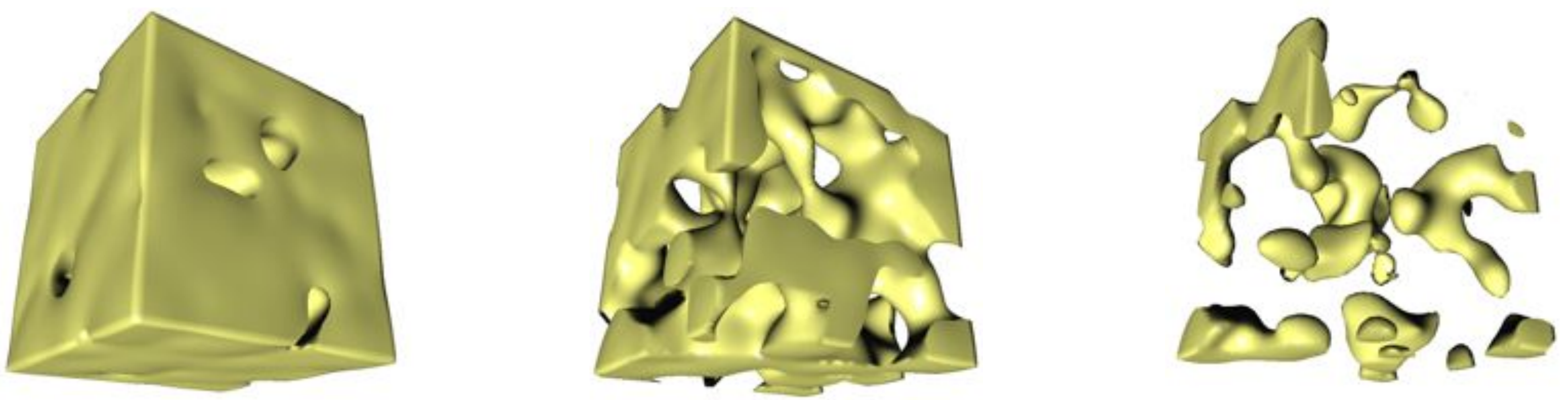
Quadratic fit of 9 male subjects over time in one particular point on the mandible surface

Worsley's random field theory based approach

$Z(x)$: Stationary isotropic random field in $x \in \Omega \subset \mathbb{R}^N$

$A_z = \{x : Z(x) > z\}$ excursion set

$\chi(A_z)$: Euler characteristic



z = -10 z = 0 z = 10

$$P\left(\max_{x \in \Omega} Z(x) > z\right) \approx \mathbb{E}\left(\chi(A_z)\right)$$

(Adler, 1984)

T random field on manifolds

$$P\Big(\max_{\mathbf{x}\in\partial\Omega_{atlas}} T(\mathbf{x}) \geq y\Big) \approx 2\rho_0(y) + \|\partial\Omega_{atlas}\|\rho_2(y)$$

Euler characteristic density

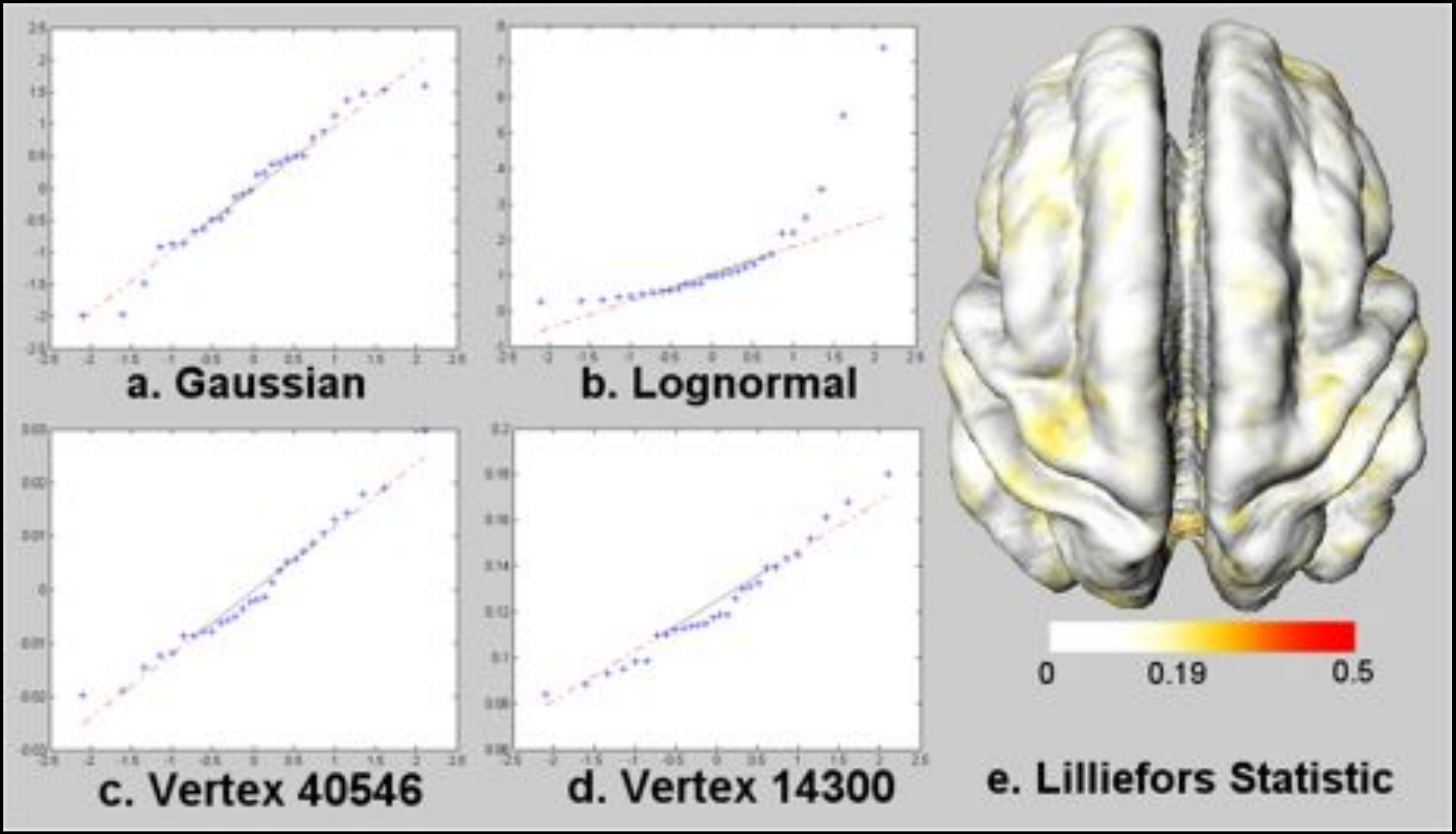
$$\rho_0(y) = \int_y^\infty \frac{\Gamma(\frac{n}{2})}{((n-1)\pi)^{1/2}\Gamma(\frac{n-1}{2})}\Big(1+\frac{y^2}{n-1}\Big)^{-n/2} dy,$$

$$\rho_2(y) = \frac{1}{\mathrm{FWHM}^2}\frac{4\ln 2}{(2\pi)^{3/2}}\frac{\Gamma(\frac{n}{2})}{(\frac{n-1}{2})^{1/2}\Gamma(\frac{n-1}{2})}y\Big(1+\frac{y^2}{n-1}\Big)^{-(n-2)/2}$$

Worsley (1995, NeuroImage)

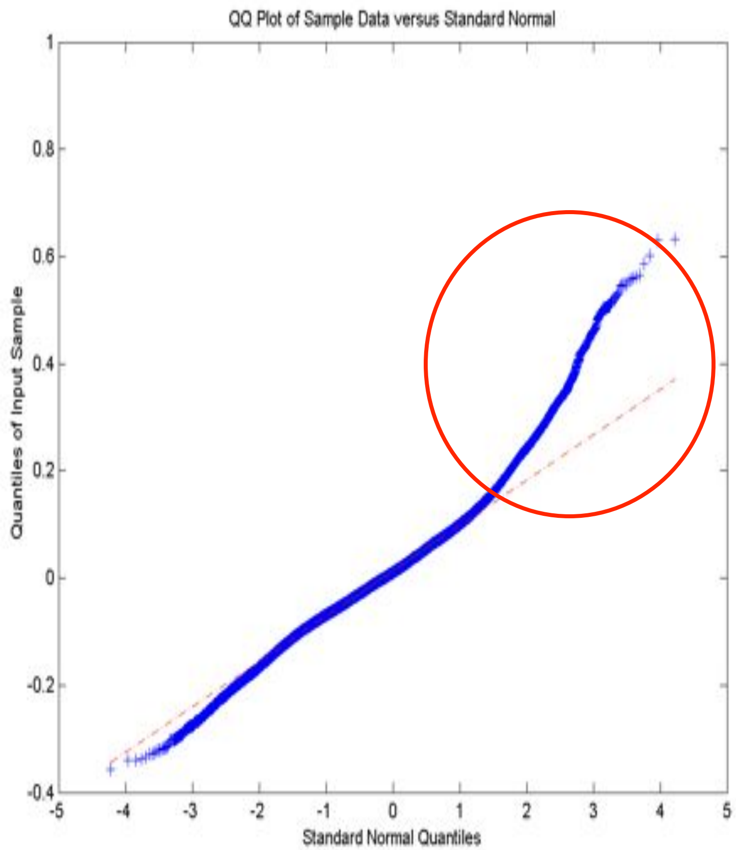
FWHM of smoothing kernel or residual field

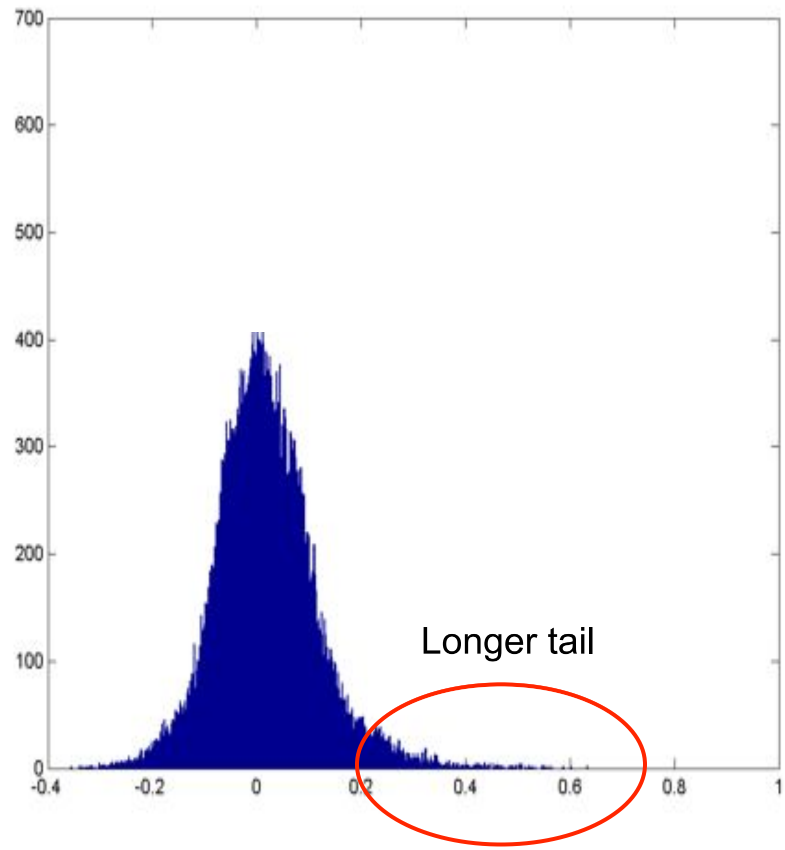
Gaussianess may not be satisfied

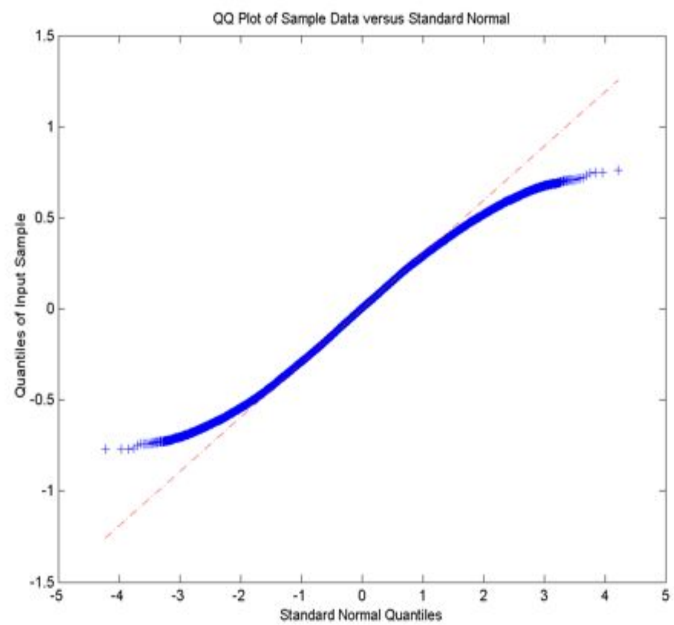


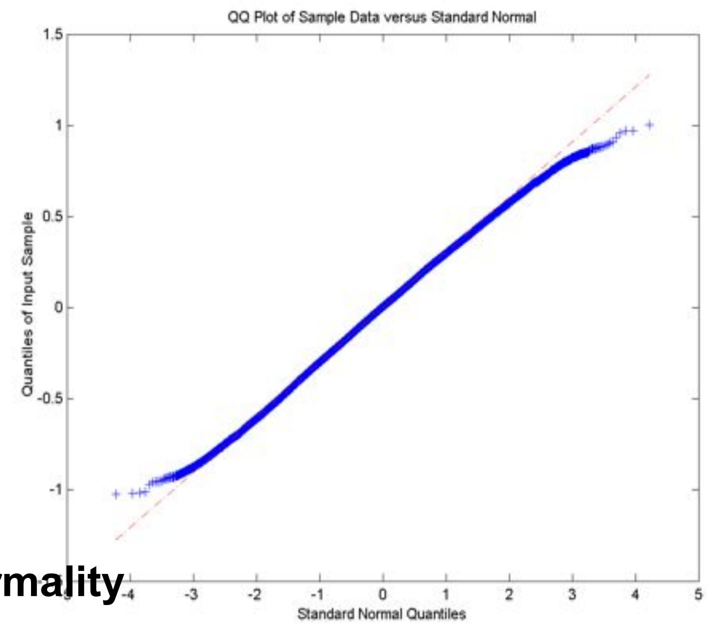
Checking normality of imaging measures

Quantile-Quantile (QQ) plot showing asymmetric distribution

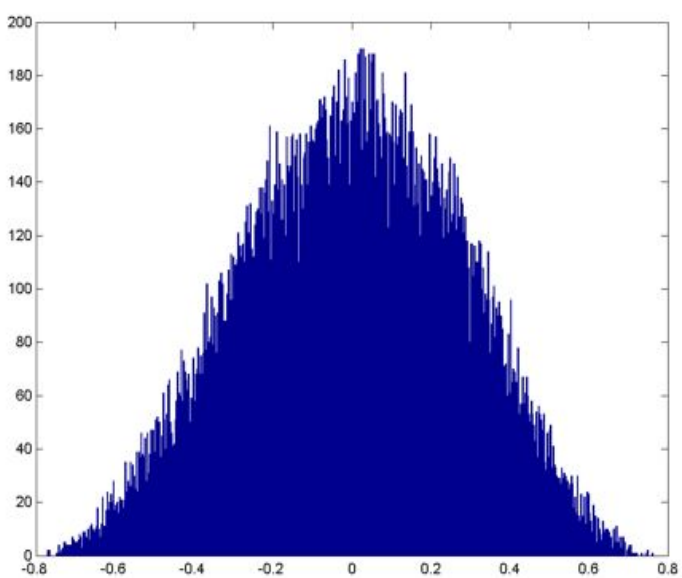




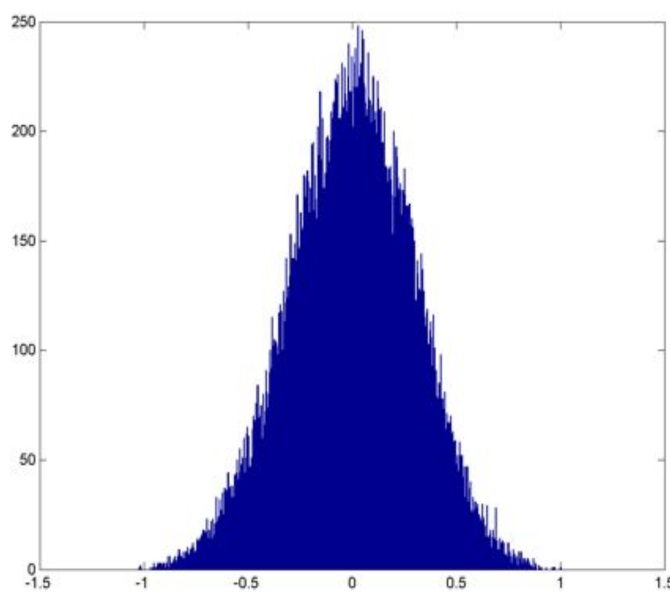




Increasing normality of data

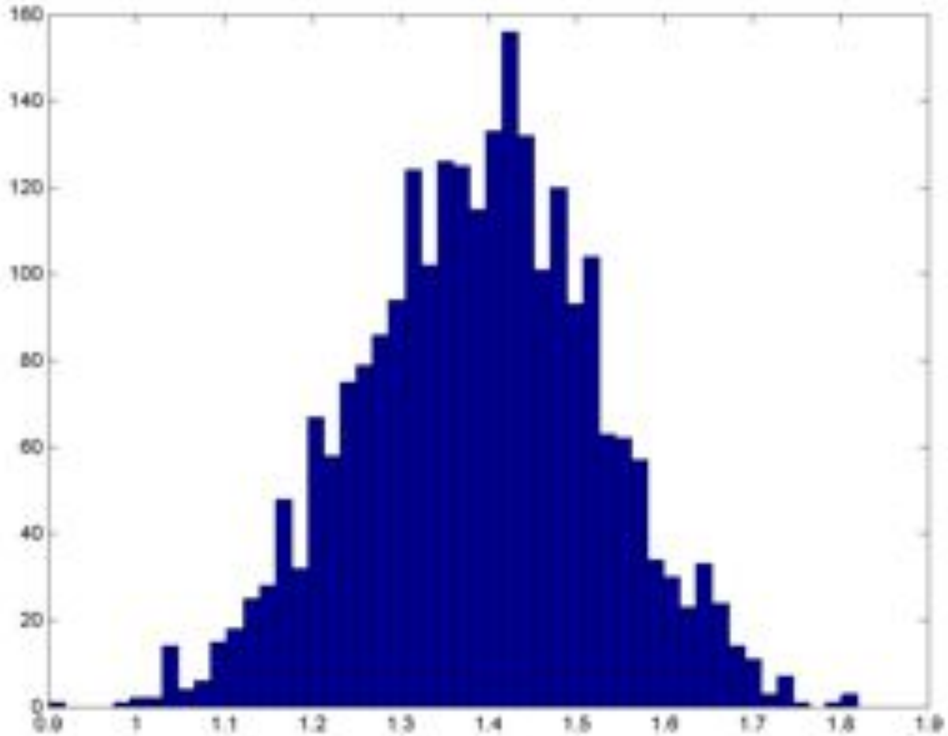


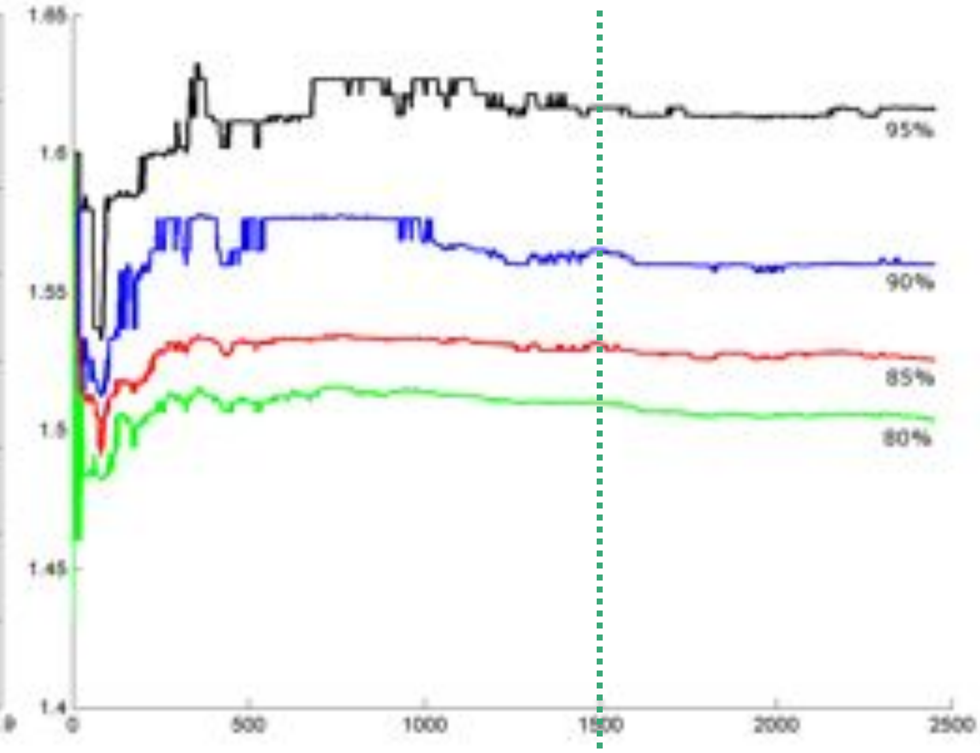
Fisher's
Z transform
on correlation



Permutation test

model free statistical inference



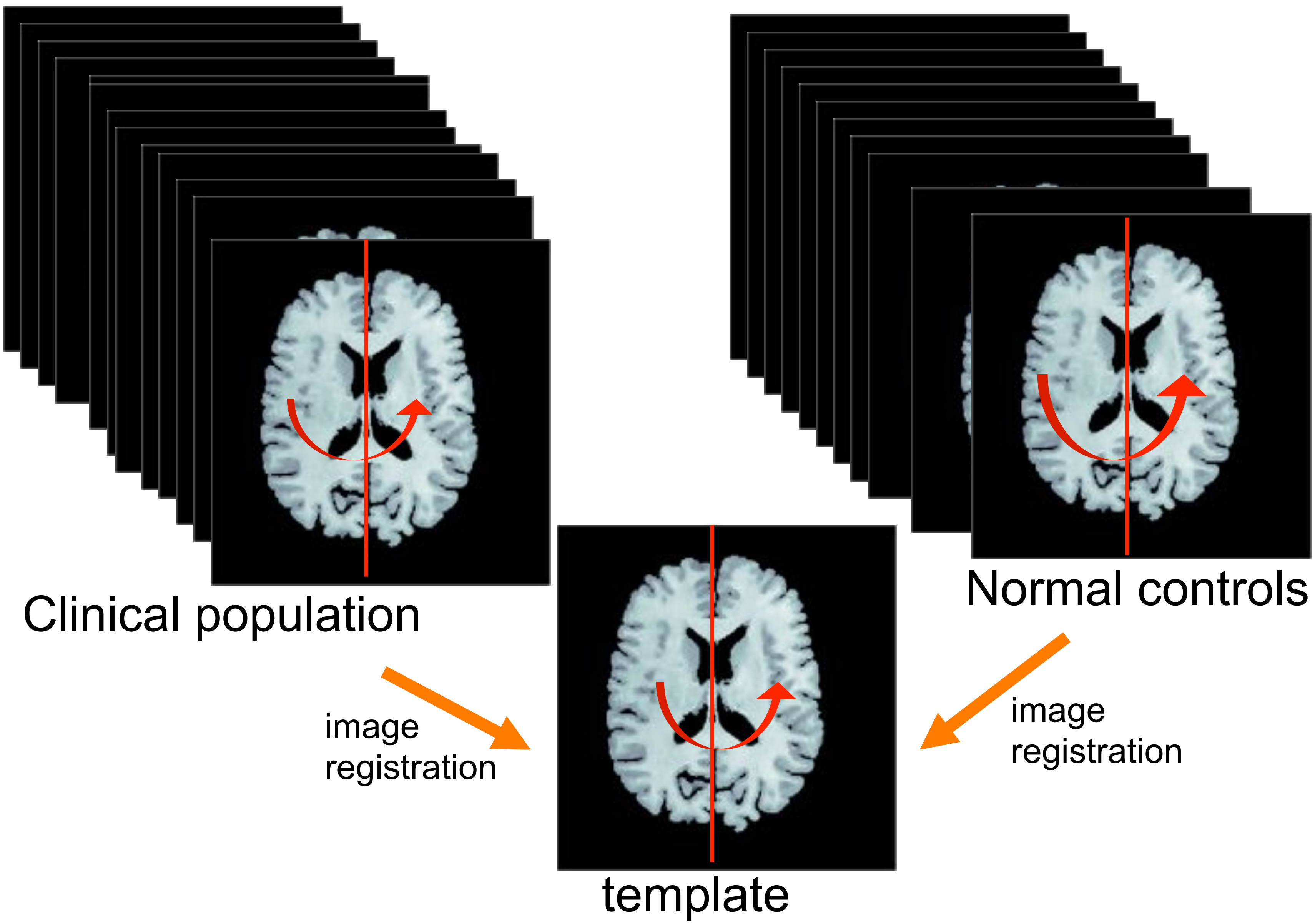


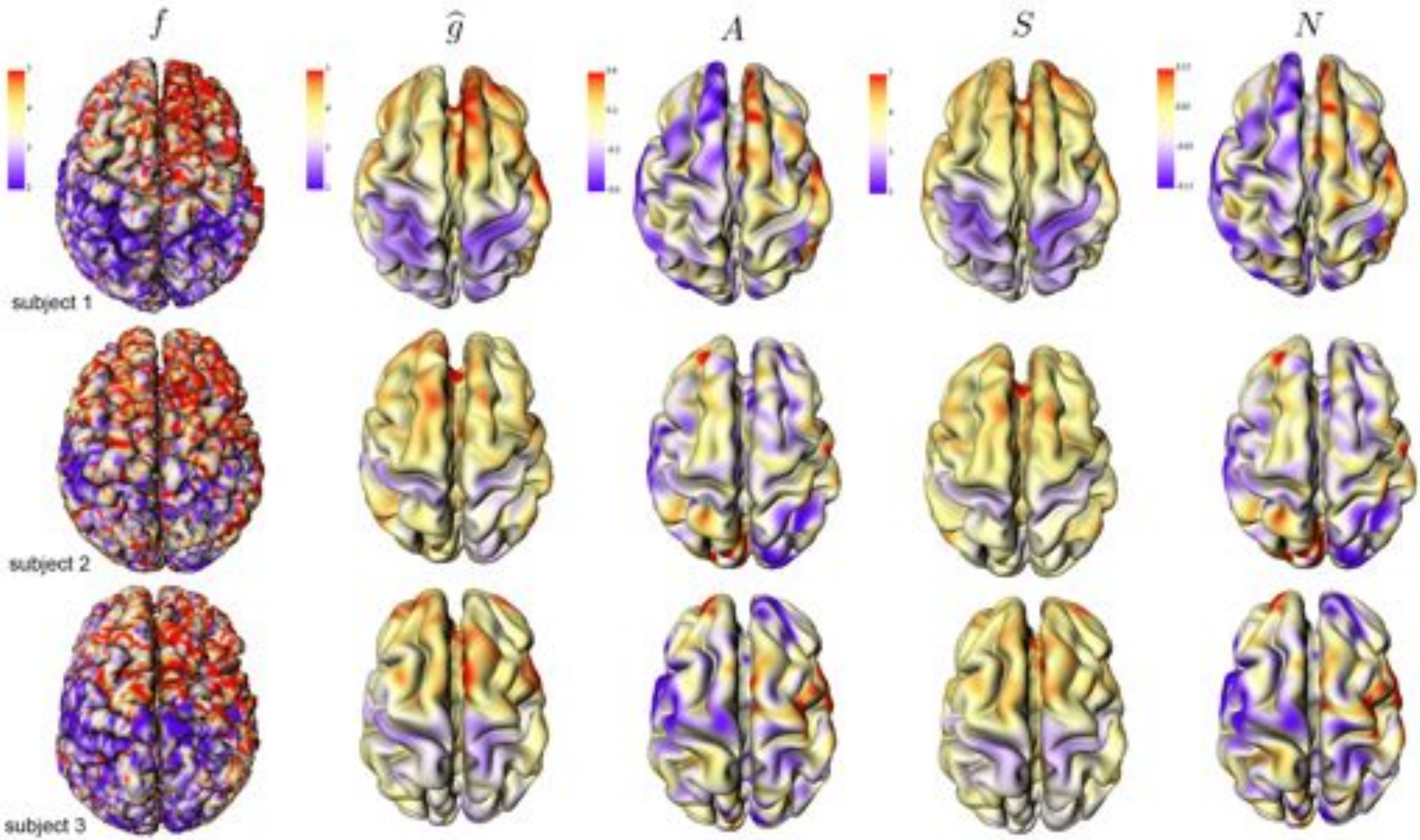
More than 1500 permutations are needed to guarantee the convergence of the thresholding.
8 hours of running time in MATLAB.

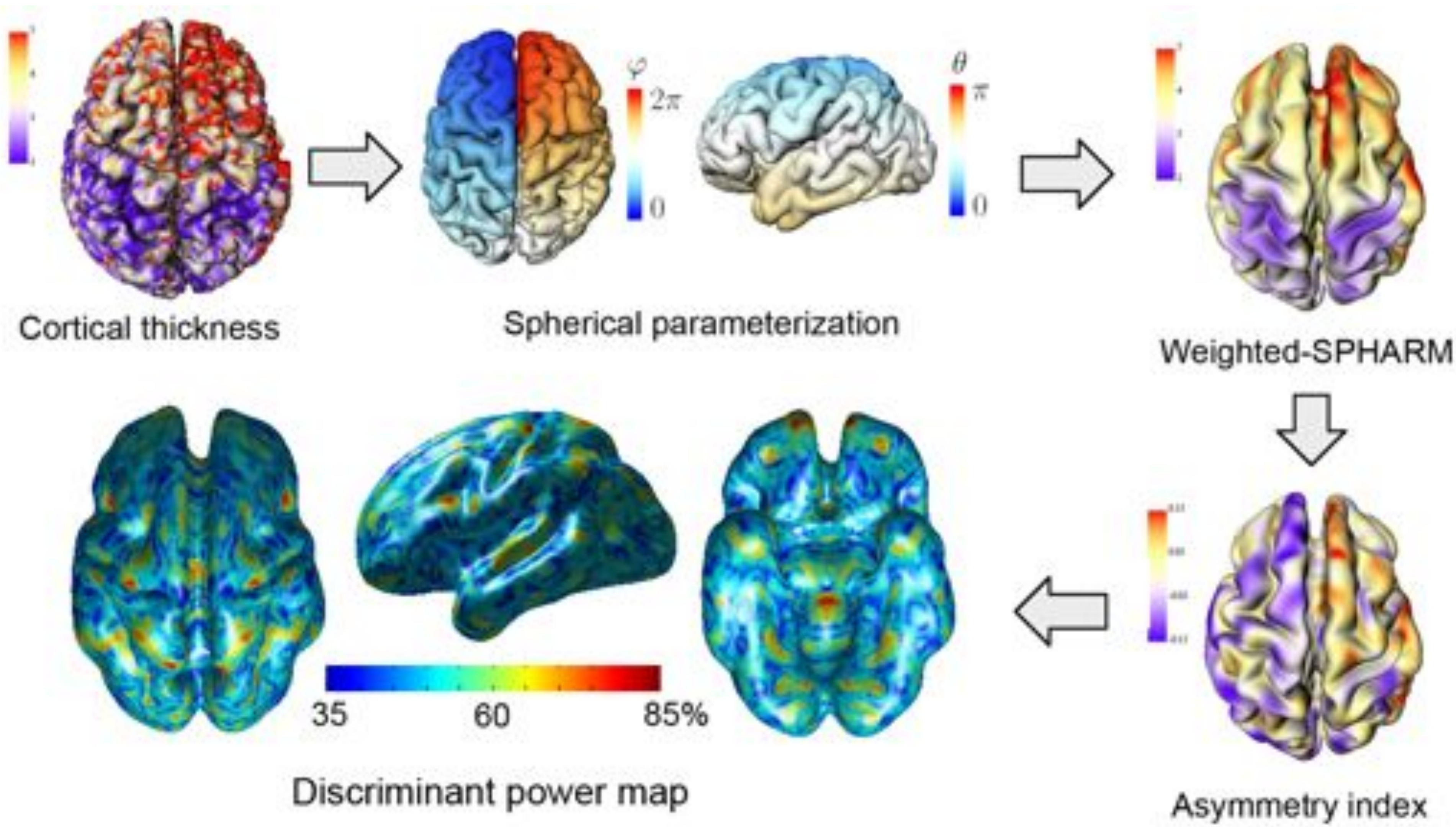
Logistic discriminant analysis

Based on logistic regression that connects categorical variables to continuous variables, we can perform a discriminant analysis

asymmetry analysis framework




f
$\widehat{g}$
A
S
N
subject 1
subject 2
subject 3



Cortical thickness
φ
2π
0
θ
π
0
Spherical parameterization
Weighted-SPHARM
Asymmetry index
35
60
85%
Discriminant power map

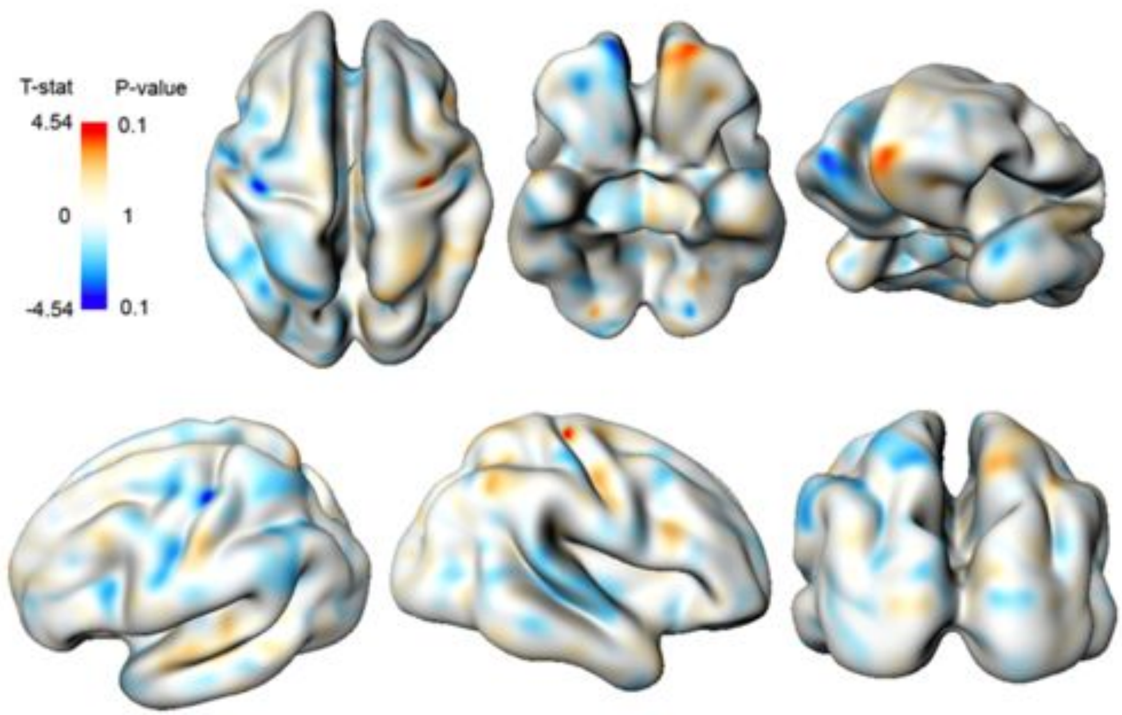
Statistical Parametric Map

multiple comparison correction via the random field theory (Worsley et al. 1995) → not so trivial

$$P(\sup_{p \in \partial\Omega} Z(p) > h) \approx \sum_{d=0}^{2} \phi_d(\partial\Omega)\rho_d(h)$$

> 40000 correlated hypotheses

Very involving mathematical derivation



T-stat resulting showing group difference between autism and control

Hypothesis & P-value free approach

Discriminant Power Map

Logistic model

$$\log \frac{\pi_i}{1-\pi_i} = \beta_0 + \beta_1 n_i$$

Probability of autism

Asymmetry index

Classification rule:

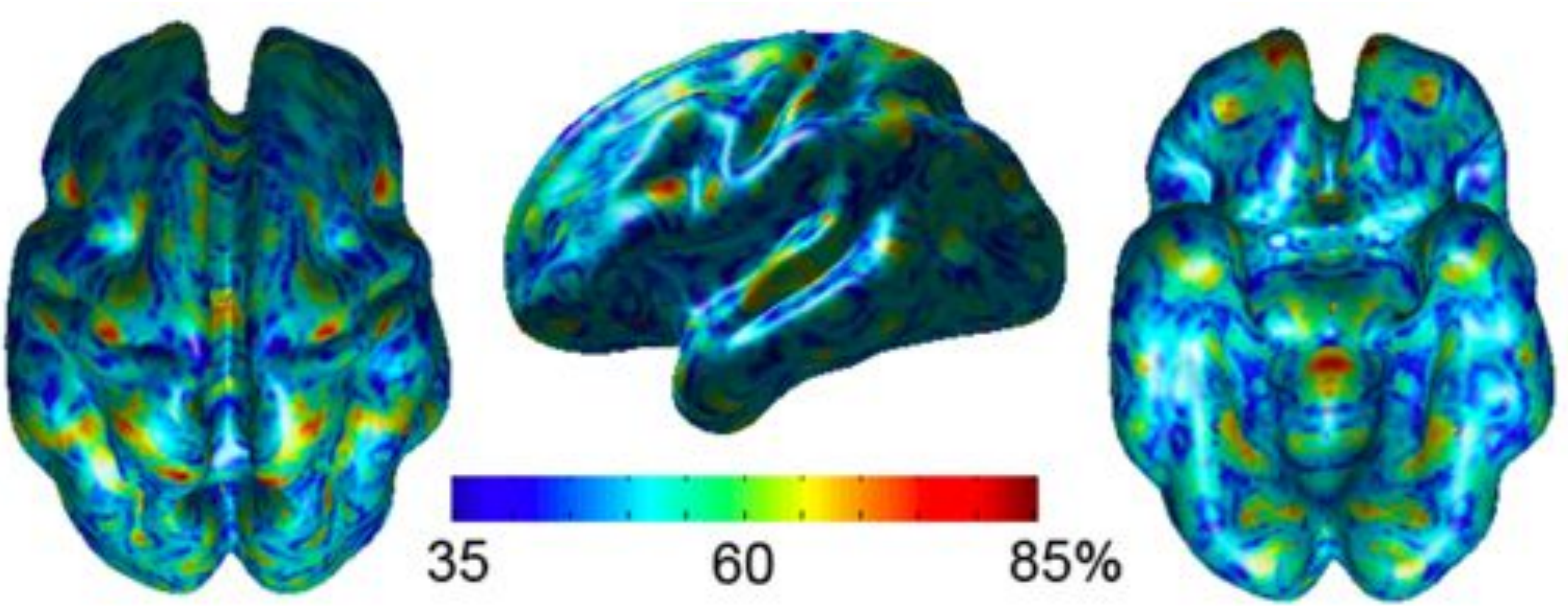
$$\pi_i > \frac{1}{2}$$

Leave-one-out cross-validation

Classification error rate

Discriminant Power Map

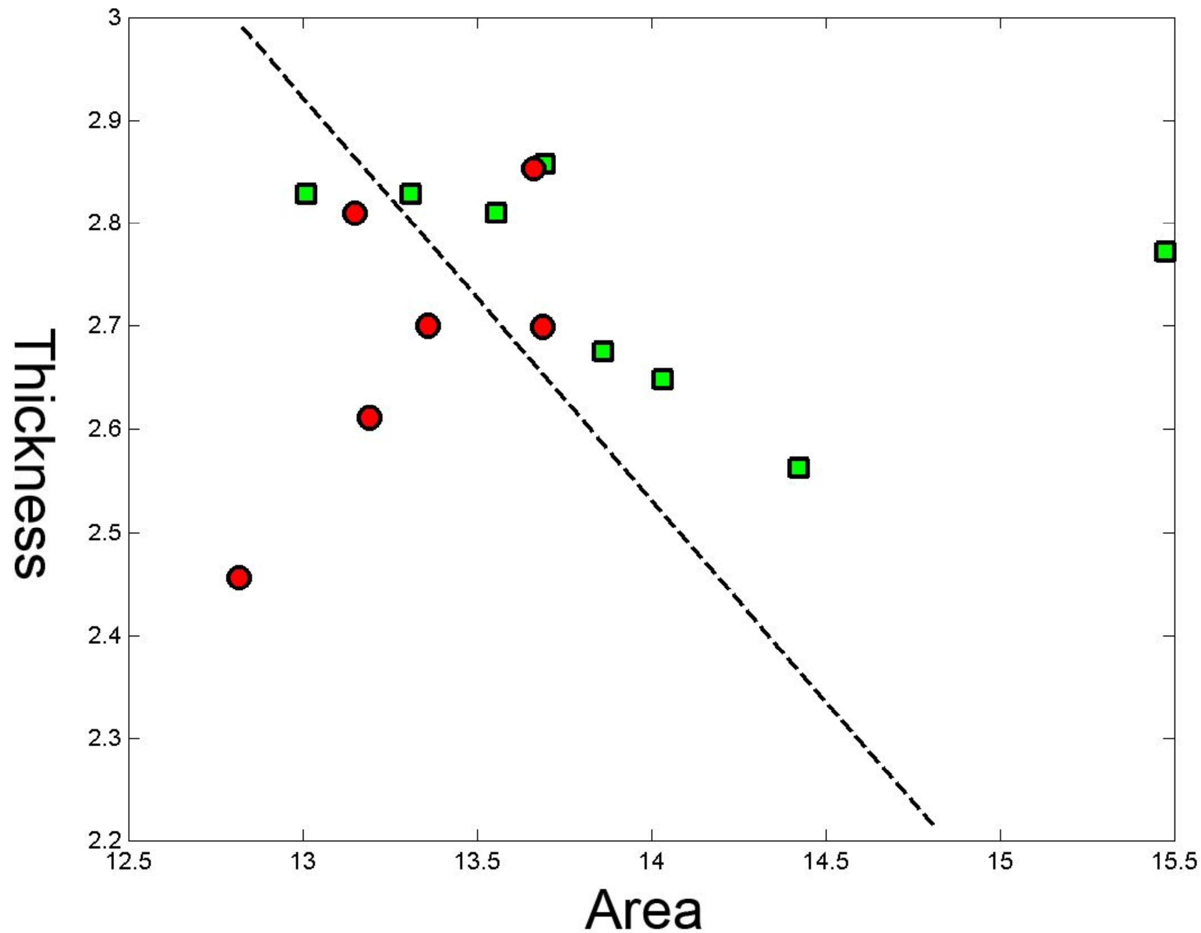
= 1 - error rate



Avoid the traditional hypothesis driven approach
No need to compute P-value → No need for random field theory

Adaboost version with spatial dependency constrain
Singh et al., MICCAI 2008.

Classification for imaging biomarkers via logistic discriminant analysis

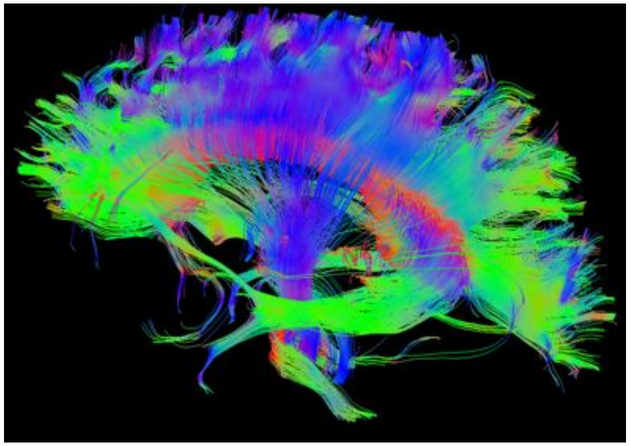


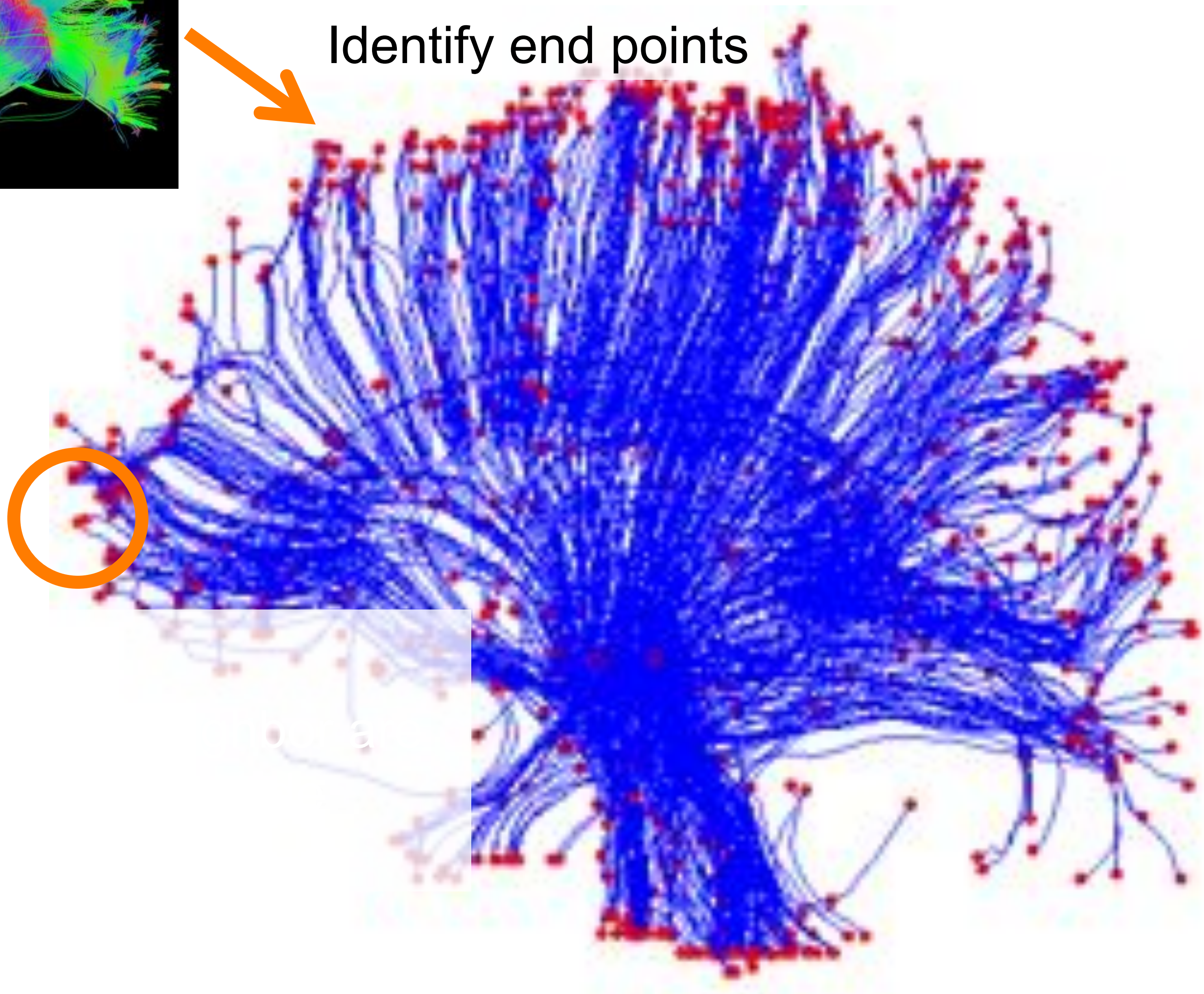
Red: mild cognition impairment (MCI)
Green: elderly normal controls

Brain Network Analysis

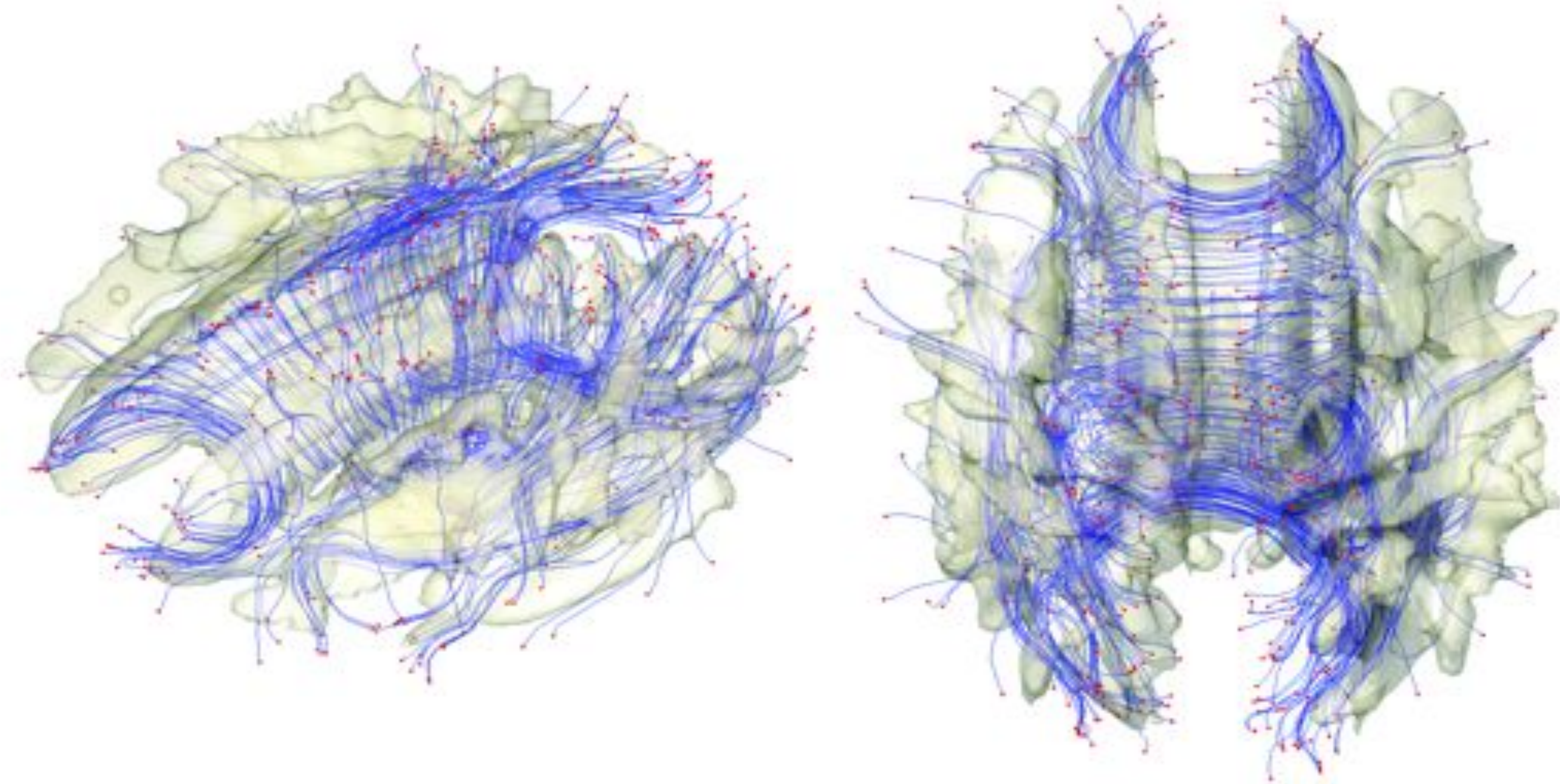
Graph theoretic approach to brain connectivity analysis. Various topological invariants are used to characterize brain connectivity.

Graph construction

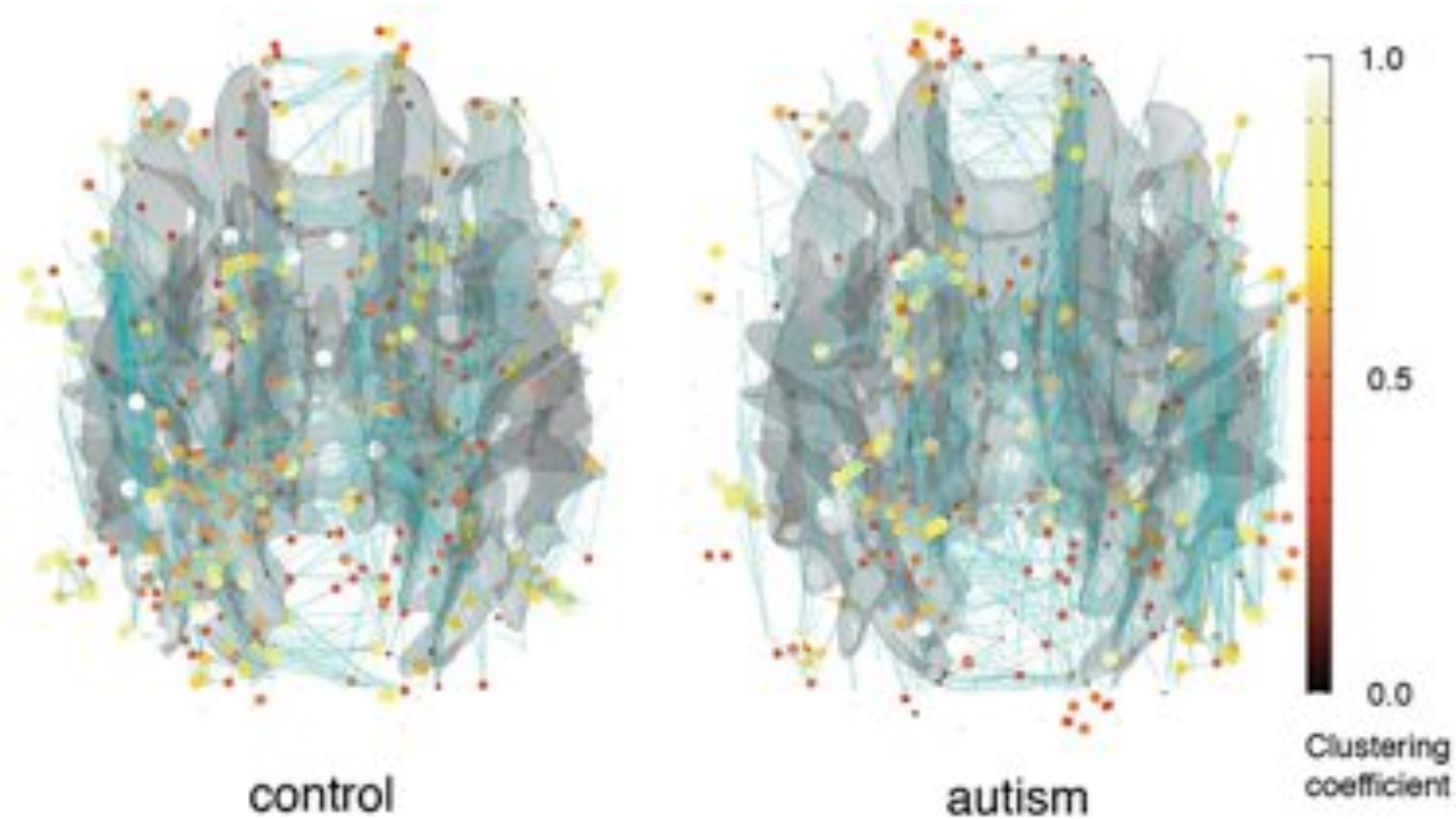




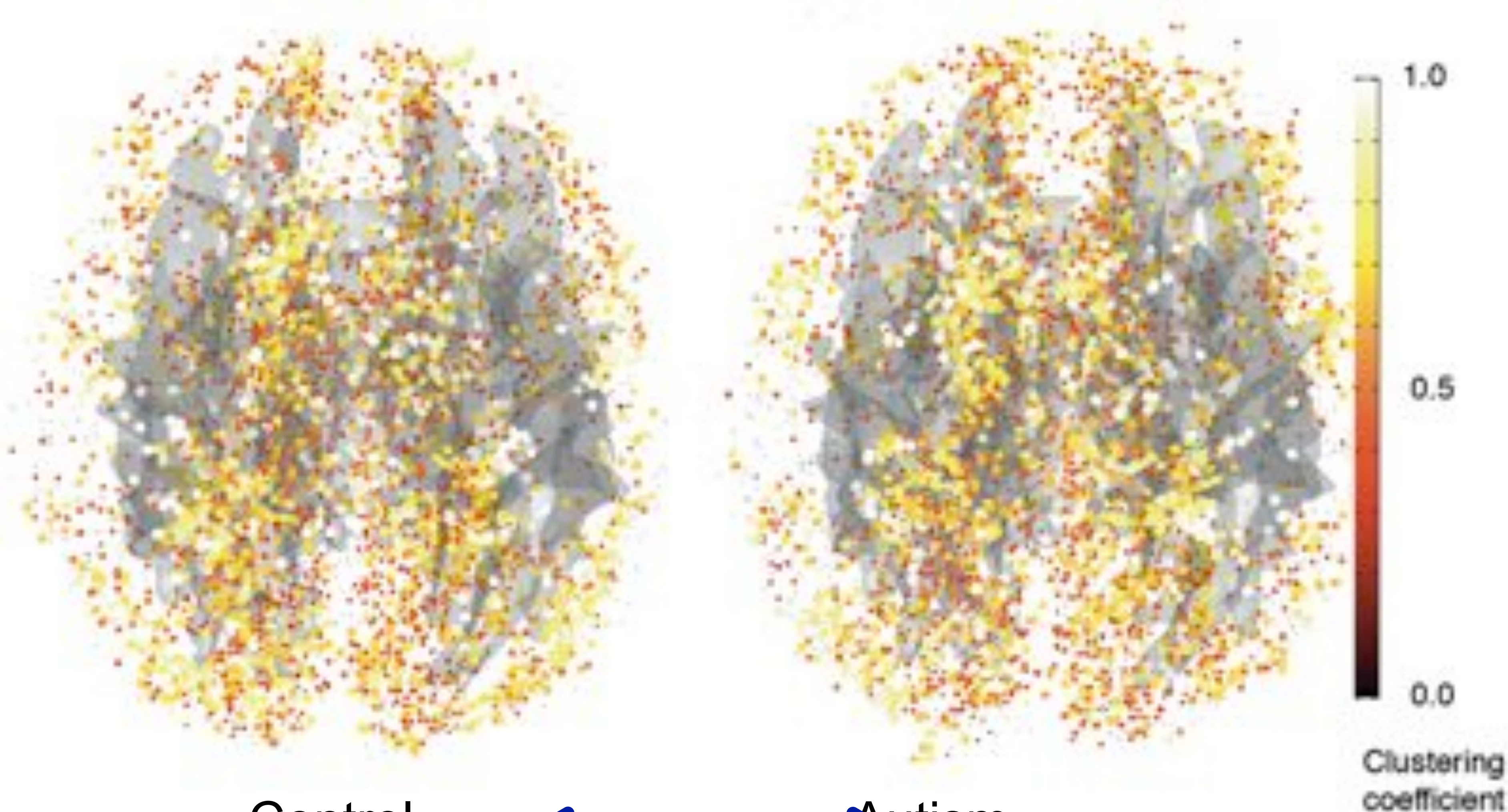
End points of tracts



Clustering coefficient



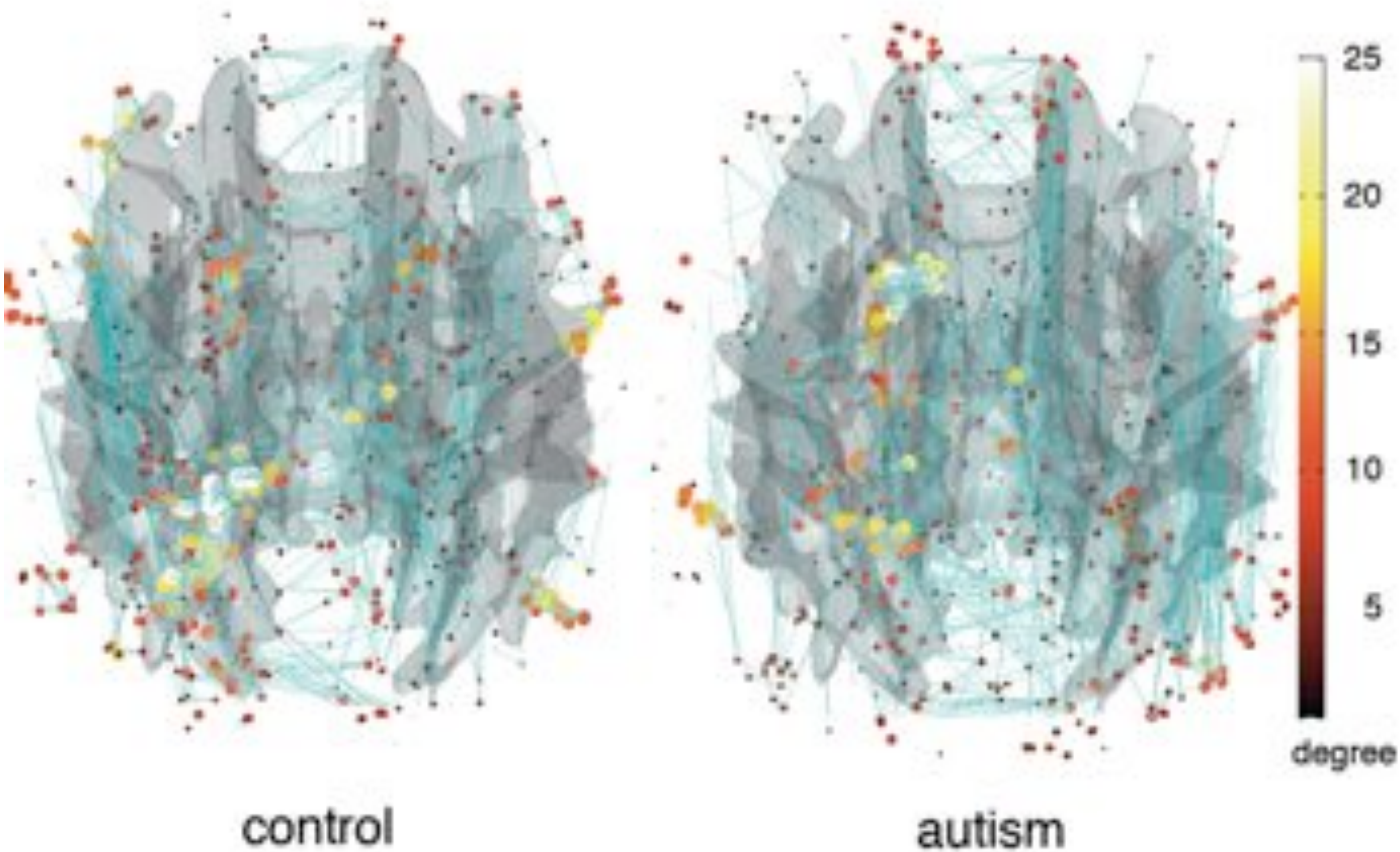
Clustering coefficients for all subjects



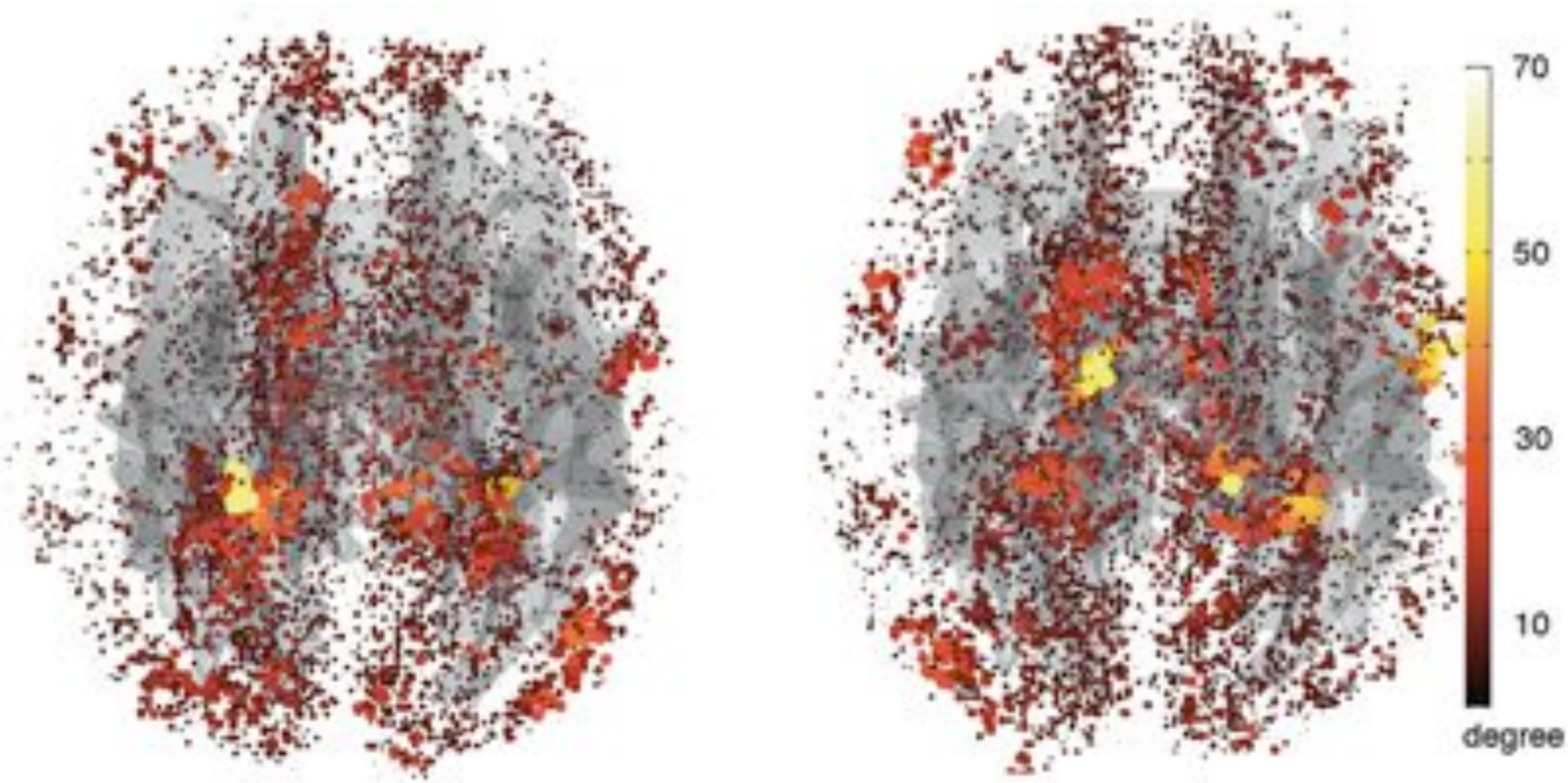
Control Autism

No group difference

Degree of nodes: measure of local network complexity



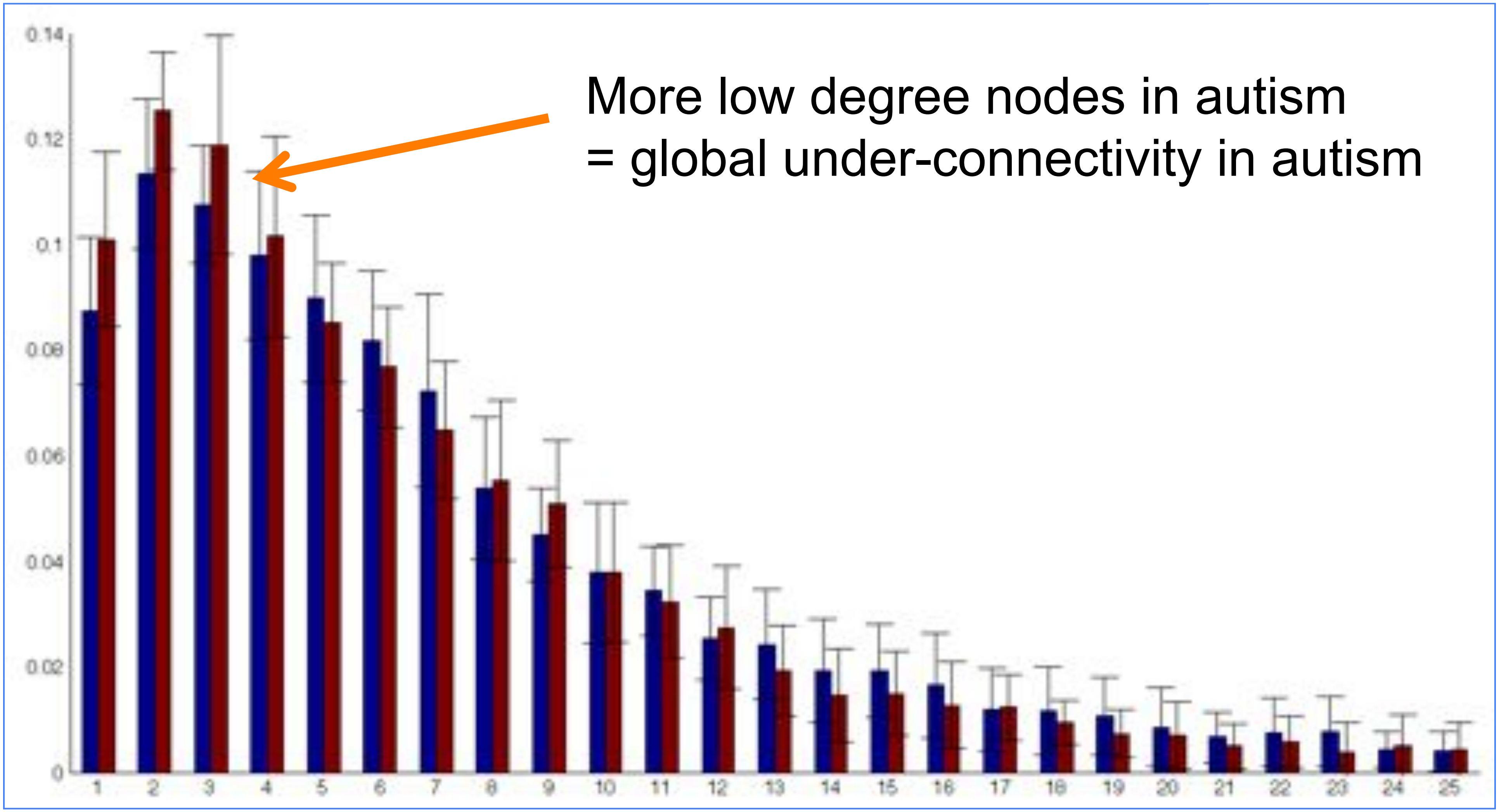
Local degree distribution for all subjects



Control Autism

Global degree distribution

red: autism
blue: control



pvalues = 0.024, 0.015 and 0.080 for degrees 1, 2 and 3.

Chung *et al.*, HBM conference 2010

Lecture 2

Least squares estimation

General linear model

Multivariate general linear model

Read two papers put in the literature directory:

chung.2004.ni.autism.pdf ← GLM
chung.2010.NI.pdf ← MGLM