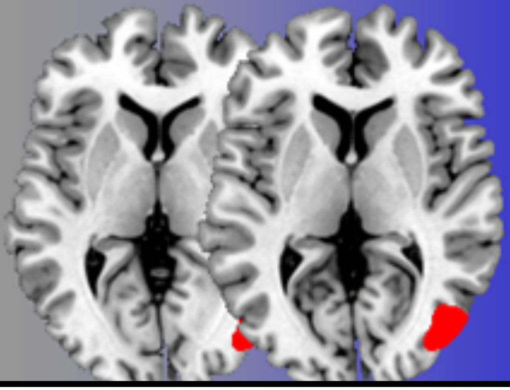
EMORY

ROLLINS SCHOOL OF PUBLIC HEALTH

CBIS

Center for Biomedical Imaging Statistics



BIOS 516

Neuroimaging Statistics

Ying Guo

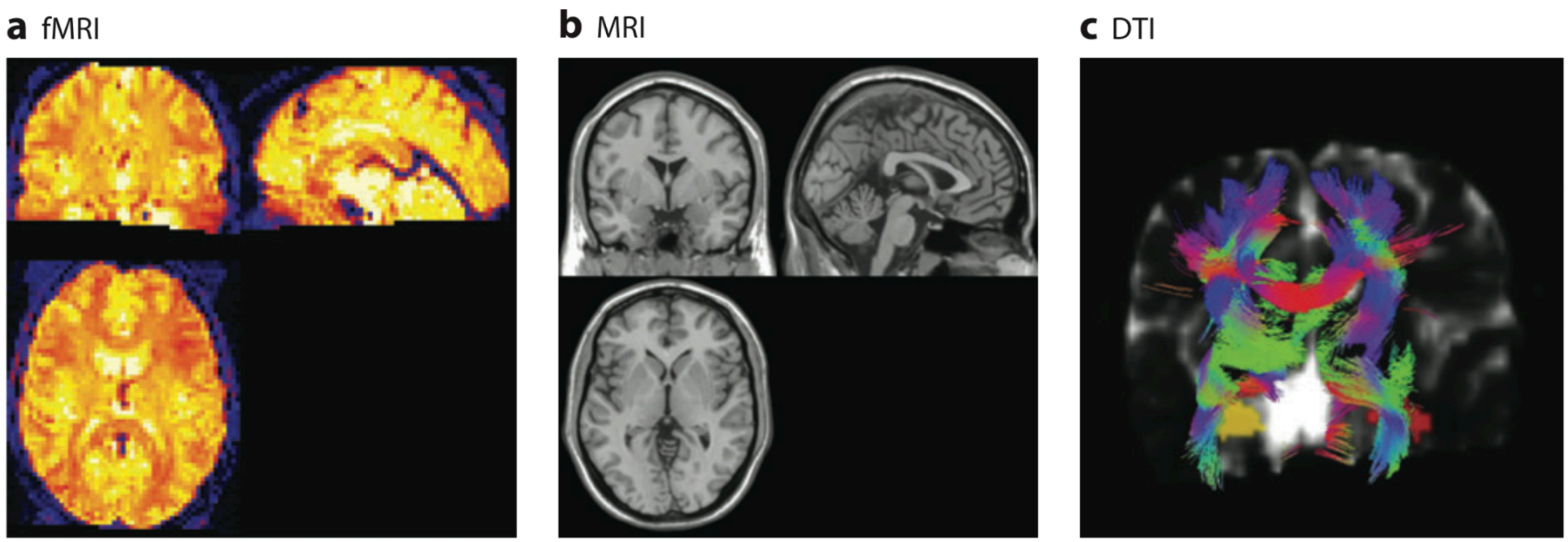
Department of Biostatistics and Bioinformatics

Emory University

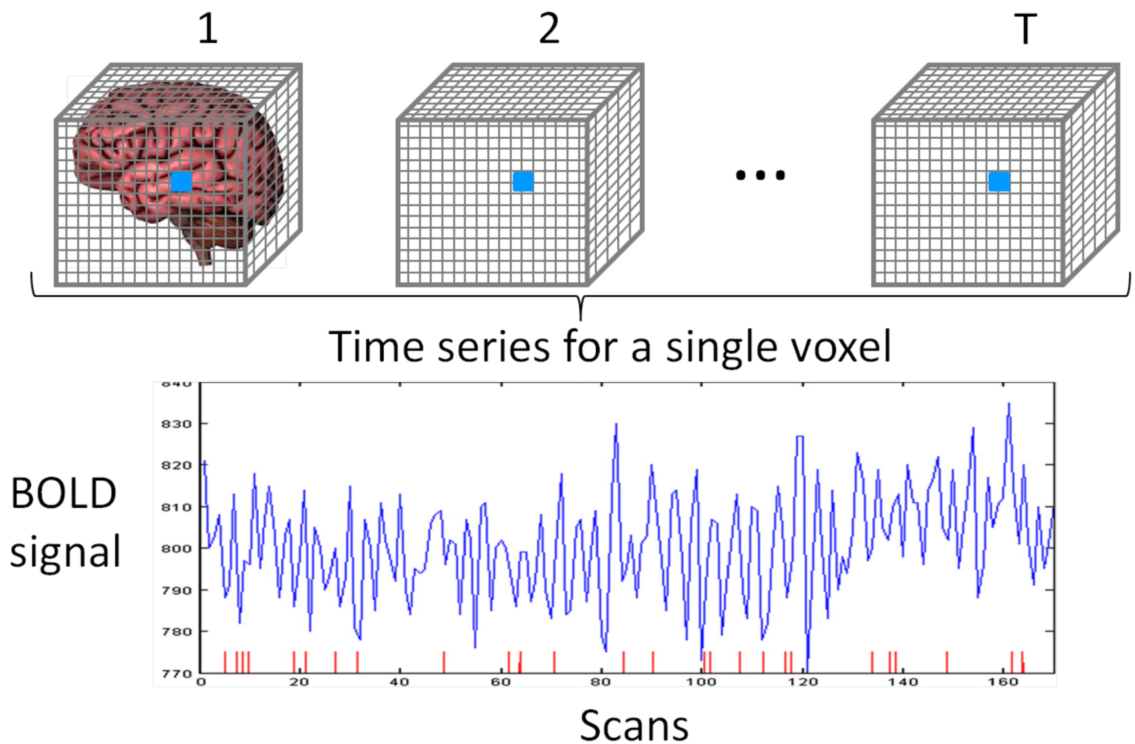
Statistical Analysis of Neuroimaging Data

Review from last time…

- Structural Imaging modalities
  – MRI, CAT, DTI (diffusion tensor imaging)
- Functional Imaging modalities
  – **fMRI,** PET, MEG & EEG



fMRI data



- Each fMRI image consists of ~100K voxels
- During the course of the experiment, hundred of images are acquired (~ one every 2 sec)
- One voxel → one BOLD time series

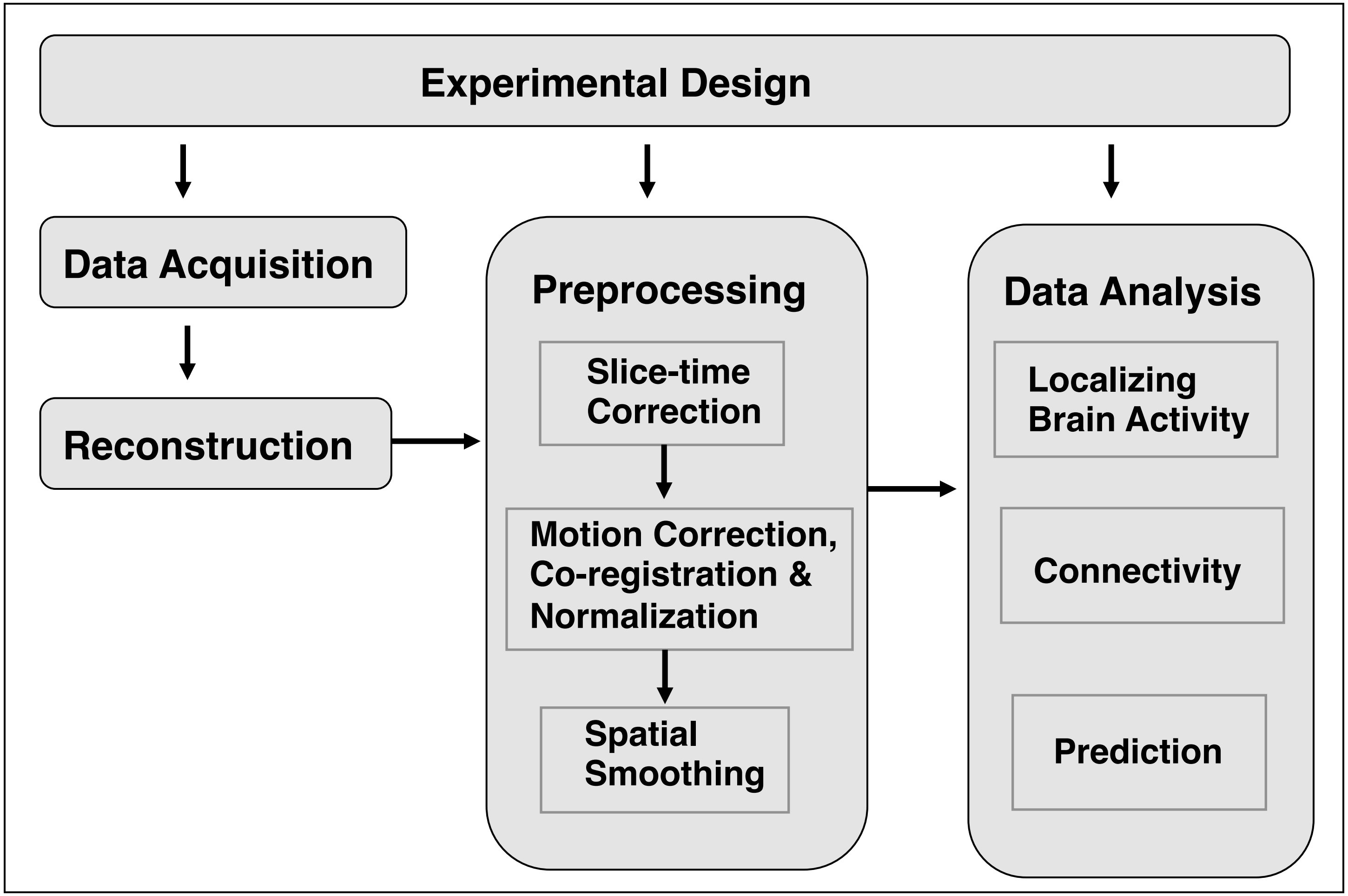
There are multiple goals in the statistical analysis of neuroimaging data:

- **ACTIVATION**: Localizing brain areas activated by the experimental task (Brain Mapping)
- **BRAIN CONNECTIVITY and NETWORK ANALYSIS**
- **PREDICTION**: making predictions about psychological or disease states

Challenges

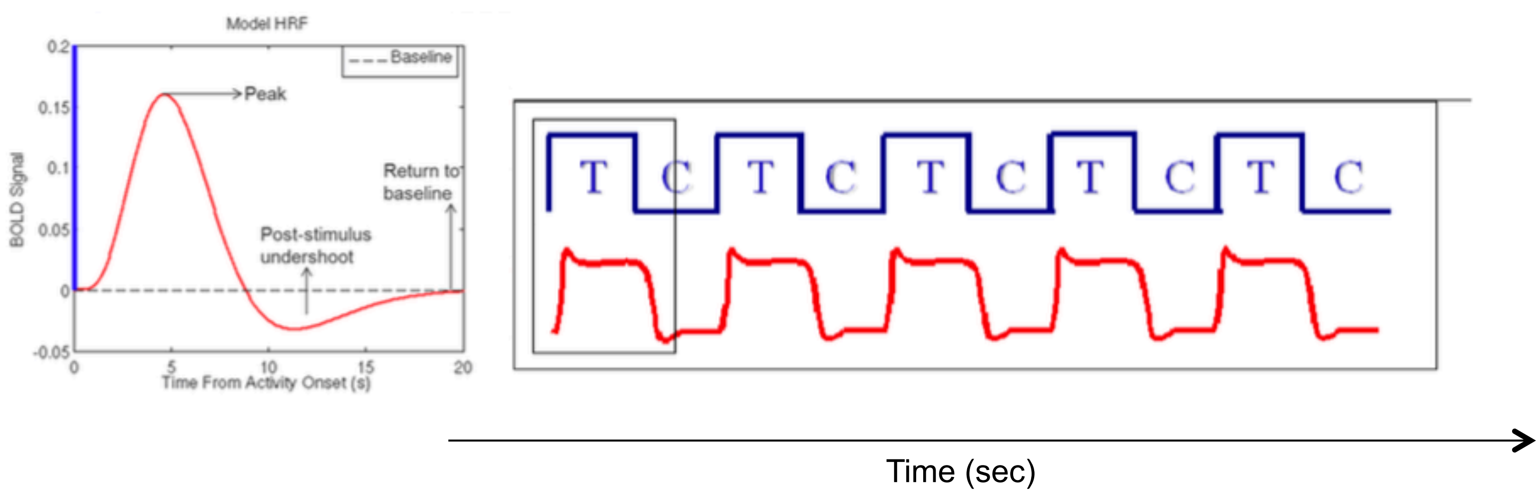
- The statistical analysis of fMRI data is challenging.
  - It is a massive data problem
  - The signal of interest is relatively weak (only 0.5-3% change in intensity with 1.5T scanner)
  - The data exhibits a complex temporal and spatial structure

Analysis Procedure



fMRI study designs

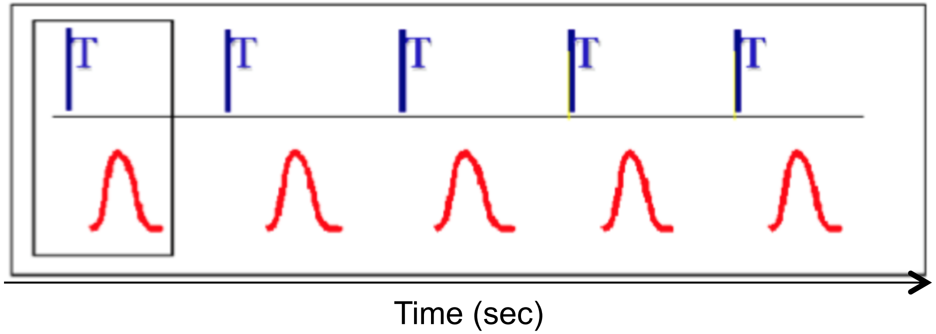
- **Block Designs**: stimuli of the same condition are grouped together in blocks



- **PRO**: Repeating the stimulus in a block causes a large total signal change – increases statistical power to detect activation
- **CON**: Can't directly estimate features of the HRF

fMRI study designs

- **Event-related Designs**: allow different stimuli to be presented in a mixed sequence



- **PRO**: Can precisely observe the actual HRF – thus allowing for the estimation of features of the HRF
- **CON**: Decreased power to detect activation (lower SNR)

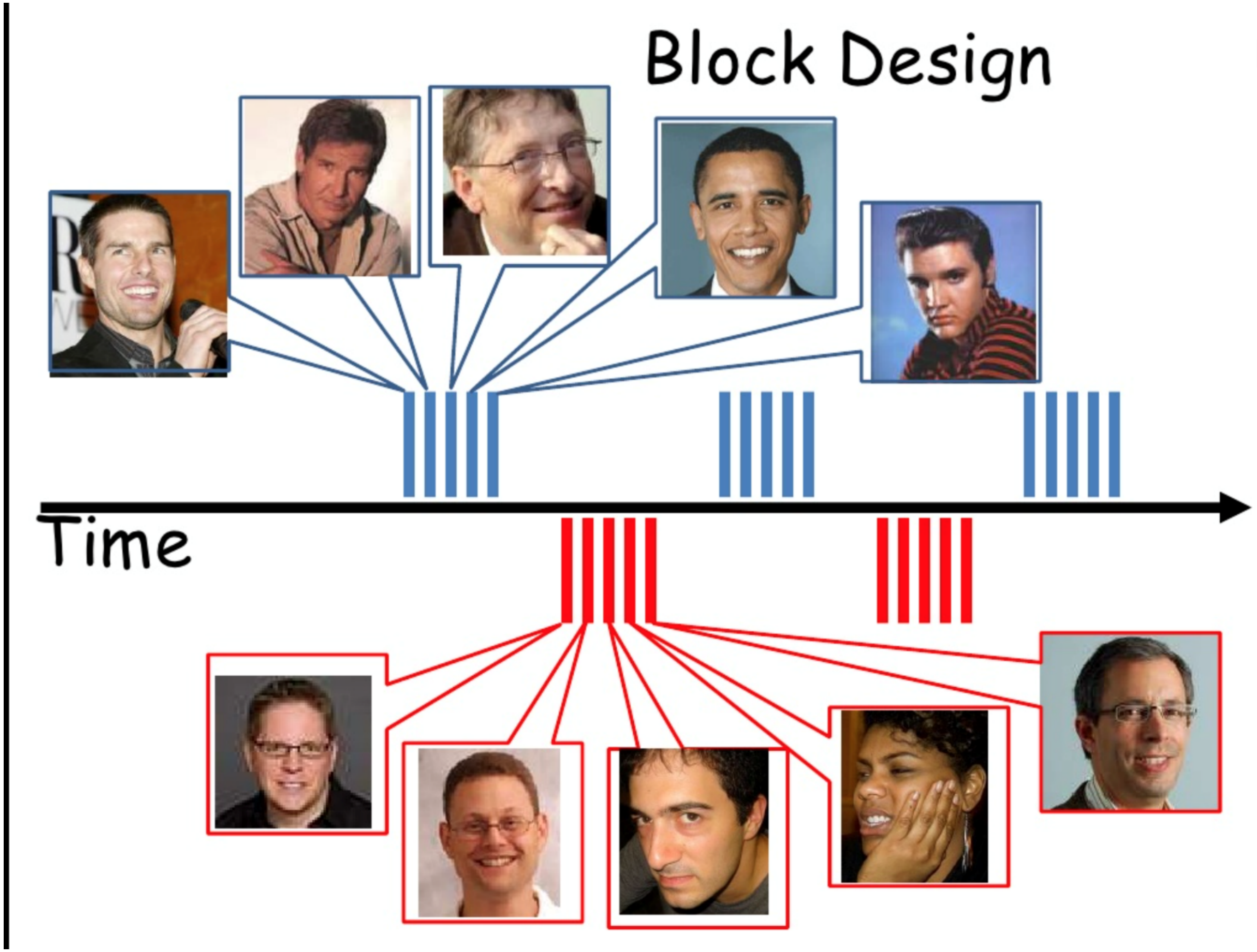
Simple task example

Facial recognition task

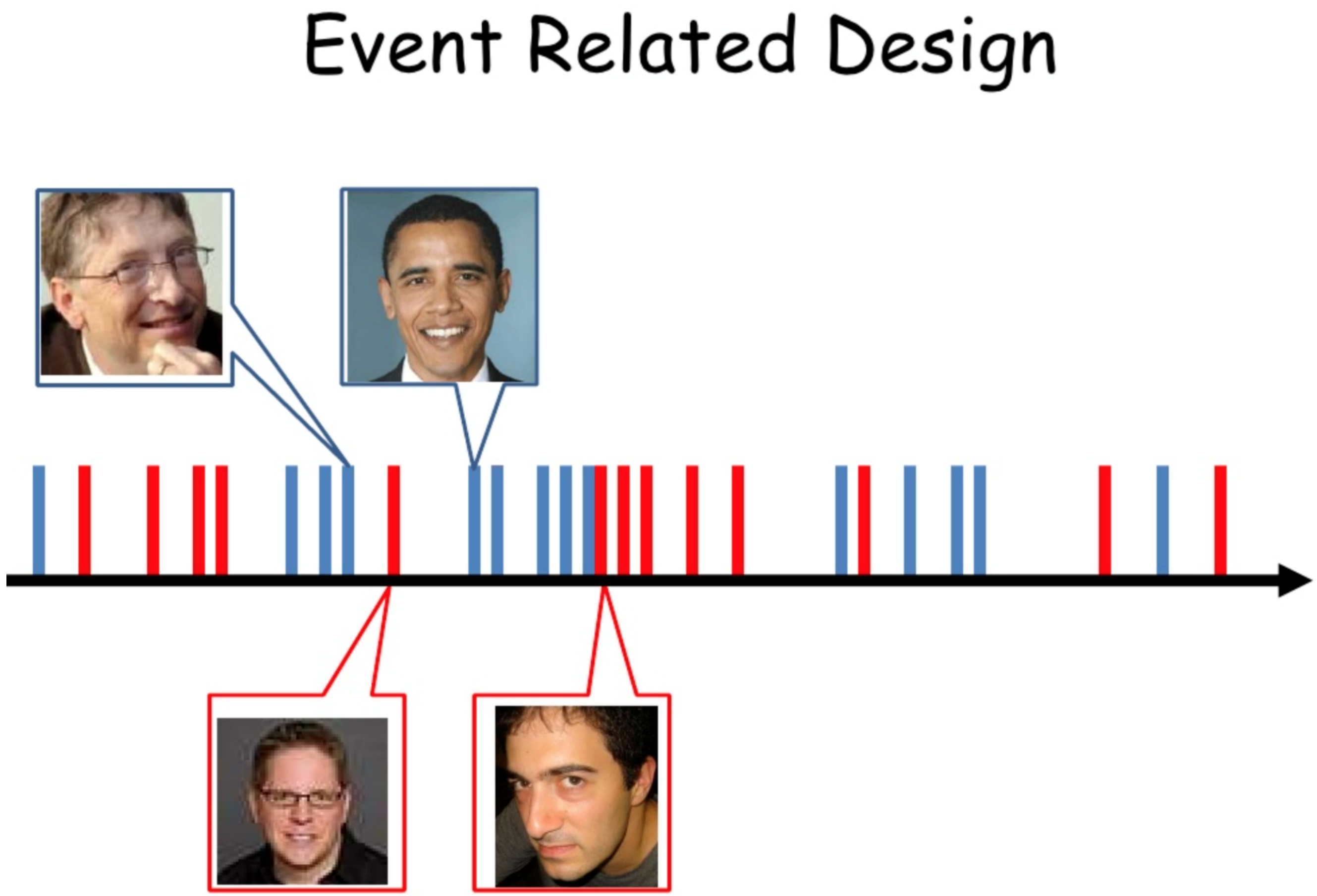
Task A: View famous faces

Task B: View non-famous faces

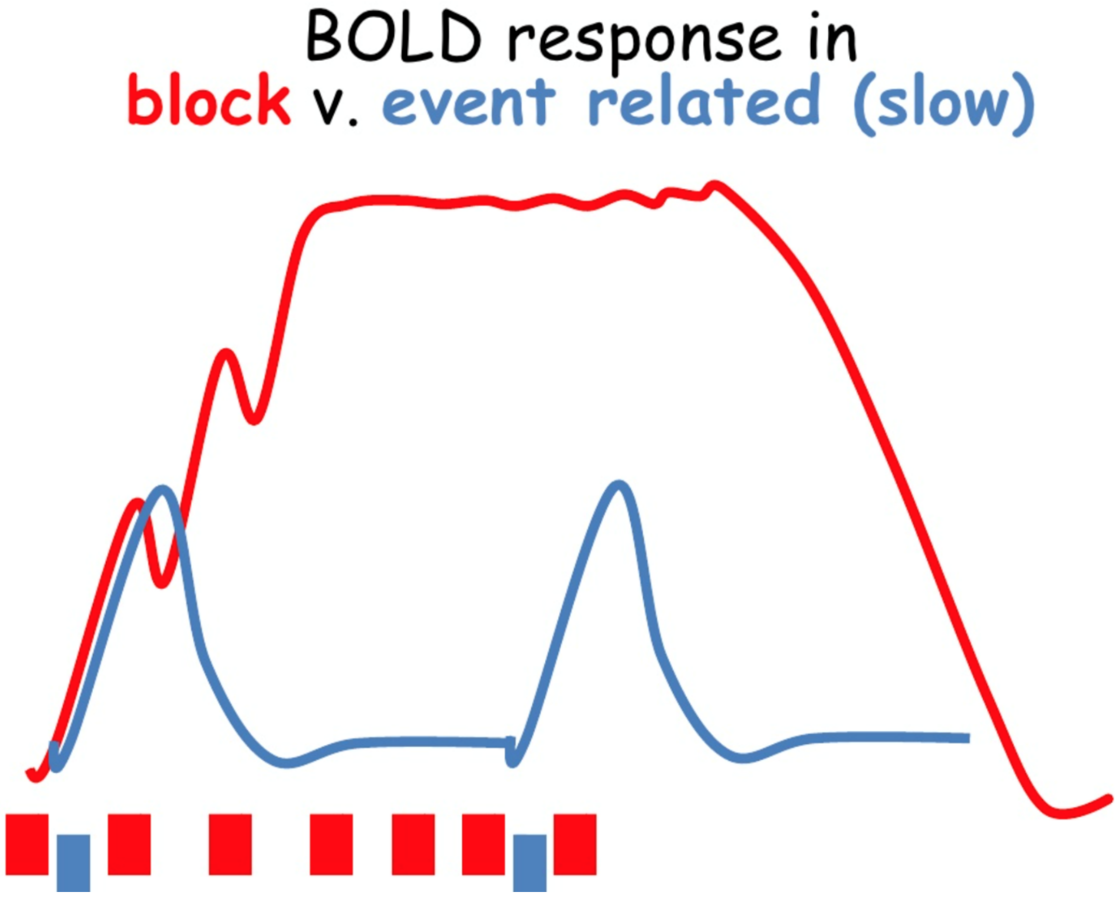
Simple task example



Simple task example



Block vs event-related design



Block Design Issues

- Repetitions can get predictable, reducing activation

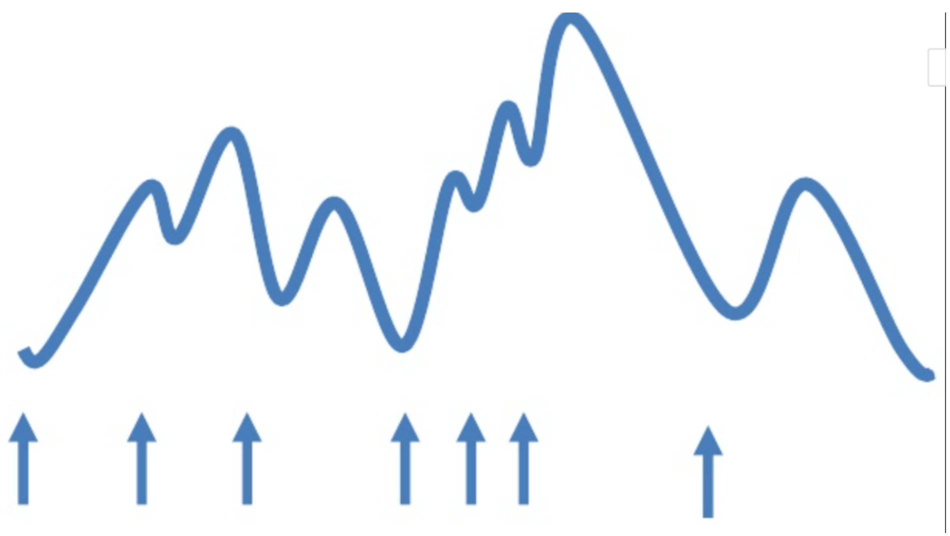
- Timing Issues:
  - Ideally 15-20 sec on, then 15-20 sec off
  - Long enough for HRF to relax in between presentations
  - Short enough for many comparison blocks within short time

Event-Related Design Issues

EMORY ROLLINS SCHOOL OF PUBLIC HEALTH

CBIS Center for Biomedical Imaging Statistics

- Slow: Waiting 12+ seconds in between each event to allow HRF to relax is inefficient
- Gap spacing >4 seconds to avoid HRF blurring
- Jitter spacing to record different parts of the HRF and avoid correlation with other functions like heart rate and breathing



Task-related activity

- How to capture task-related activity in a noisy brain? Use cognitive subtraction/contrasts (task vs. control)

Image of task A

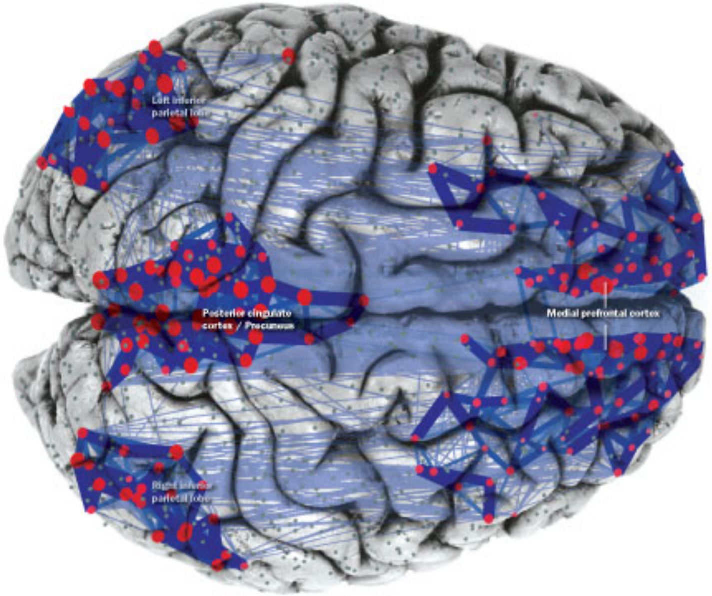
Image of task B

Image of task A- Image of task B

Resting-state fMRI

- **Resting-state fMRI studies:**
  - **No task**/stimulus
  - Acquire scans while subjects are left to think for themselves
  - May reflect a natural or more common mode of neural processing

Brain is **not** silent at rest!

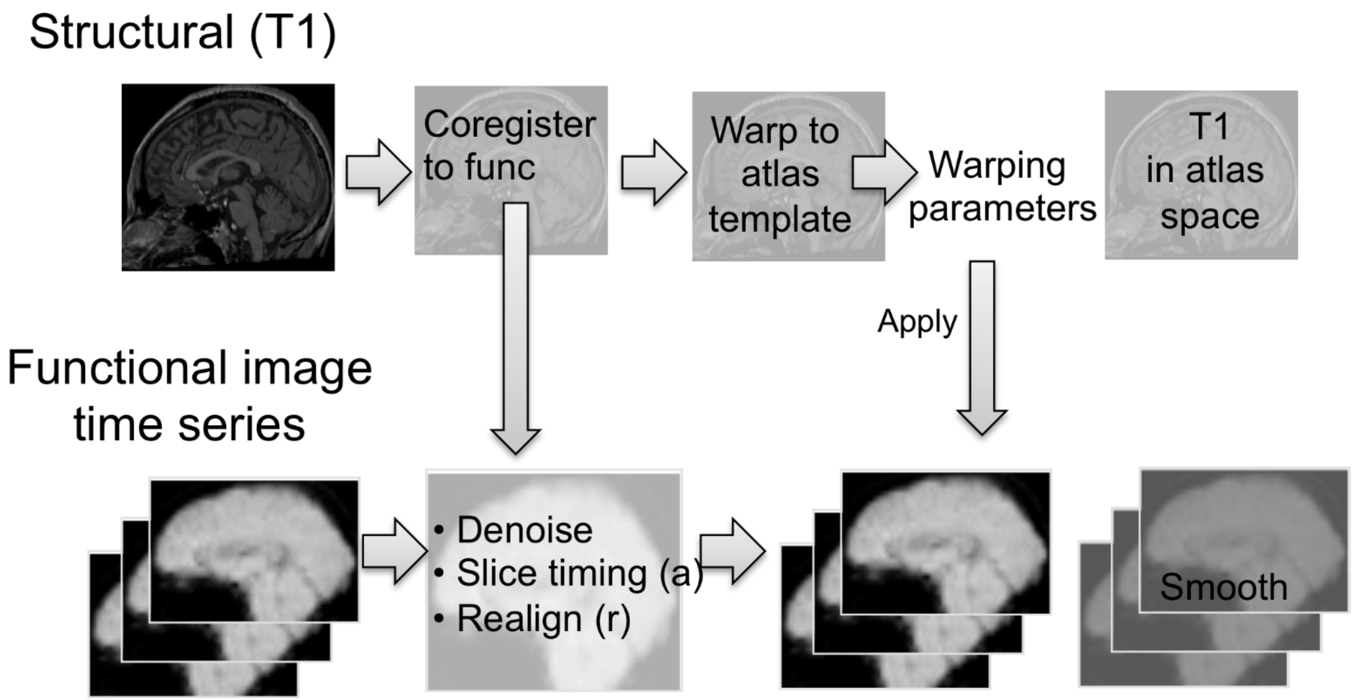


Default Mode Network

Sources of Noise

- Noisy brain:
  - Random neural activity
  - signal of interest is relatively weak
- Noisy scanner:
  - Scanner Drift – the magnetic field can slowly rise and fall
  - Non-uniformities in magnetic field
- Physiological noise:
  - head/brain movement due to heartbeat, breathing, subject fidgeting, etc.
- Solutions:
  - Limit subject movement in the scanner
  - **Preprocessing** steps to minimize artifacts and standardize <u>before</u> conducting further analysis

Preprocessing Pipeline



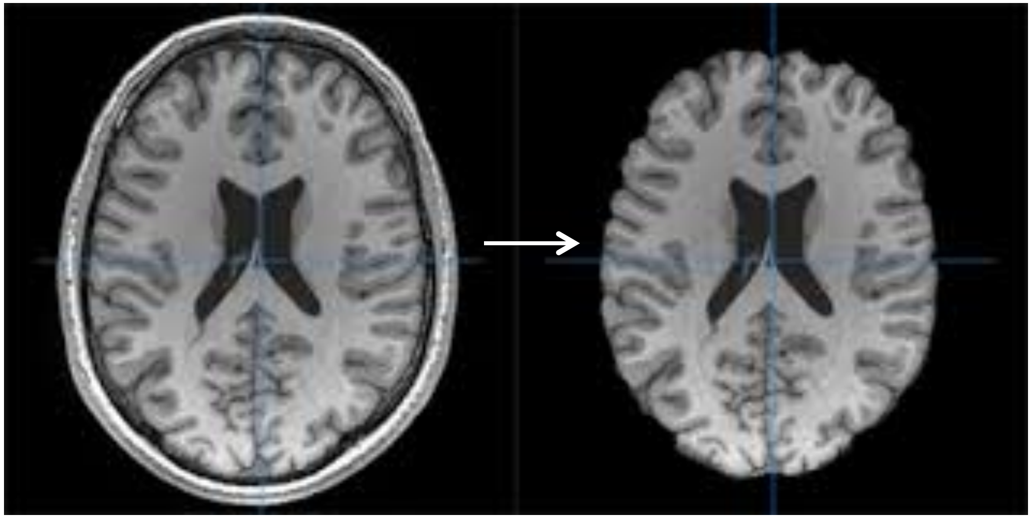
- Preprocessing is performed both on the fMRI data and structural (MRI) scans, collected prior to the experiment.

Preprocessing Steps

- Brain Extraction
- Slice timing correction
- Motion correction
- Co-registration
- Normalization
- Spatial Filtering/Smoothing
- Temporal Filtering

Brain Extraction

- Remove non-brain tissue and skull from the image, so that we only use voxels located in the brain.
- Easy to implement with brain extraction tool (BET) in FSL, or 3dSkullStrip in AFNI

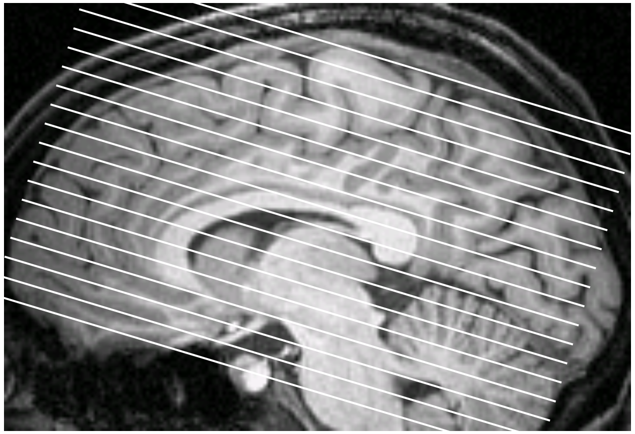


Preprocessing Steps

- Brain Extraction
- Slice timing correction
- Motion correction
- Co-registration
- Normalization
- Spatial Filtering/Smoothing
- Temporal Filtering

Slice Timing Correction

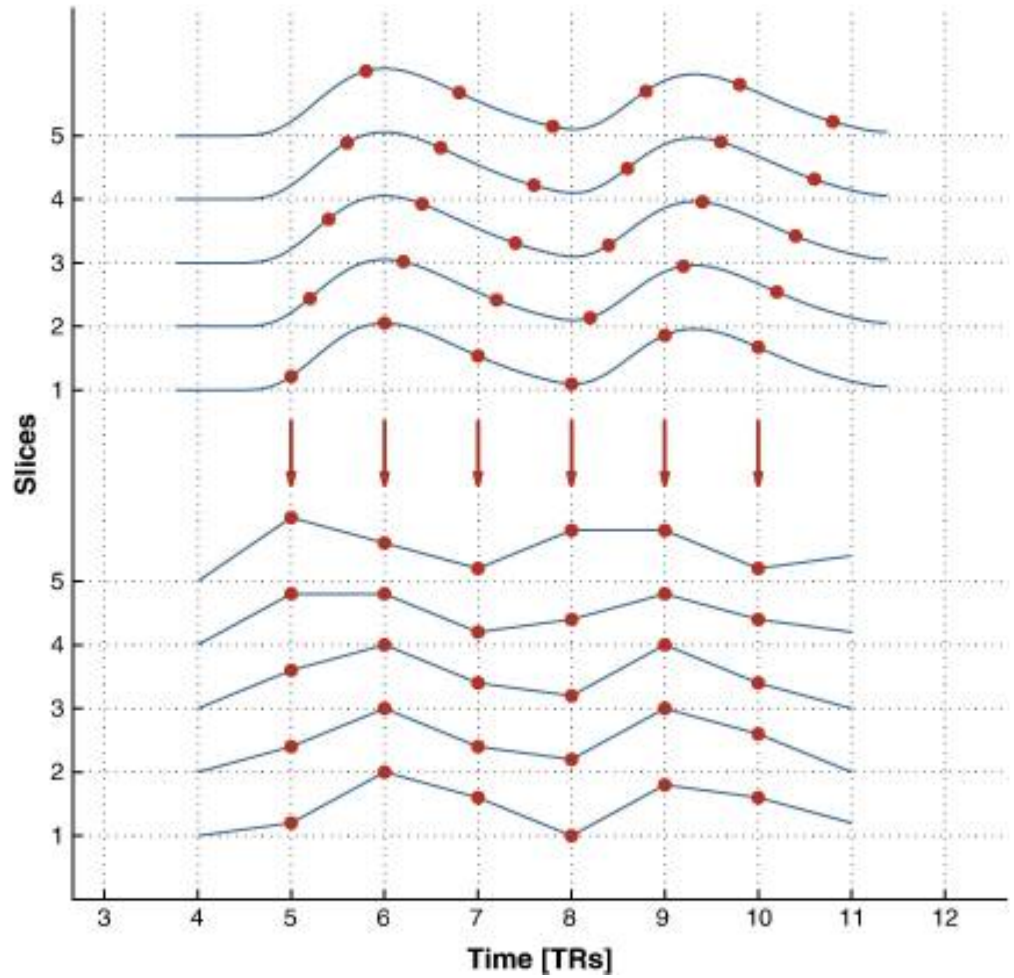
- We often sample multiple slices of the brain during each repetition time (TR) to construct a brain volume.
- Each slice is sampled at slightly different time points.
- 2D slices → 3D brain volume



Axial slices

Slice Timing Correction

- **Slice timing correction** shifts each voxel's time series so that they all appear to have been sampled simultaneously.

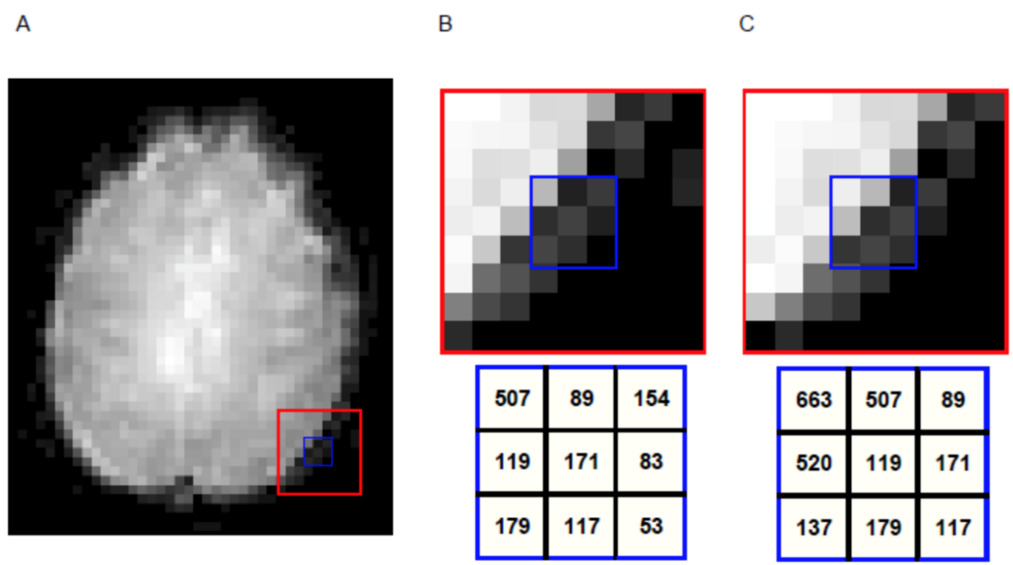


Preprocessing Steps

- Brain Extraction
- Slice timing correction
- Motion correction
- Co-registration
- Normalization
- Spatial Filtering/Smoothing
- Temporal Filtering

Head Motion

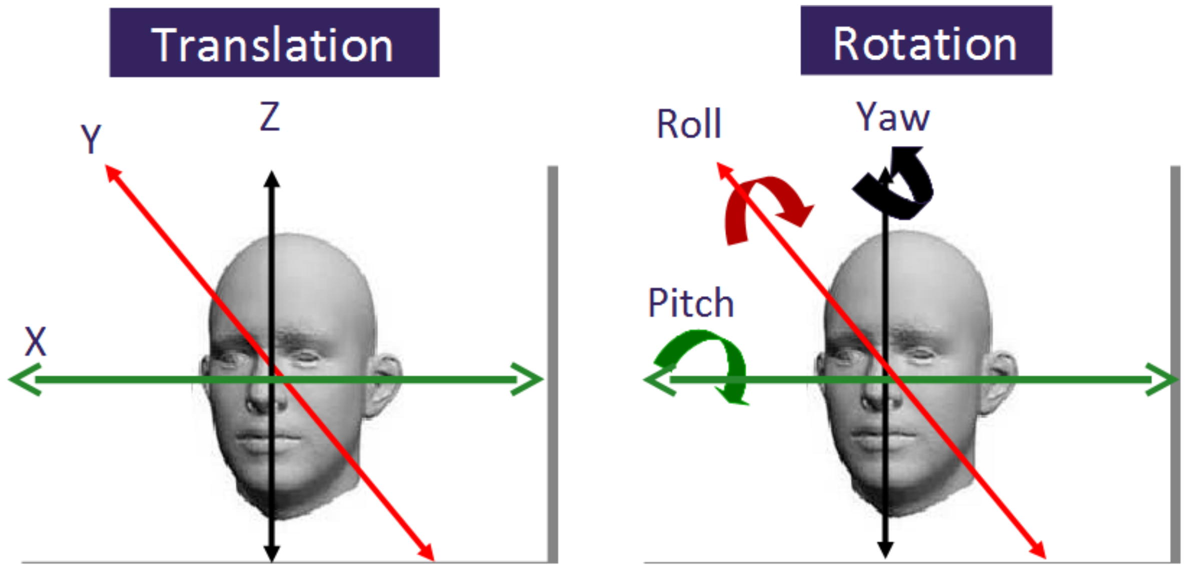
- Small head movements during a scan can be a major source of error if not treated correctly.
- When analyzing a voxel's time series, we assume that the voxel represents the same location in the brain at every time point.
  - Head motion may make this assumption incorrect



Huettel et al. Functional Magnetic Resonancel Imaging

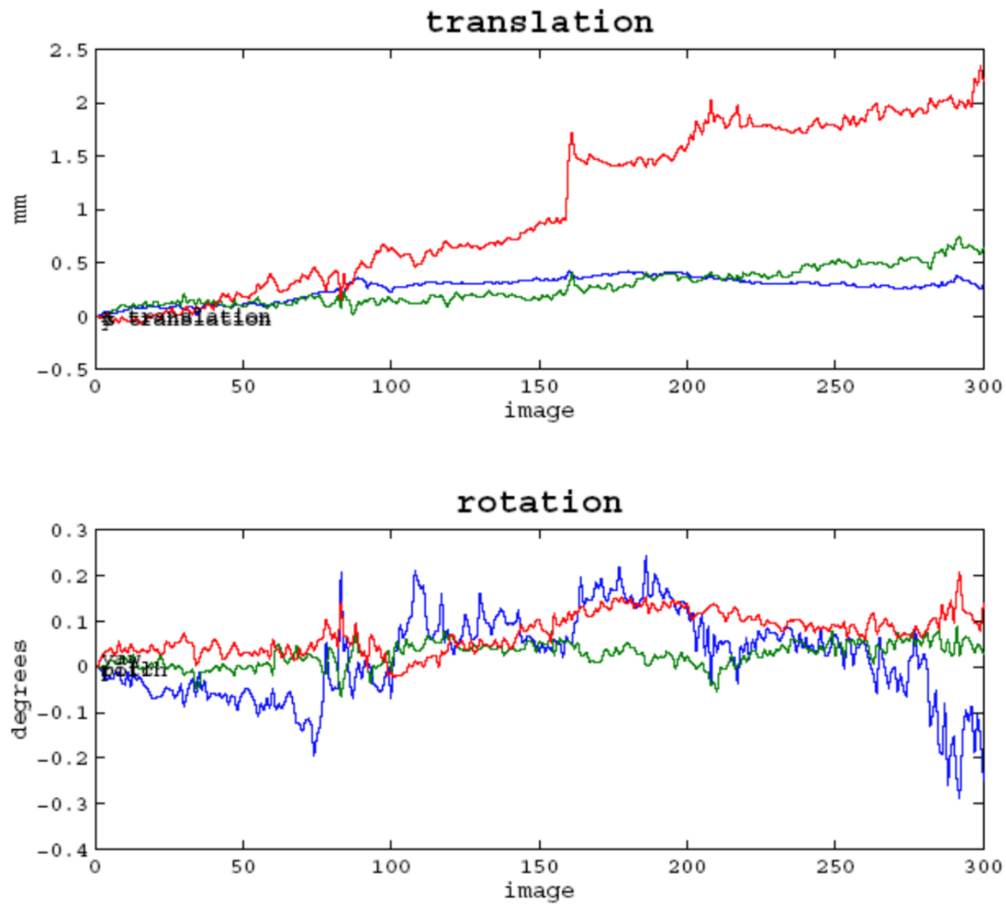
Motion Correction

- Motion can be corrected using a **rigid body transformation**:
  - Choose a reference volume to register all the other volumes to. (e.g. first volume, middle volume for FSL)
  - Re-aligns to reference volume to minimize variance
  - 6 DOF: translation (x, y, z) and rotation (roll, pitch, yaw)



Motion Correction

EMORY
ROLLINS SCHOOL OF PUBLIC HEALTH
CBIS
Center for Biomedical Imaging Statistics



Preprocessing Steps

- Brain Extraction
- Slice timing correction
- Motion correction
- Co-registration
- Normalization
- Spatial Filtering/Smoothing
- Temporal Filtering

Coregistration

- Functional MRI (T2) image has low spatial resolution
- It is common to map the results obtained from fMRI onto a high-res structural MRI (T1) image, collected at the start of the scanning session.
- The process of aligning the structural and functional image is called **coregistration**
  - Affine transformation (12 DOF) – translation, rotation, as well as scaling and shearing (non-linear allows for warping)

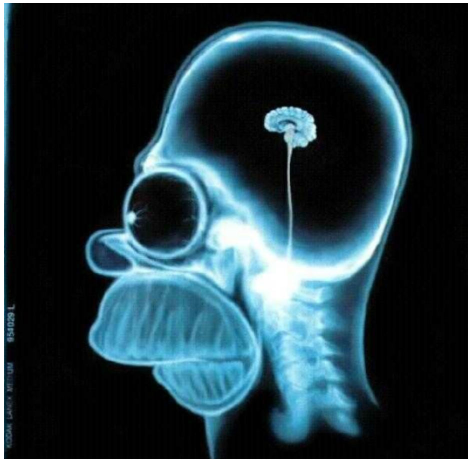
Preprocessing Steps

- Brain Extraction
- Slice timing correction
- Motion correction
- Co-registration
- **Normalization**
- Spatial Filtering/Smoothing
- Temporal Filtering

Normalization

EMORY ROLLINS SCHOOL OF PUBLIC HEALTH | CBIS Center for Biomedical Imaging Statistics

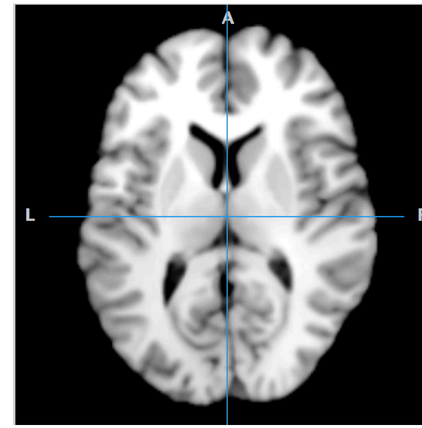
- Everyone's brain is different. The brain size of subjects can differ in size by up to 30%!



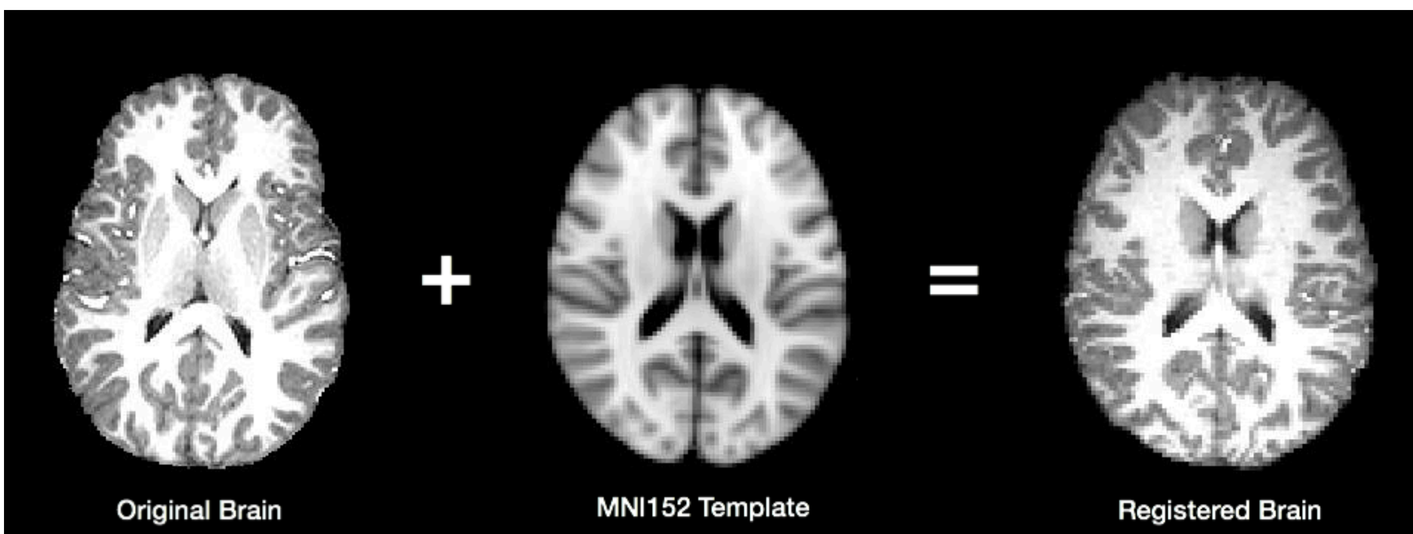
- There is also substantial variation in brain shapes
- **Normalization** attempts to register each subjects anatomy to a standard coordinate space defined by a **template brain**
  - Affine transformation (12 DOF)

Standard Brain Templates

- **Talairach**
  - Talairach and Tournoux (1988)
  - Based on dissection and photography of a single subject (cadaver of a 60 y.o female)
- **MNI** (Montreal Neurological Institute)
  - Based on MRI scans of hundreds of normal controls (all RH)



Talairach Template



Preprocessing Steps

- Brain Extraction
- Slice timing correction
- Motion correction
- Co-registration
- Normalization
- Spatial Filtering/Smoothing
- Temporal Filtering

Spatial Smoothing

- **Spatial smoothing of fMRI data**: improves inter-subject registration and overcomes limitations in spatial normalization by blurring any residual anatomical differences.

- **PROs**: can increase SNR by decreasing variance and remove artifacts
- **CONs**: may reduce signal if small activations; reduces spatial resolution

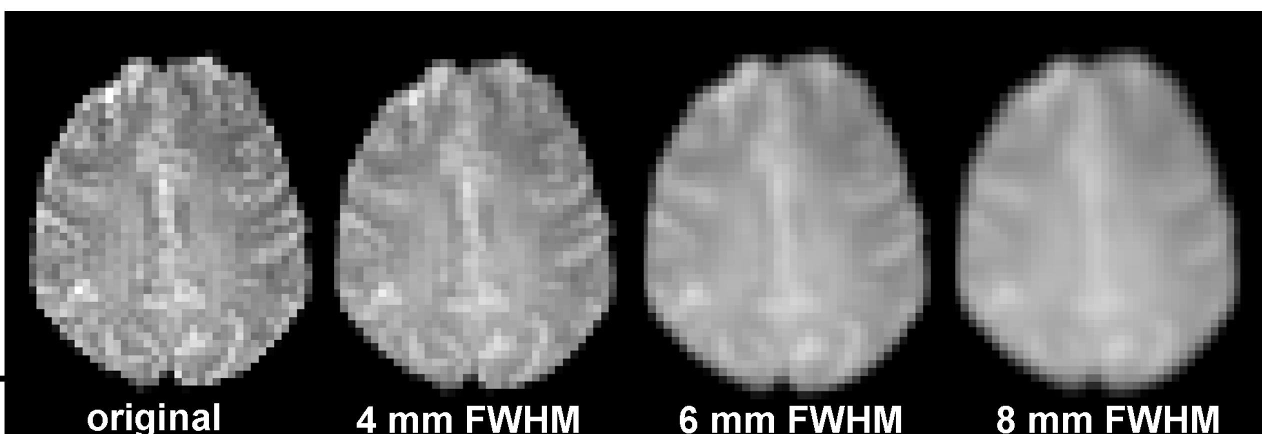
Spatial Smoothing

- Average one voxel's values with its neighbors
- Gaussian Full Width Half Maximum (FWHM) kernel
  - Each voxel intensity is replaced by a weighted average of neighboring intensities
  - Gaussian function specifies weightings and neighborhood size
  - Usu. 4-12 mm FWHM

Weights

0.1	0.3	0.4	0.3	0.1
0.3	0.6	0.8	0.6	0.3
0.4	0.8	1.0	0.8	0.4
0.3	0.6	0.8	0.6	0.3
0.1	0.3	0.4	0.3	0.1

FWHM

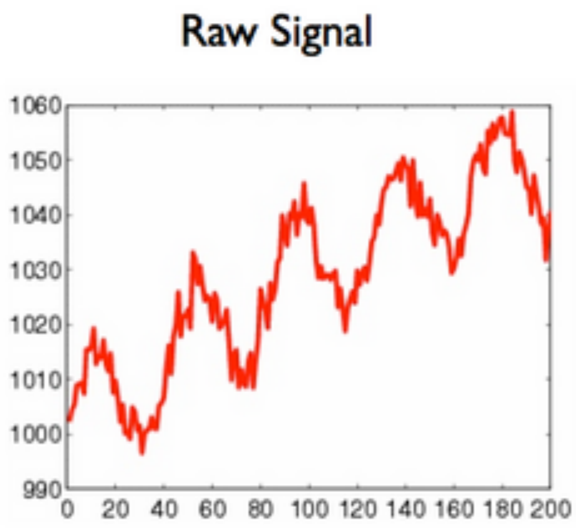


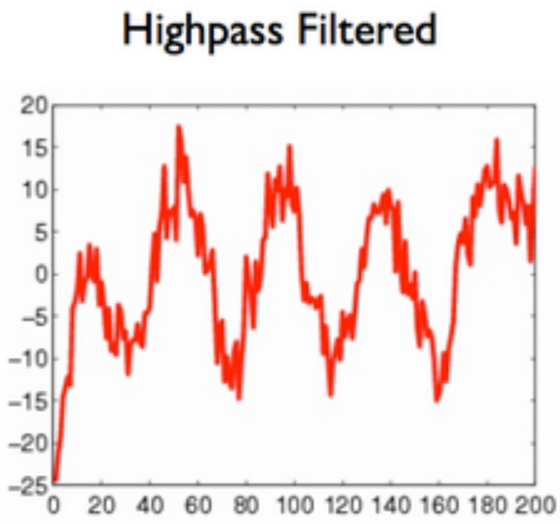
Preprocessing Steps

- Brain Extraction
- Slice timing correction
- Motion correction
- Co-registration
- Normalization
- Spatial Filtering/Smoothing
- Temporal Filtering

Temporal Filtering

- Temporal noise due to drift from scanner, subject's heartbeat and breathing
- These can mask your actual signal!
- Use a high-pass filter to remove low frequency (i.e. long, slow) noise



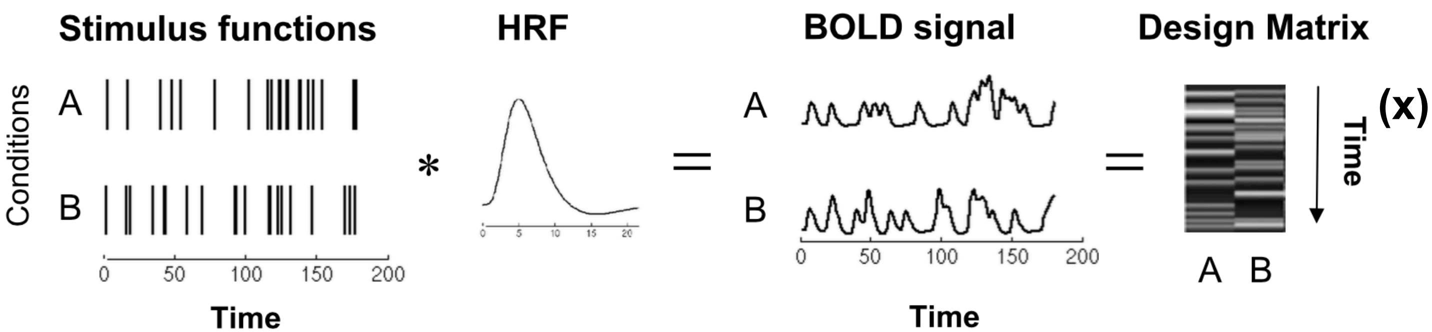


Statistical Analysis

- After the images have been preprocessed, we can begin statistical analysis!
- Goals of statistical analysis of fMRI data:
  - **ACTIVATION**: Localizing brain areas activated by the experimental task
  - **FUNCTIONAL CONNECTIVITY**: Determining networks corresponding to brain function
  - **PREDICTION**: making predictions about psychological or disease states

Activation

- **Goal**: identify regions that are active during a specific task or related to a certain behavioral measure
- Step 1: construct a model for each voxel
  - "Massive univariate approach"
  - Regression models (GLM) commonly used

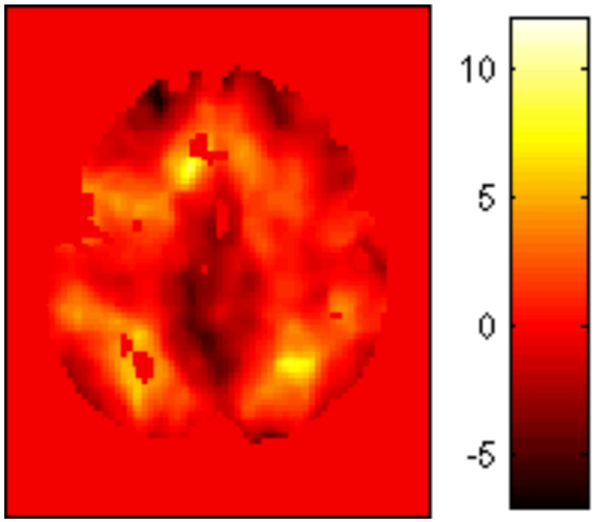


$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \qquad \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{V})$$

Activation

- Step 2: perform a statistical test to determine whether task-related activation is present in each voxel

$$H_0 : \mathbf{c}^T \boldsymbol{\beta} = 0$$

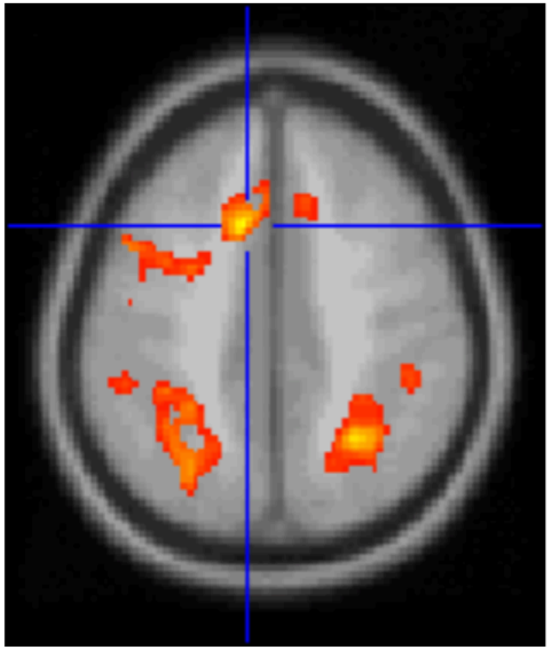


For contrast c, task vs. control

Statistical map: map of t-test statistics across all voxels (a.k.a. t-map)

Activation

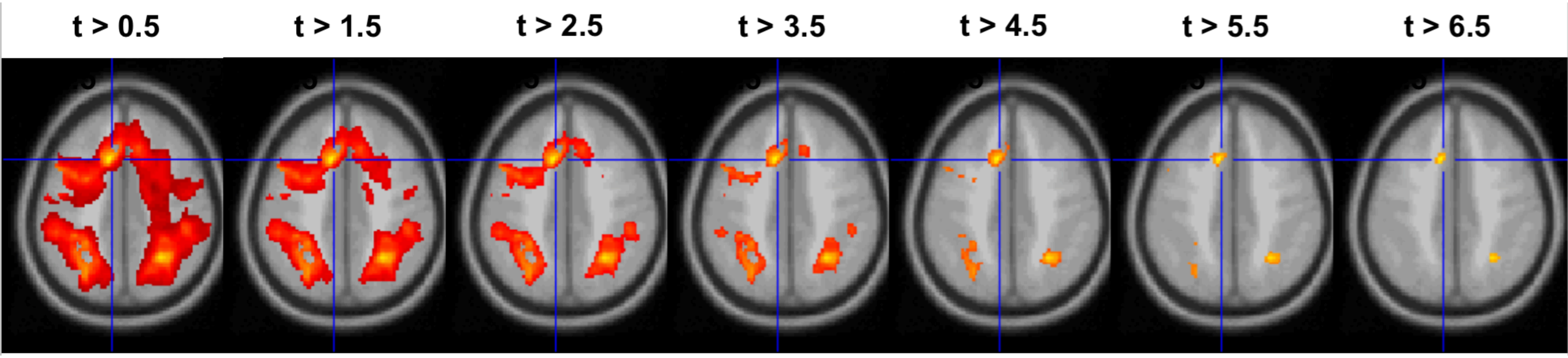
- Step 3: Choose an appropriate threshold for determining statistical significance



Thresholded t-map:
Each significant voxel is color-coded according to the size of its p-value

Multiple Comparison Problem

- Which of 100,000 voxels are significant?
  - α=0.05 → 5,000 false positive voxels
- Bonferroni correction is overly conservative
- Choosing a threshold is a balance between sensitivity (true positive rate) and specificity (true negative rate)

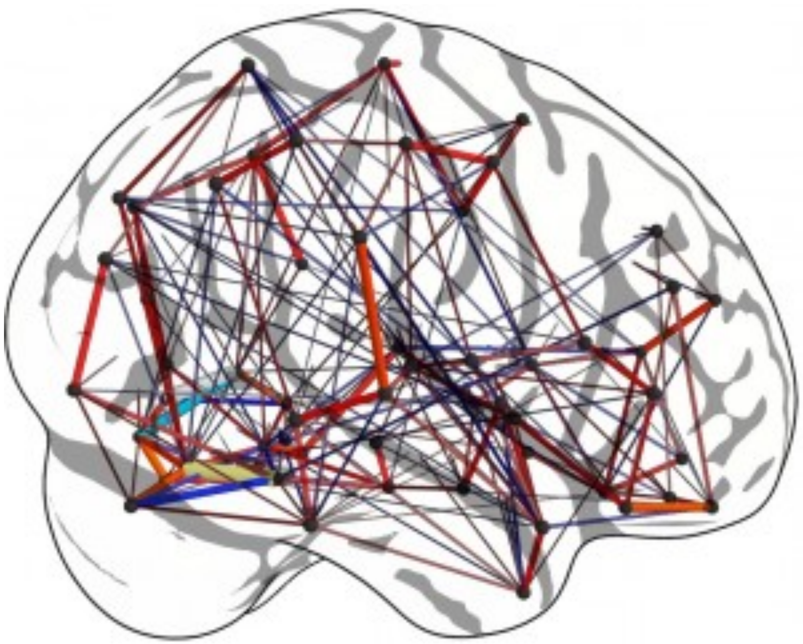


Statistical Analysis

- Goals of statistical analysis of fMRI data:
  - **ACTIVATION**: Localizing brain areas activated by the experimental task
  - **FUNCTIONAL CONNECTIVITY**: Determining networks corresponding to brain function
  - **PREDICTION**: making predictions about psychological or disease states

Connectivity

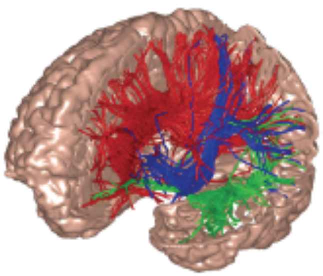
- Recently, there has been increased interest in augmenting activation analyses with **connectivity studies**, which describe how various brain regions interact.



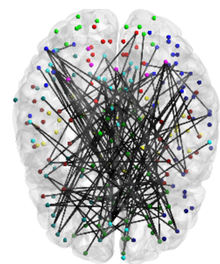
Connectivity

BRAIN CONNECTIVITY

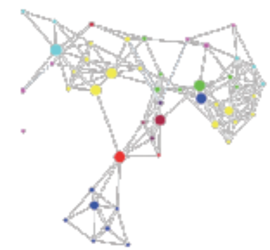
- **Structural connectivity**
  - Diffusion MRI tractography



- **Functional connectivity**
  - seed-based analysis, graphical models, ICA

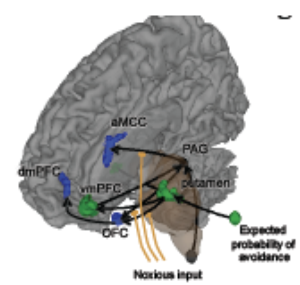


Wang et al. 2016



Wager et al. 2015 graphical model

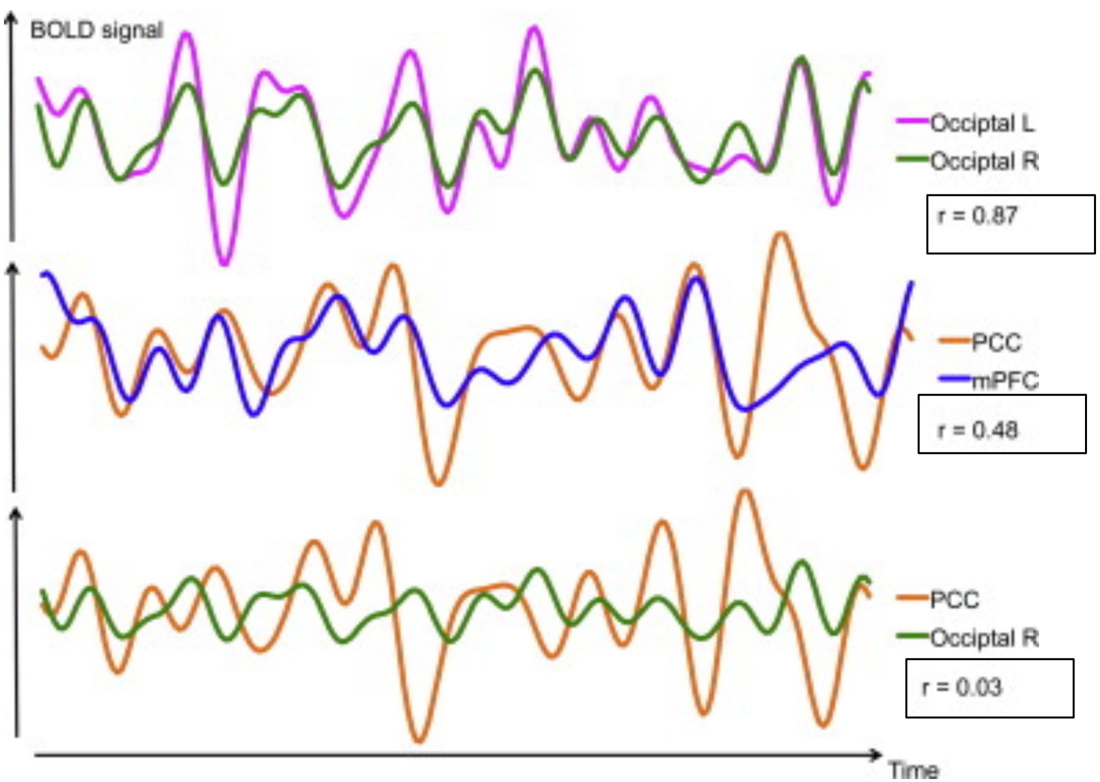
- **Effective connectivity**
  - Granger causality, Dynamic Causal Modeling (DCM)
- **Dynamic connectivity**
  - sliding window, hidden Markov model, change-point method



Roy et al. 2014 DCM

Functional Connectivity

- Temporal coherence of brain regions
- Undirected association (usu. correlation)



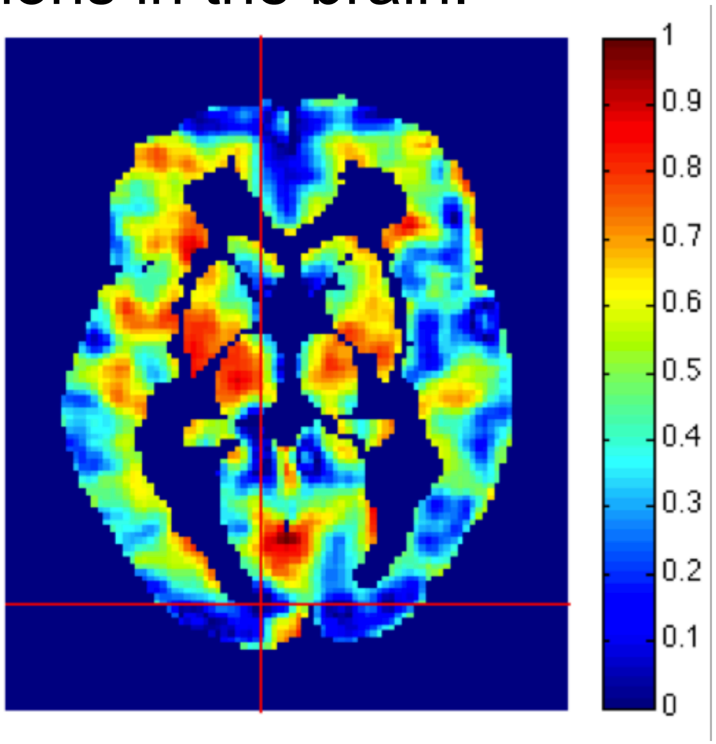
Functional Connectivity

- **Functional Connectivity Analysis** is usually performed using data-driven methods which make no assumptions about the underlying biology

- Methods include:
  - Seed analysis
  - Network analysis
  - Partitioning methods: Clustering, PCA, ICA

Seed Analysis

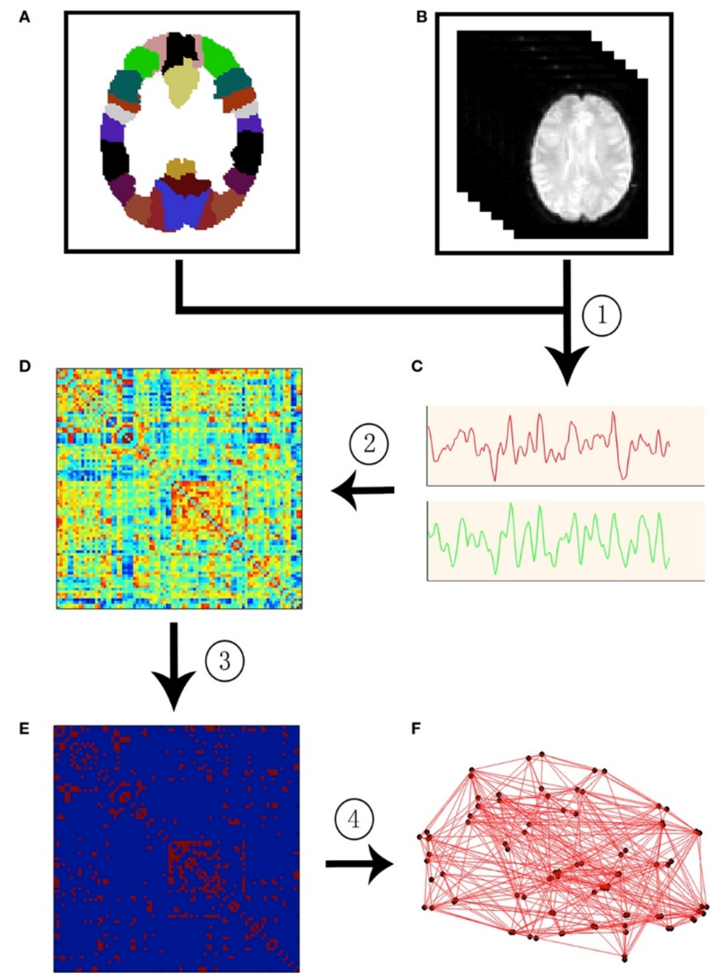
- Calculates the correlation between the temporal brain activity profile in a selected ("seed") voxel/region and the profiles from other voxels/regions in the brain.
- Simple and easy to implement
- BUT… requires careful selection of seed voxel/region
- Provides a limited view of the brain, since it is restricted to connectivity involving the seed voxel.



Network Analysis

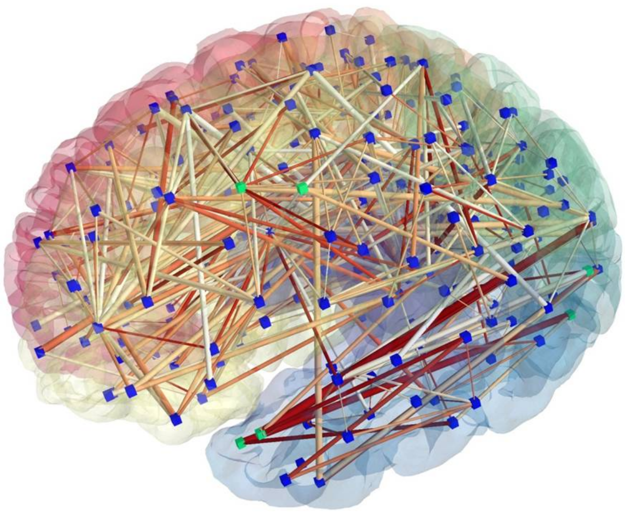
Whole-Brain region-to-region approach:

1. Parcellate the brain and extract the “average” fMRI time course for each region
2. Calculate correlation between regions → correlation matrix
3. Threshold → binary adjacency matrix
4. Graph theory analysis



Network Analysis

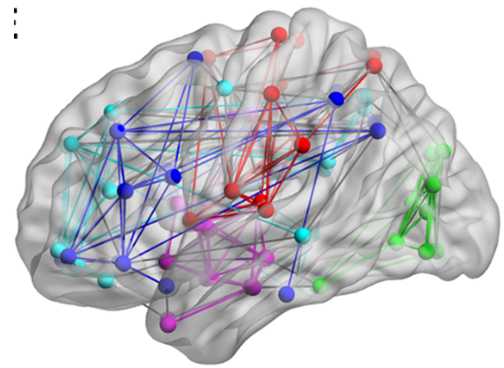
- **Network/Graph Theory analysis** tries to characterize networks using a small number of meaningful summary measures
- Comparing network topological measures (ex: node degree, clustering coefficient, etc.) between groups of subjects may reveal connectivity abnormalities related to brain disorders



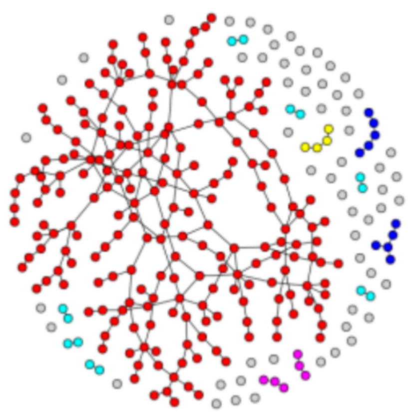
A network is a system of **nodes** (regions) and **edges** (connections between regions)

Network Analysis

- Network Visualization tools
  - BrainNet MATLAB toolbox



  - igraph R package



- Brain connectivity toolbox (MATLAB) for calculating graph theory metrics to characterize networks.

Partitioning Algorithms

- **Partitioning algorithms** identify spatially distinct components or clusters in the brain
- Each of these components represents a functionally connected network
- Methods:
  - Clustering
  - PCA
  - ICA

Clustering

- **Cluster analysis:** identifies “clusters” of voxels with similar brain activity patterns.
- Clusters may consist of noncontiguous voxels, offering the potential of identifying associations between anatomically distant voxels
- Several algorithms: K-means approach, fuzzy clustering, hierarchical clustering, etc.

The image cannot be displayed. Your computer may not have enough memory to open the image, or the image may have been corrupted. Restart your computer, and then open the file again. If the red x still appears, you may have to delete the image and then insert it again.

3 clusters (red, orange, yellow) based on mean brain activity of cocaine addicts in inhibitory control study.

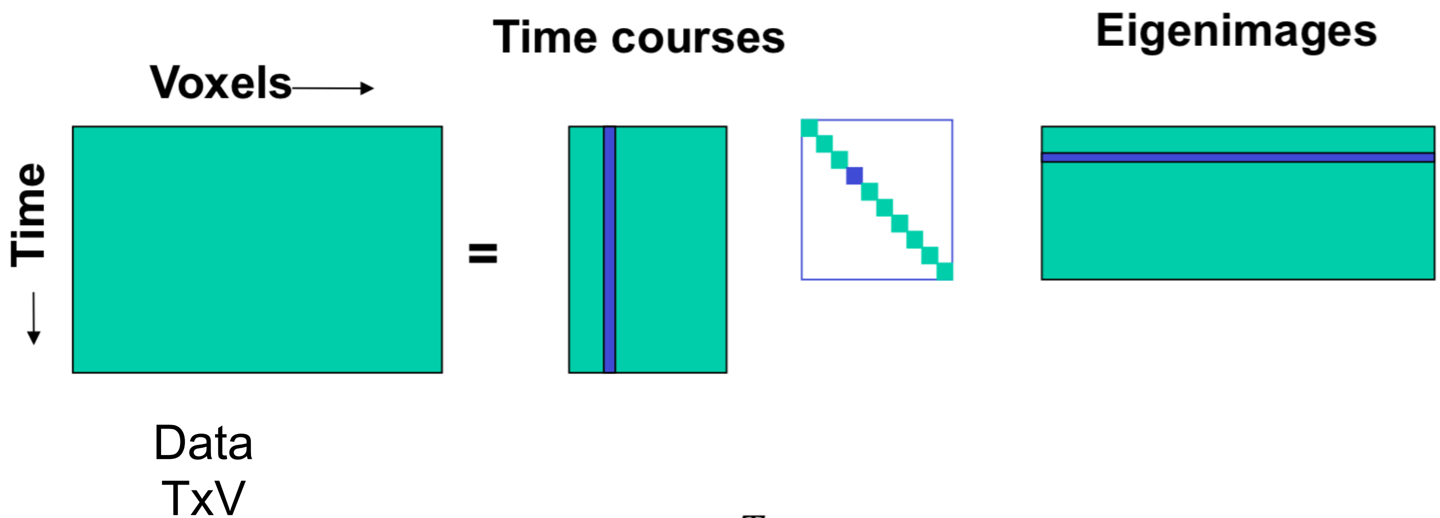
Each cluster contains voxels with similar patterns of brain activity

PCA

- **Principal components analysis (PCA)** involves finding spatial modes, or eigenimages, in the data
  - These are the patterns that account for most of the variance-covariance structure in the data, ranked in order

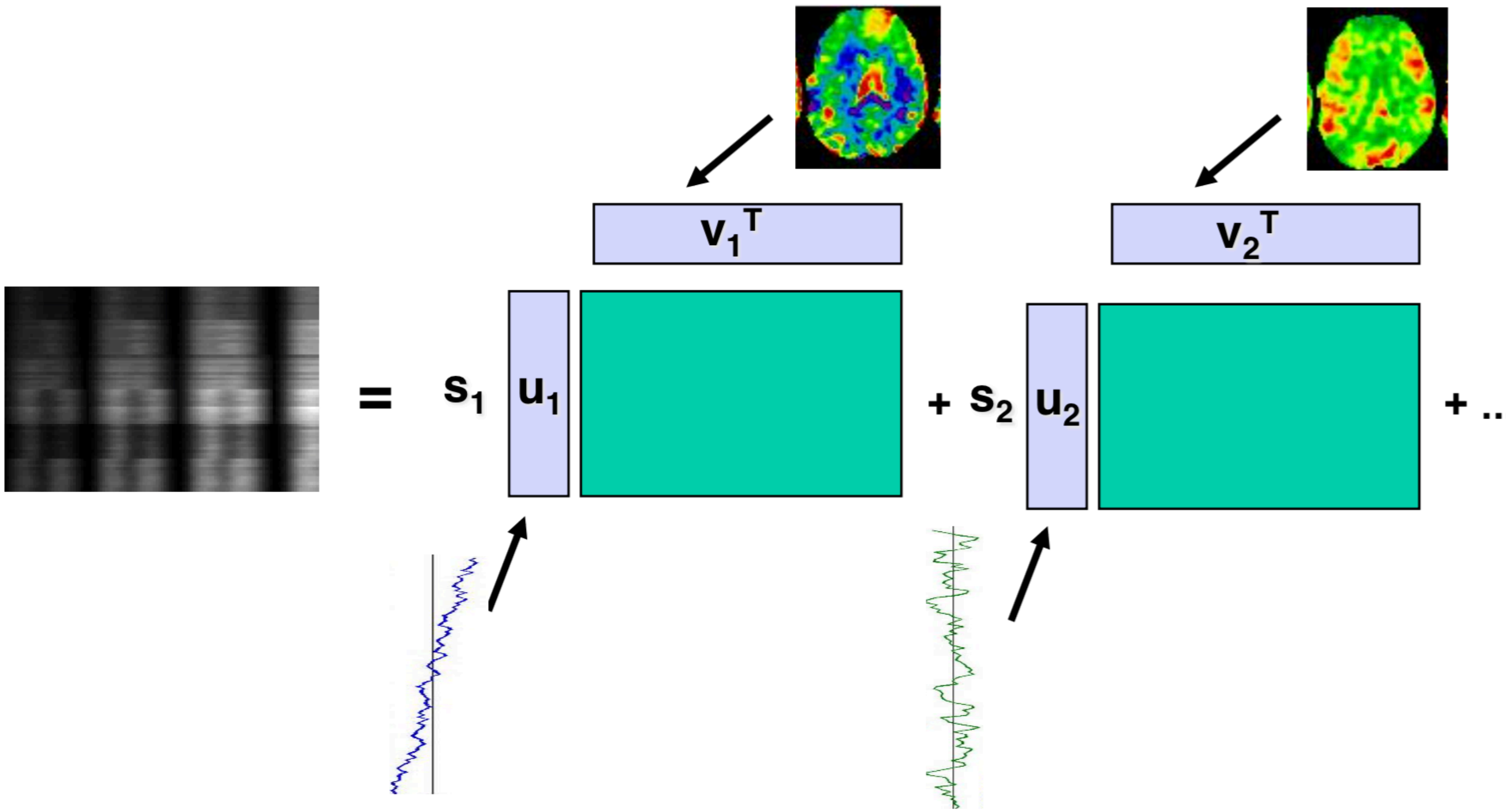
- The eigenimages can be obtained using singular value decomposition (SVD), which decomposes the data into two sets of orthogonal vectors that correspond to patterns in space and time.

PCA

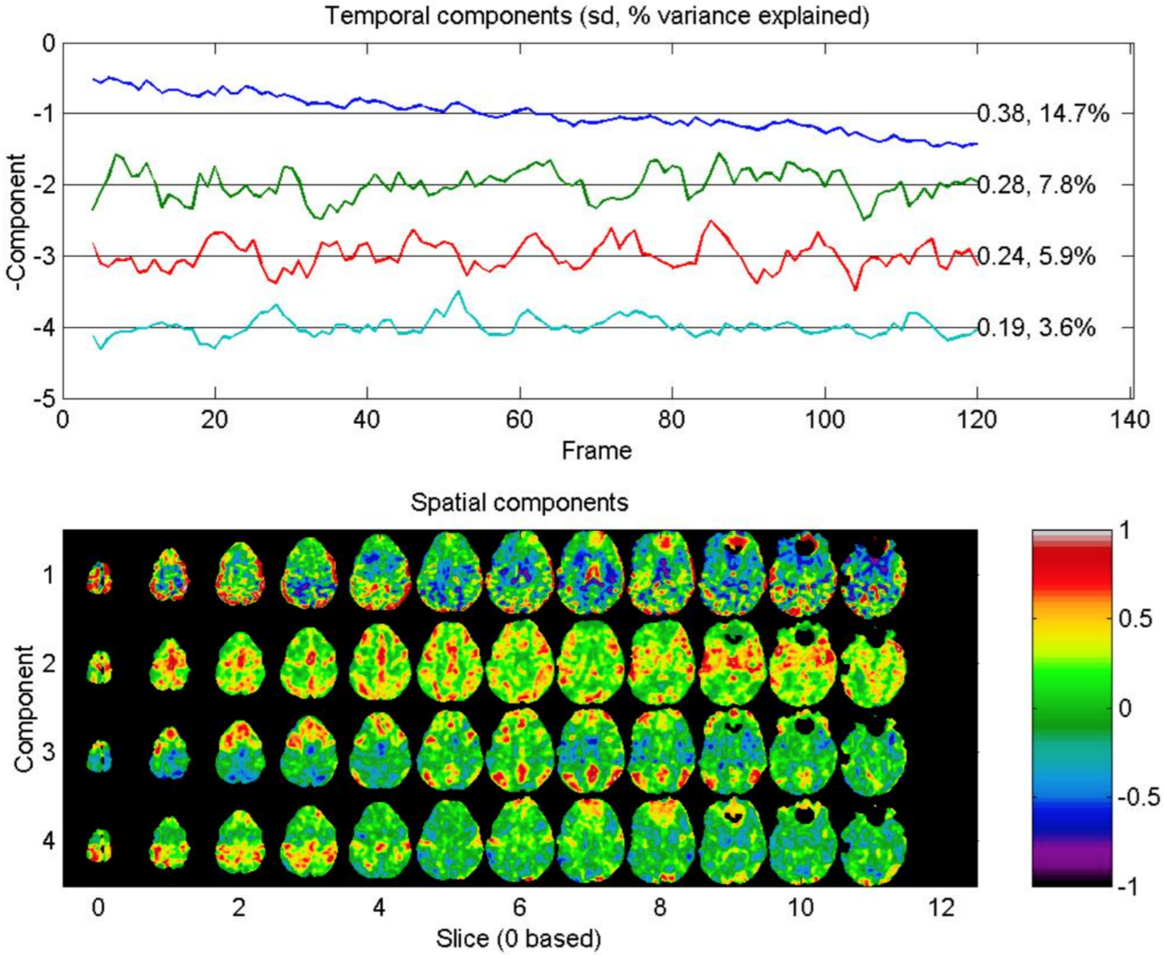


$$\mathbf{X} = \mathbf{U}\mathbf{S}\mathbf{V}^T$$

$$\mathbf{X} = s_1\mathbf{u}_1\mathbf{v}_1^T + s_2\mathbf{u}_2\mathbf{v}_2^T + \ldots + s_N\mathbf{u}_N\mathbf{v}_N^T$$


=
s_1
u_1
v_1^T
+ s_2
u_2
v_2^T
+ ...

PCA


Temporal components (sd, % variance explained)
-Component
Frame
0.38, 14.7%
0.28, 7.8%
0.24, 5.9%
0.19, 3.6%
0
-1
-2
-3
-4
-5
0
20
40
60
80
100
120
140
Spatial components
Component
Slice (0 based)
1
2
3
4
0
2
4
6
8
10
12
1
0.5
0
-0.5
-1

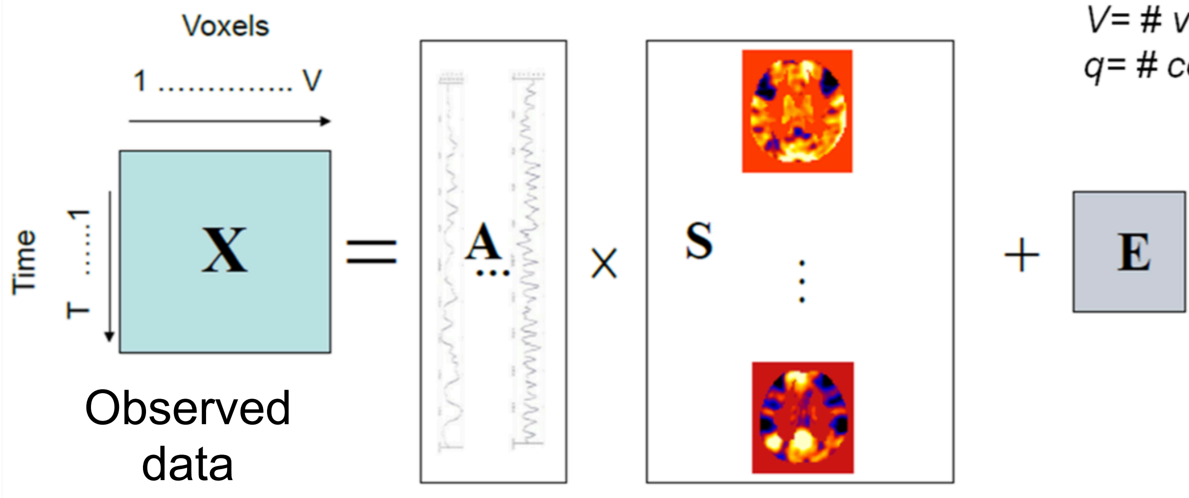
Worsley

ICA

- **Independent components analysis (ICA)** decomposes the observed fMRI signal to estimate q statistically independent component maps and their associated time courses.

$$\mathbf{X}_{T\times V} = \mathbf{A}_{T\times q}\,\mathbf{S}_{q\times V} + \mathbf{E}_{T\times V}$$

T= # time pts
V= # voxels
q= # components



Observed data (TxV)

Mixing Matrix (time courses)

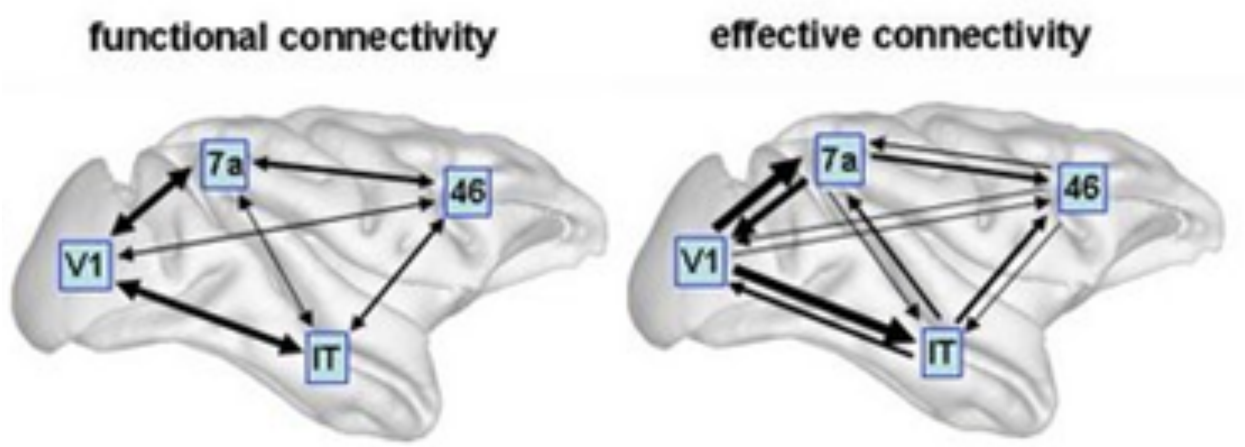
Component Maps

Advantages of ICA

- Does not require any apriori assumptions about the spatiotemporal structure underlying the observed brain activity
- Can be used for fMRI data with any paradigm; esp. useful for resting-state data where no clear task-related activations exist
- Simultaneously separates neuronal and non-neuronal sources (e.g. respiration) into different components
- ICA is more effective than PCA at identifying functional networks (Beckmann et al, 2005)
- Easy to extend to multi-subject case for group inference – use GIFT toolbox in Matlab

Effective Connectivity

- Directed influence of one brain region on the activity recorded in another brain region.



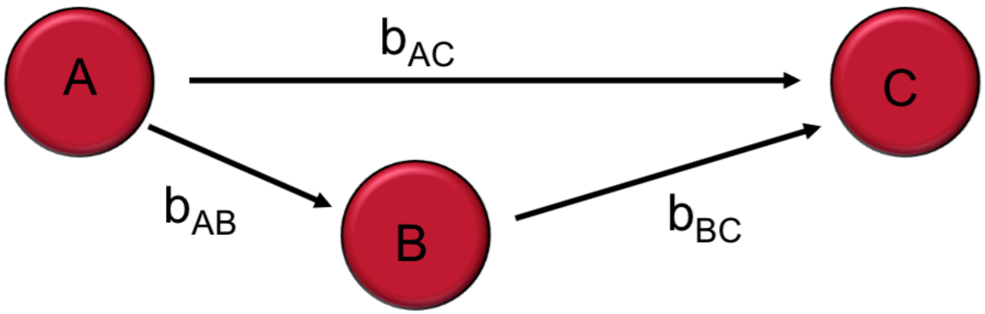
Undirected associations

Directed associations

- Methods: SEM, DCM, Granger Causality

SEM

- Structural Equation Models comprise a set of regions and a set of directed connections



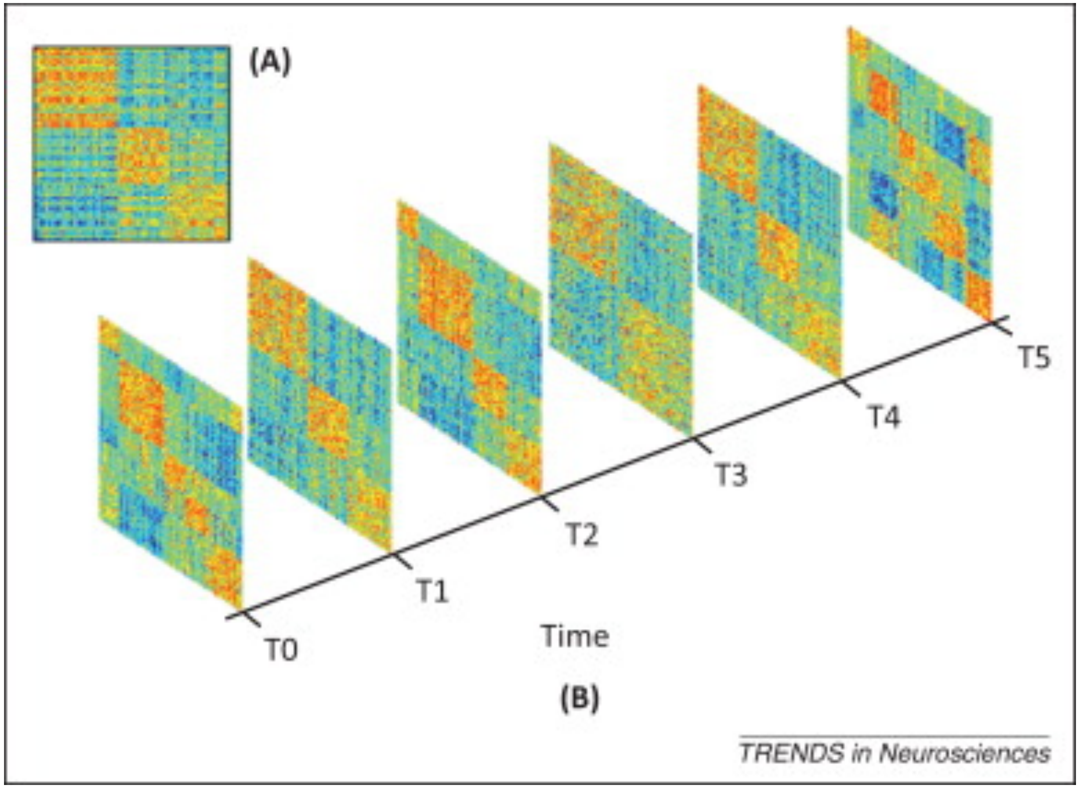
- Focuses on the covariance structure that reflects associations between variables

DCM

- Dynamic Causal Modeling estimates effective connectivity in a Bayesian framework.
- attempts to model latent neuronal interactions using hemodynamic time series
- Effective Connectivity is parameterized in terms of the coupling among unobserved neuronal activity in different regions.

Dynamic FC

- Dynamic FC attempts to model changes in FC over time
- Sliding Window approach



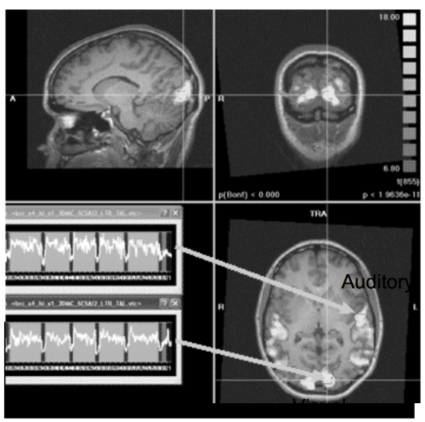
Statistical Analysis

- Goals of statistical analysis of fMRI data:
  - **ACTIVATION**: Localizing brain areas activated by the experimental task
  - **FUNCTIONAL CONNECTIVITY**: Determining networks corresponding to brain function
  - **PREDICTION**: making predictions about psychological or disease states

Prediction/Classification

- Predicting future neural activity based on baseline functional brain images.

- Predicting experiment conditions, cognitive states and group membership (psychiatric conditions, treatment response) based on functional brain images.



Experimental condition, cognitive states

Pittsburgh Brain Activity Interpretation Competition

Clinical outcomes:
•Diseased (e.g. ADHD) vs. normal
•Treatment Response vs. non-response

Prediction/Classification

- There is a growing interest in using fMRI data for classification of mental disorders and prediction of neural activity.

- This application of machine learning techniques is often referred to as multi-voxel pattern analysis (MVPA)
  - A classifier is trained to discriminate between different brain states and used to predict the states in a new set of data

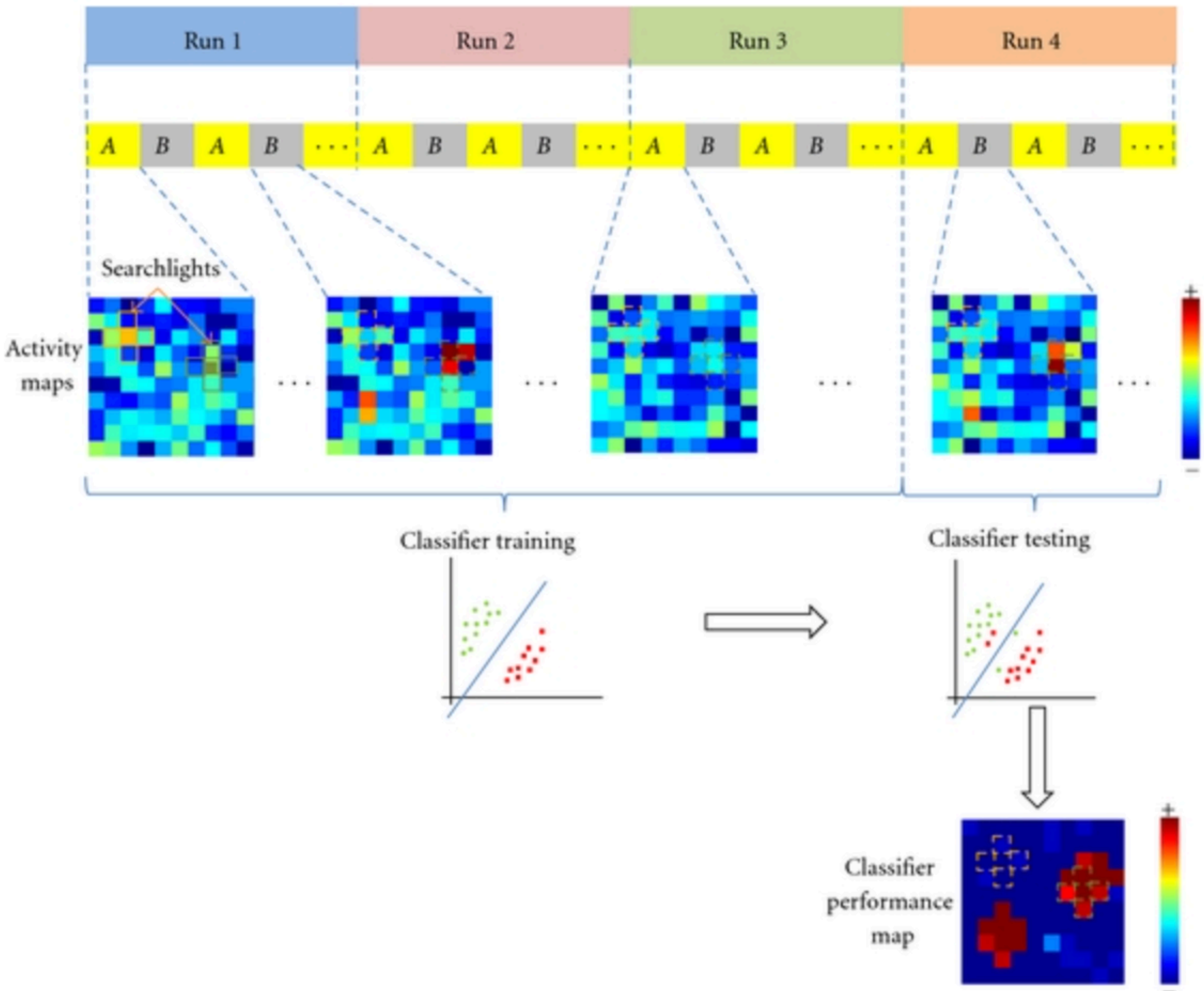
Machine Learning

- When applied to fMRI data, the result is often a pattern of weights across brain regions that quantify the degree to which the pattern of brain activity responds to a particular type of event. (Ex: SVM)

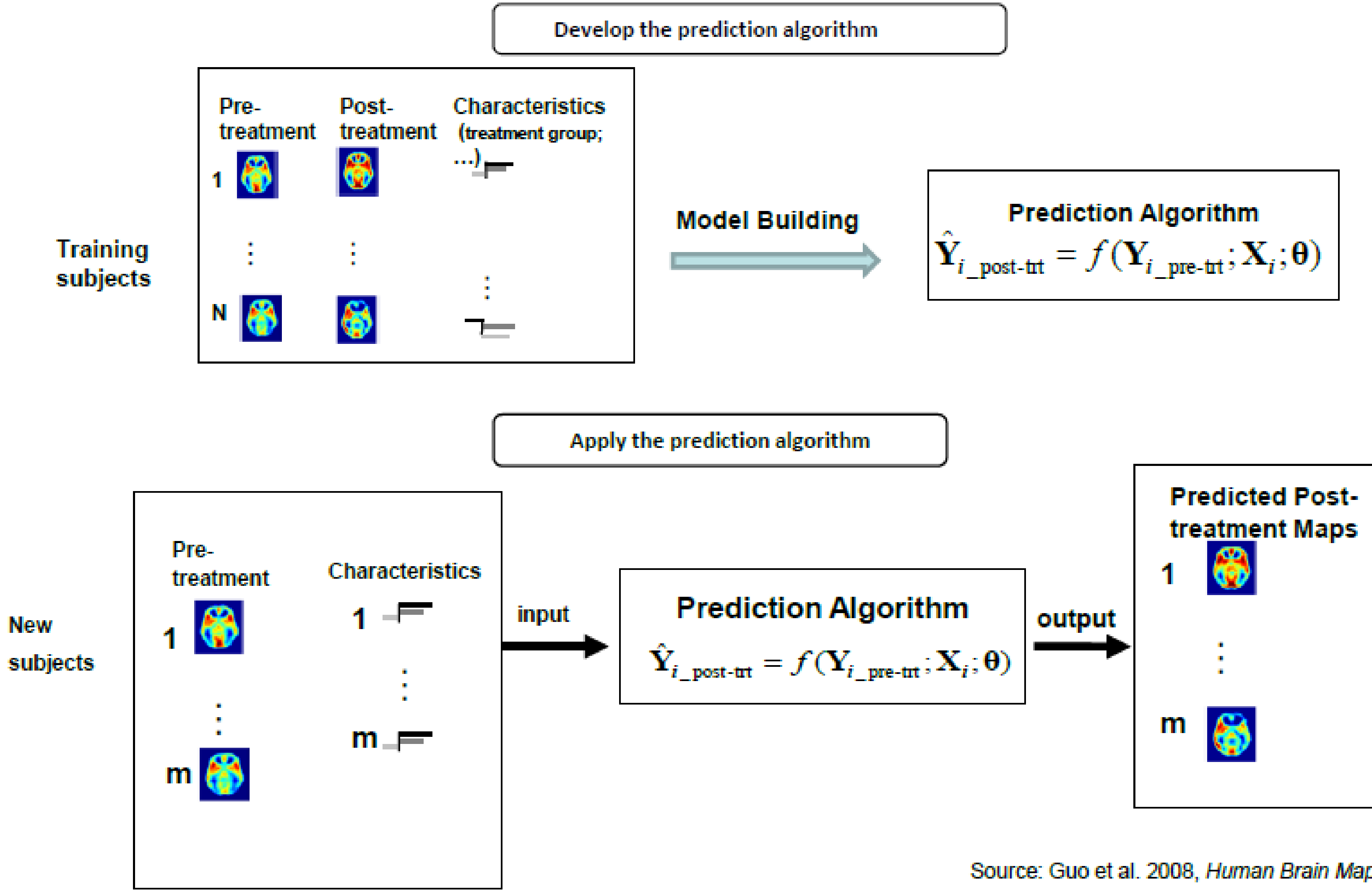
Performing MVPA

- The process of performing MVPA follows a series of steps:
  - Defining features and classes
  - Feature selection
  - Choosing a classifier
    - SVD, LDA, logistic regression
  - Training and testing the classifier
    - Cross validation
  - Examining results
    - Prediction accuracy

Performing MVPA



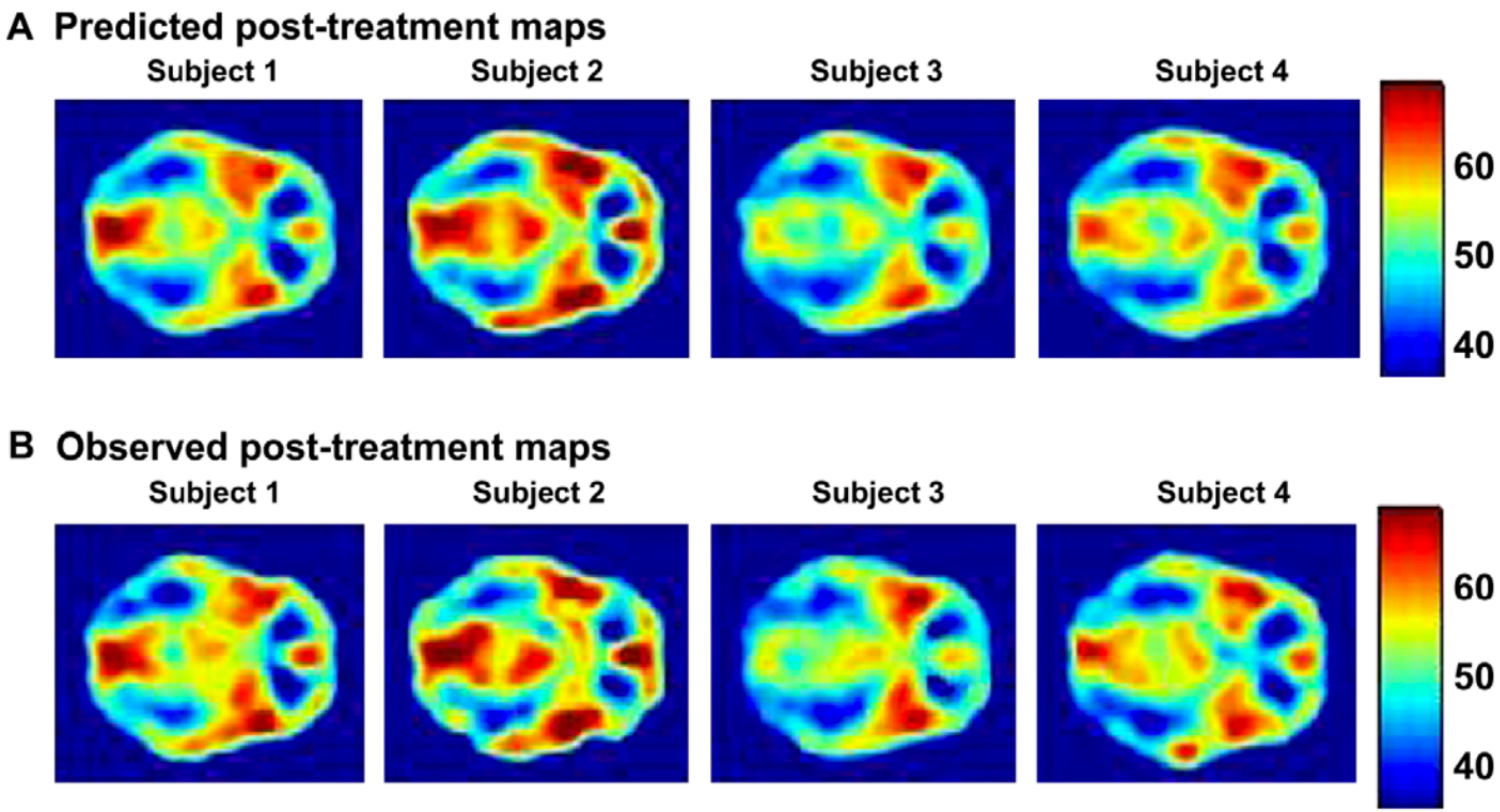
Predicting future neural activity



Source: Guo et al. 2008, *Human Brain Mapping*.

Prediction Example

Prediction of treatment response (Guo et al., 2008)



Prediction for clinical outcomes

imaging data

$\mathbf{Y}_i$

Feature selection

$\mathrm{h}(\mathbf{Y}_i)$

Building a predictive model

$f(\mathrm{h}(\mathbf{Y}_i), \mathbf{z}_i; \boldsymbol{\theta})$

Prediction, training and testing

$\mathbf{x}_i$

non - imaging outcomes

e.g.
1.Between-region functional connectivity (Shuo et al.);
2. cluster-specific principal features extracted using kernel PCA (Guo, 2010)

For categorical outcome: Support vector classifier

$$f(\mathrm{h}(\mathbf{Y}_i); \boldsymbol{\theta}) = sign[\mathrm{h}'(\mathbf{Y}_i)\boldsymbol{\beta} + \beta_0]$$

For continuous outcome: regression models such as ridge regression

$$f(\mathrm{h}(\mathbf{Y}_i); \boldsymbol{\theta}) = \beta_0 + \mathrm{h}'(\mathbf{Y}_i)\boldsymbol{\beta}$$

$$\text{with } \hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} \left\{ \sum_{i=1}^{N} \| \mathbf{x}_i - [\beta_0 + \mathrm{h}'(\mathbf{Y}_i)\boldsymbol{\beta}] \|^2 + \lambda \| \boldsymbol{\beta} \|^2 \right\}$$