ANALYSIS OF IMAGING DATA: FMRI

Mekibib Altaye, PhD.

Division of Biostatistics and Epidemiology

CCHMC

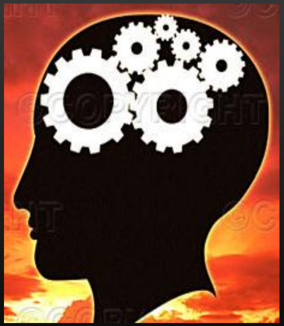
- What do we want to measure
  - Interested in neural activity of single neurons or ensembles of neurons
  - No direct way of measuring neural activity in normal humans or most cases of pathology
  - Instead, need <u>indirect measures</u> that are correlated with neural activity
- what is measured
  - popular neuroimaging methods measure <u>correlates</u> of blood flow and metabolism as a proxy for brain activity
  - Functional magnetic resonance imaging (fMRI)

- What is fMRI (Functional Magnetic Resonance Imaging)?
  - A noninvasive technique for measuring changes in brain activity over time using the principle of magnetic resonance.
  - Typically a series of brain images are acquired during which the subject performs a given set of tasks.
  - Changes in the measured signal are used to make inferences regarding task related brain activation.

Acquiring fMRI data

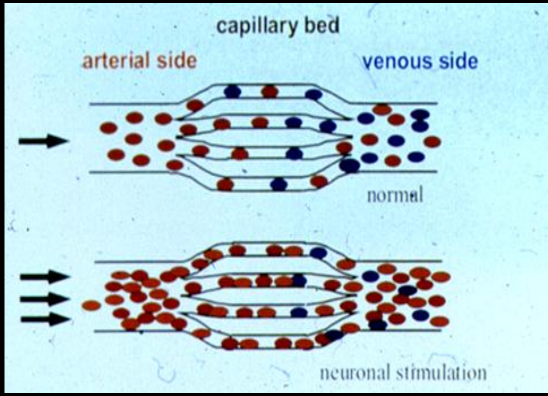
Magnetic Resonance Imaging Scanner

How does fMRI work?

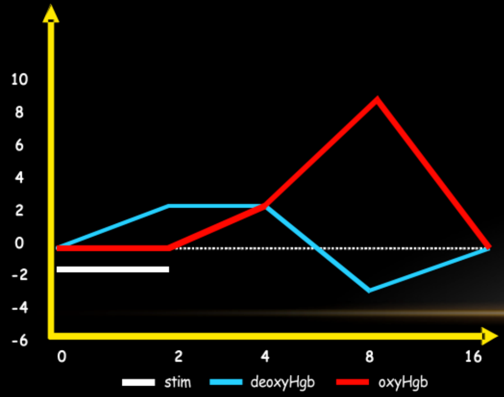


Brain stimulation

Local neuronal activity

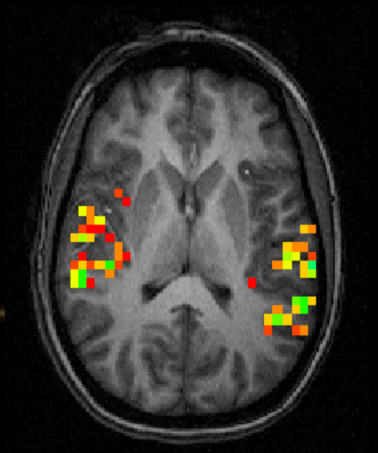


Local **increase** in blood flow



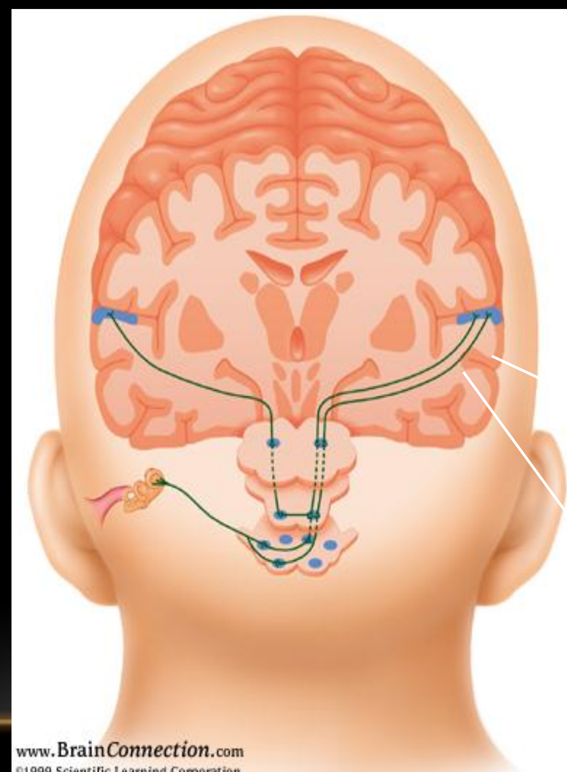
Increase in oxyhemoglobin

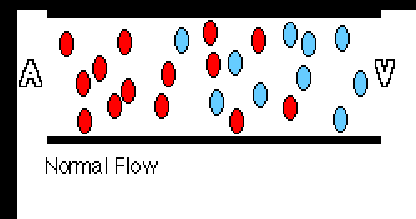
MRI pixel intensity change



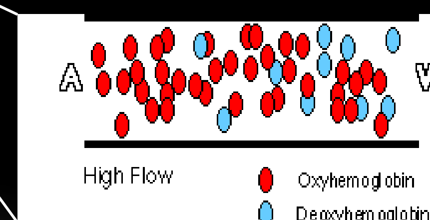
What is measured is BOLC Effect-Blood Oxygen Level Dependent MRI

- Brain activation causes local increase in blood flow
- The BOLD effect measures blood oxygenation in bulk neurons





at rest



stimulated

Afferent projections to auditory cortex

- Analysis of fMRI data is challenging because
  - it is NOISY!
  - Signal of interest represent only a small part of what is measured by the scanner.
  - The changes we are looking are small 3 to 5%
  - The signal is inherently correlated both in time and space
- One of the task here is to reduce noise and maximize signal detection
- How to reduce the noise partly depend on identification of source of noise

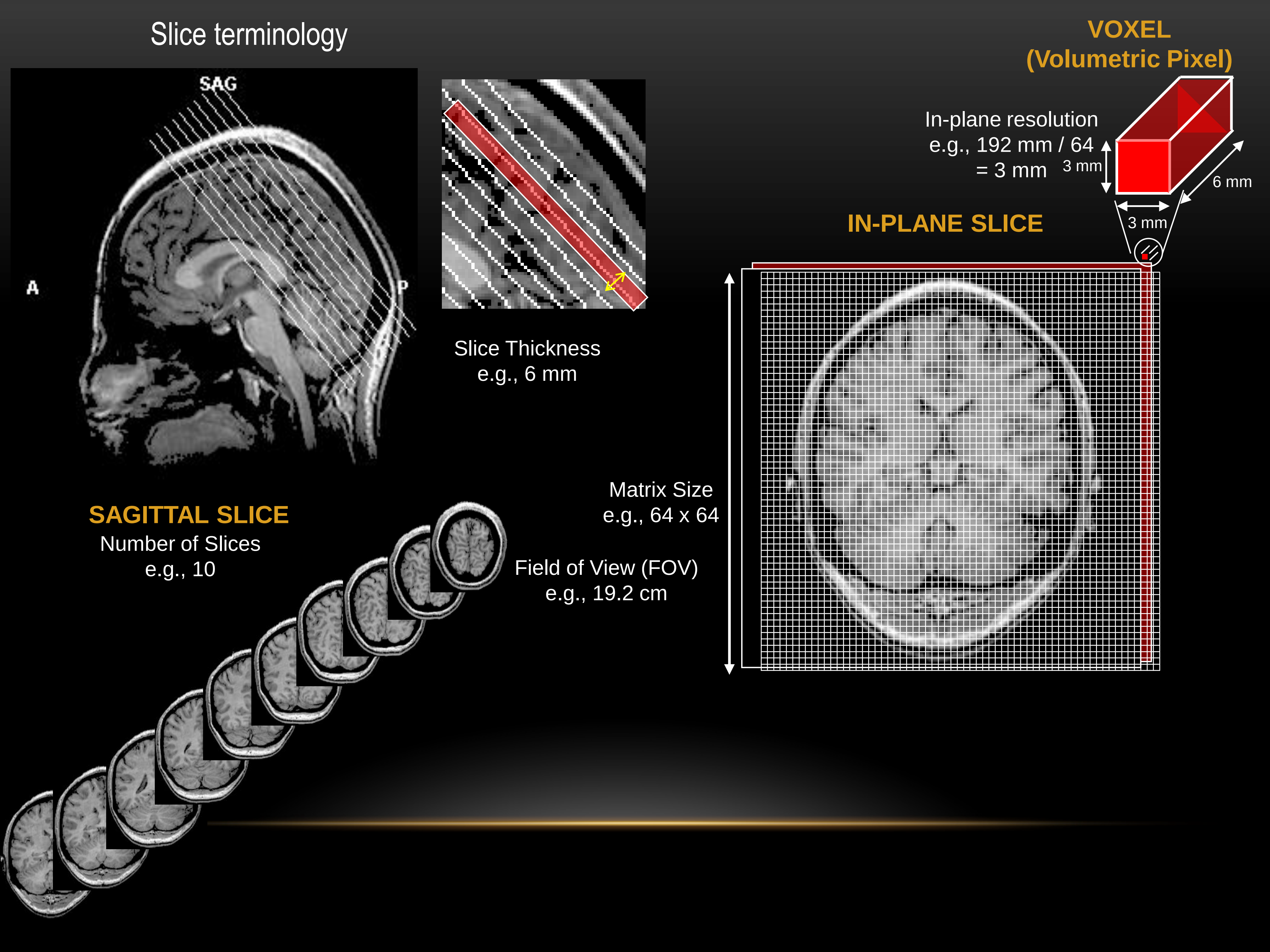
- Sources of noises in fMRI data
  - System noise
    - Thermal noise: motion of electrons in subject & RF eqp.
    - Signal drift: magnetic field drift over time
  - Subject dependent noise
    - Subject movement (<u>largest source of noise in fMRI data</u>)
  - Physiological noise
    - Cardiovascular, respiratory effect
  - Variability in BOLD response
    - Pulsatile motion: influx of blood into brain induced movement
  - Variability across sessions within the same subject
  - Variability across subjects

- Prevention is the Best Remedy
  - Tell your subjects how to be good subjects
    - “Don’t move” is too vague
    - Make sure the subject is comfy going in
    - avoid “princess and the pea” phenomenon

  - Emphasize importance of not moving at all during beeping
    - do not change posture
    - if possible, do not swallow
    - do not change mouth position
    - do not tense up at start of scan

  - Discourage any movements that would displace the head between scans

- Data pre-processing steps
  - to reduce noise maximize signal
    - Slice timing correction
    - Motion correction
    - Coregistration
    - Spatial normalization
    - Spatial smoothing
    - Temporal filtering

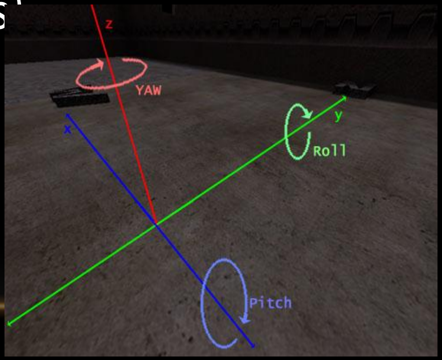


Slice terminology
VOXEL
(Volumetric Pixel)
In-plane resolution
e.g., 192 mm / 64
= 3 mm
3 mm
6 mm
3 mm
IN-PLANE SLICE
SAG
A
P
Slice Thickness
e.g., 6 mm
Matrix Size
e.g., 64 x 64
SAGITTAL SLICE
Number of Slices
e.g., 10
Field of View (FOV)
e.g., 19.2 cm

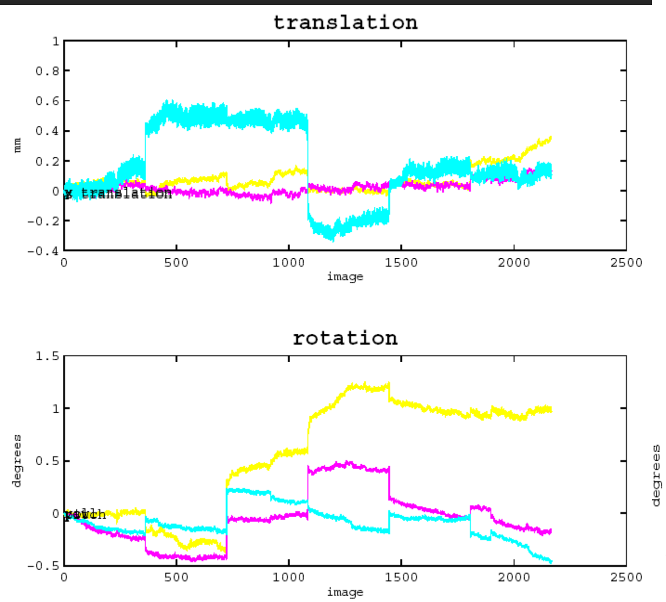
- Slice timing correction
  - Corrects for sampling of different slices at different time
    - Interpolation
- Motion correction
  - <u>Major source of variability</u>
  - Adjust for movement between slices
  - Typically take the first image and align the rest of the images to it
  - Rigid body transformation (6 parameters: 3 translation, 3 rotation)
  - Minimization of cost function
  - *(Use motion parameters as a covariate in statistical analysis)*

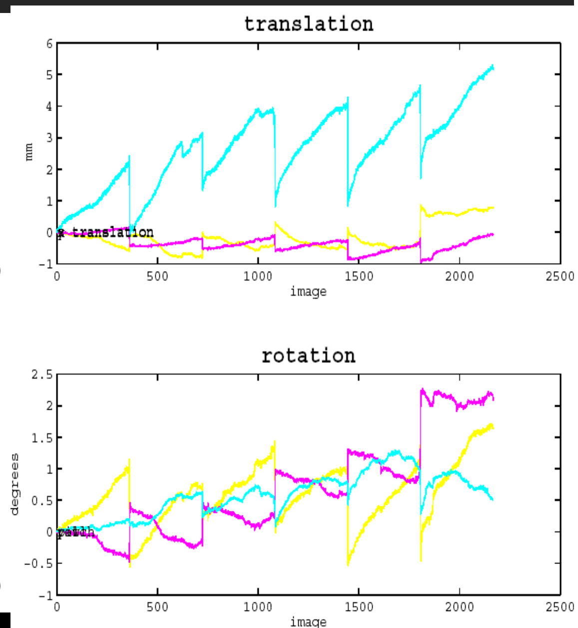
$$\begin{pmatrix} 1 & 0 & 0 & X\text{trans} \\ 0 & 1 & 0 & Y\text{trans} \\ 0 & 0 & 1 & Z\text{trans} \\ 0 & 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\Phi & \sin\Phi & 0 \\ 0 & -\sin\Phi & \cos\Phi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} \cos\Theta & 0 & \sin\Theta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\Theta & 0 & \cos\Theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} \cos\Omega & \sin\Omega & 0 & 0 \\ -\sin\Omega & \cos\Omega & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

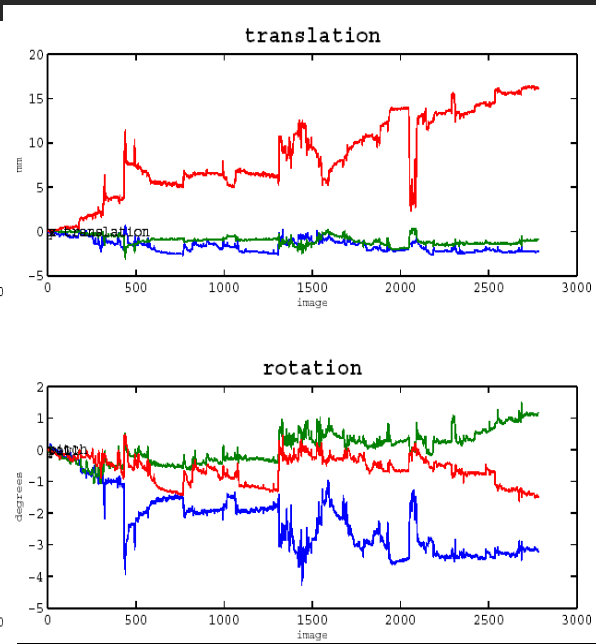
Translation *pitch (X-axis)* *Roll(Y-axis)* *Yaw (Z-axis)*



Head motion: Alignment with first image







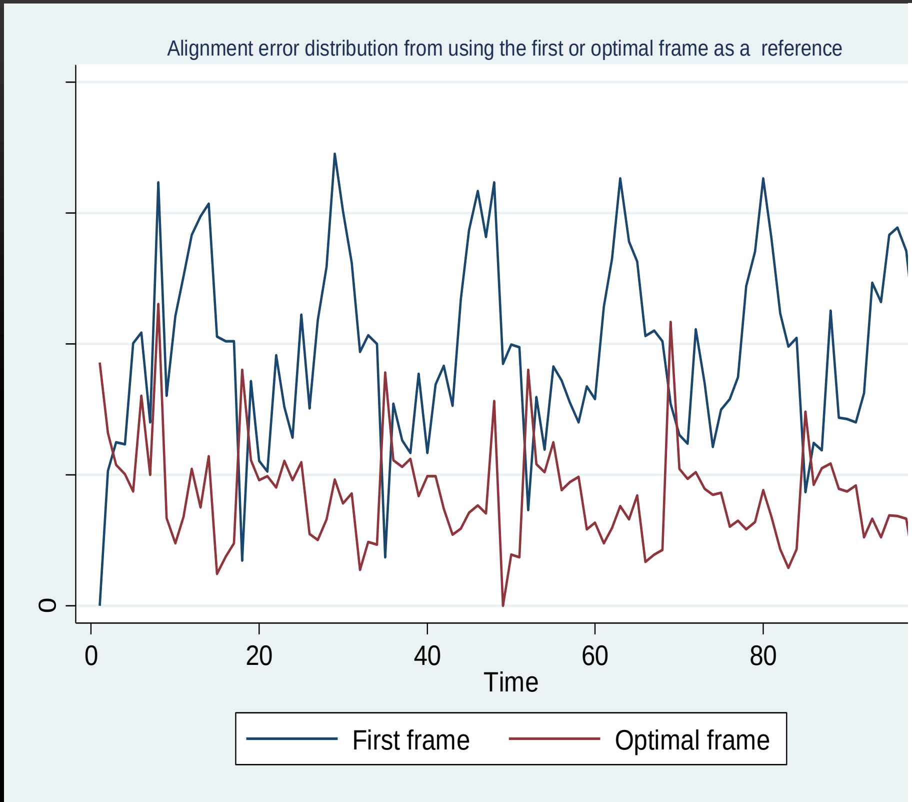
Good

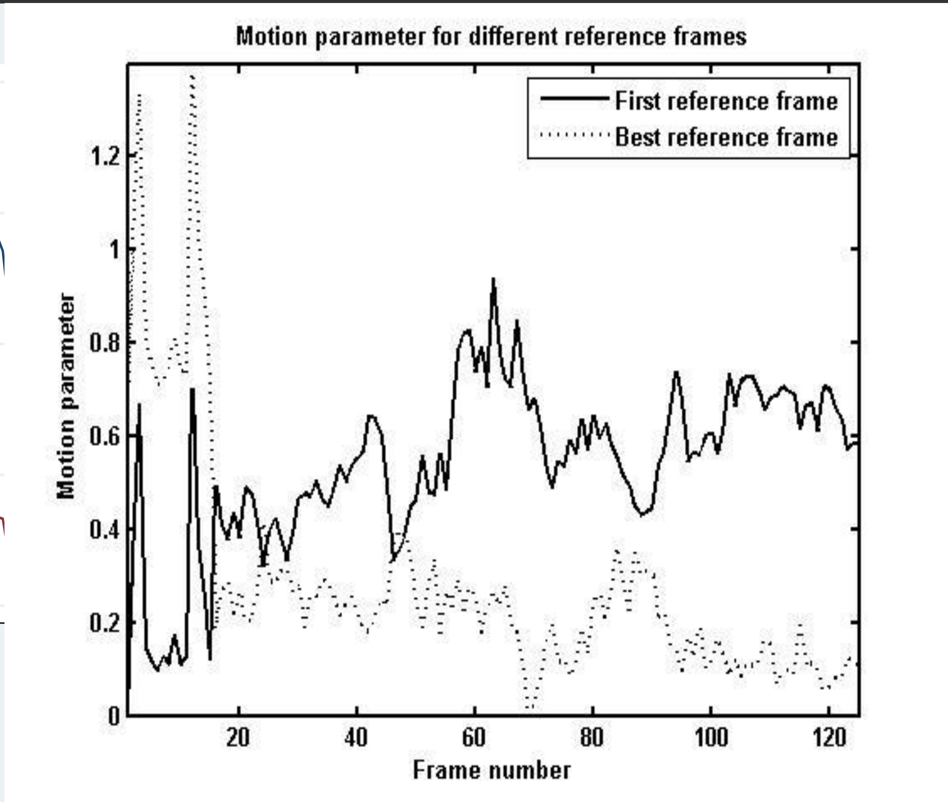
Bad

Ugly

CAN THIS BE OPTIMIZED ?

Slide from Duke course





Cost function $C(n) = \sqrt{\sum_i (S_i(n) - T_i)^2} / \sum_i T_i$

We propose to pick up the slide with minimum median error to be the reference

Also exclude frames that are misaligned with the rest of the frames

Altaye, Rajapola, Eaton, Holland (in progress)

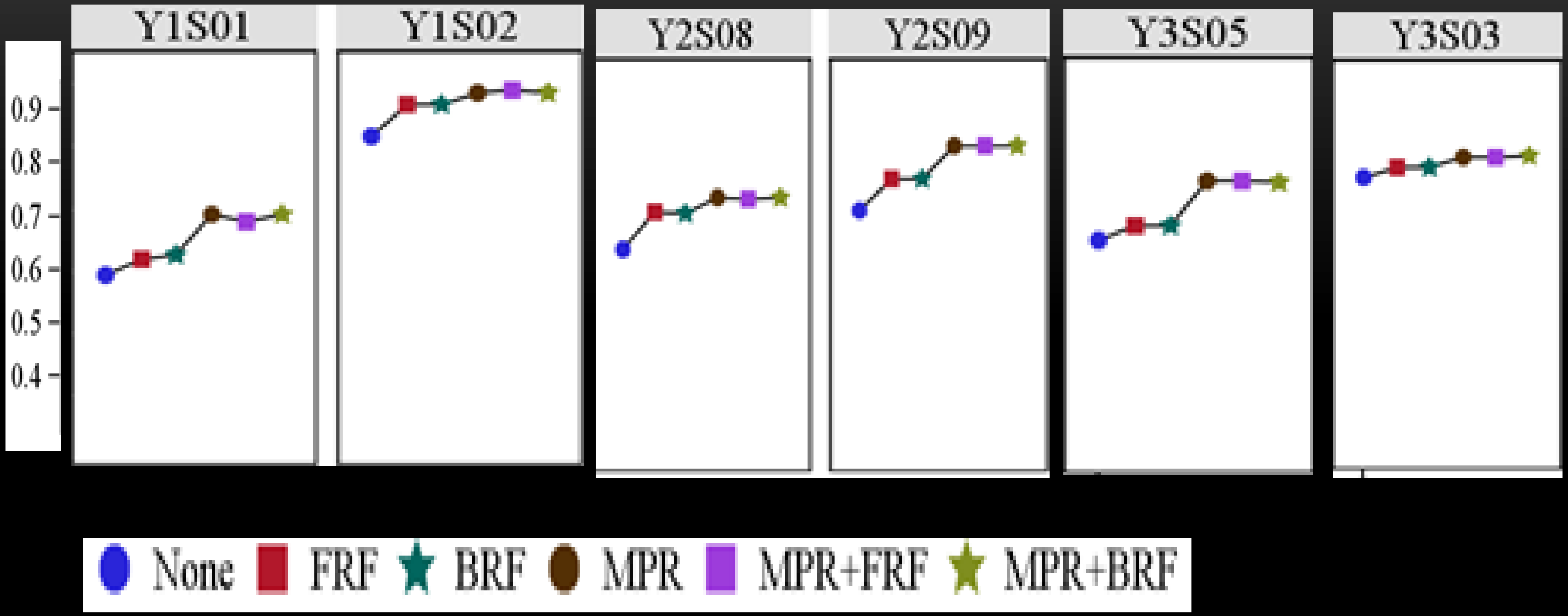
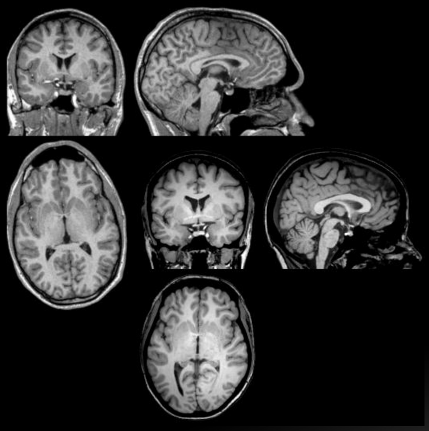


Figure 6. Individual Plots of the rSNR for six preprocessing pipelines for ten subjects

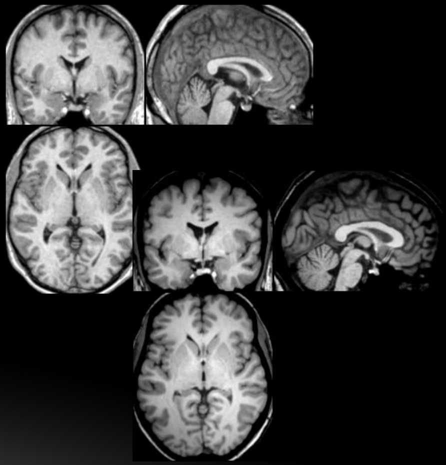
Meng, Altaye, Lin, Holland (in progress)

- Corgistration of functional and anatomical data
  - For displaying the result of functional maps
  - Used for later spatial normalization
  - Based on mutual information

- Spatial normalization
  - Warps images from different subjects into one template
  - Allows for group inference by combining data
  - Allows generalization of a study result

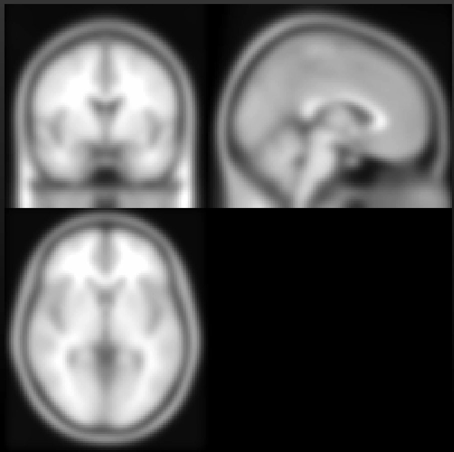
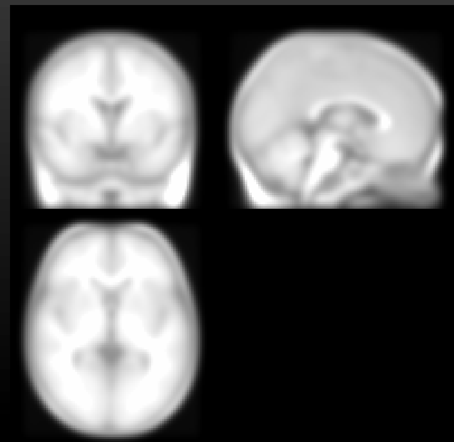
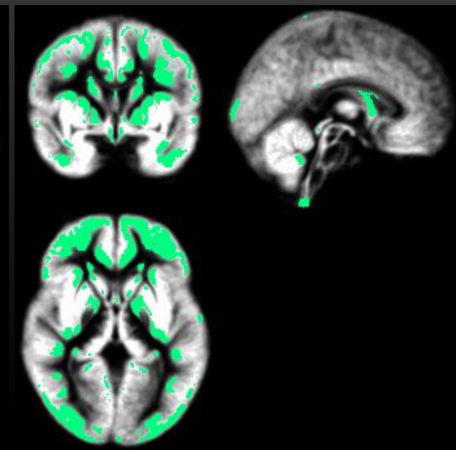
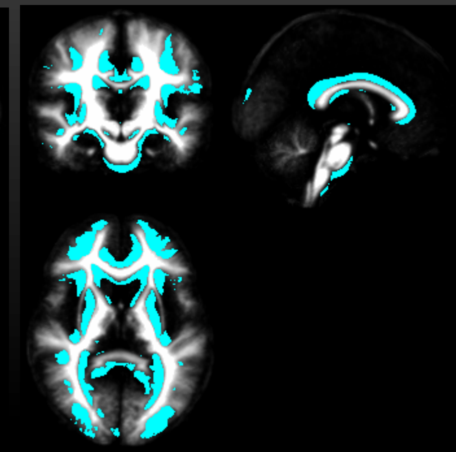


Raw



Normalized

- Question: WHICH TEMPLATE TO USE?
  - Talairach space: based on a single subject
  - MNI average of 153 brains age 18 to 64?
  - ICBM average of 452 brains of normal young adult
- Adult Standard – Talairach, ICBM, MNI is not appropriate for children (misclassification occur!)
- We develop pediatric and infant brain templates to be used for spatial normalization

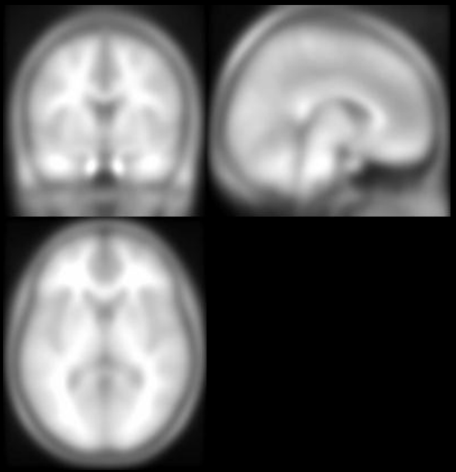
- Adult template

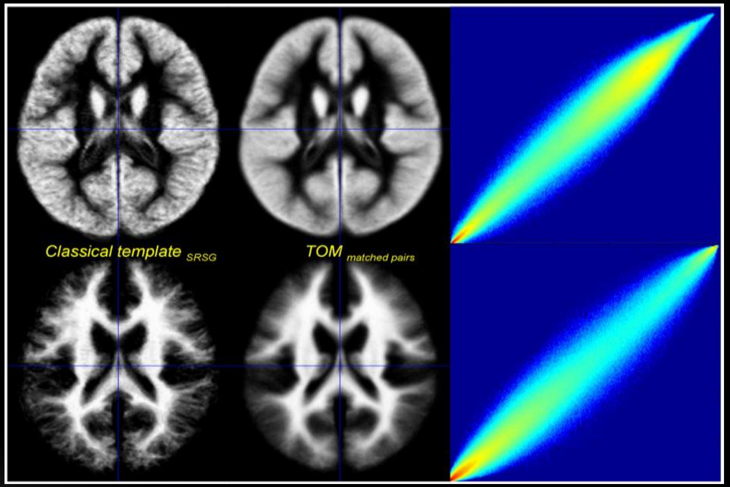
Infant template

Difference in GM (20%)

Difference in WM

Altaye M, Holland SK, Wilke M, Gaser C. Infant brain probability templates for MRI segmentation and normalization. NeuroImage (2008) 43(4): 721-30.

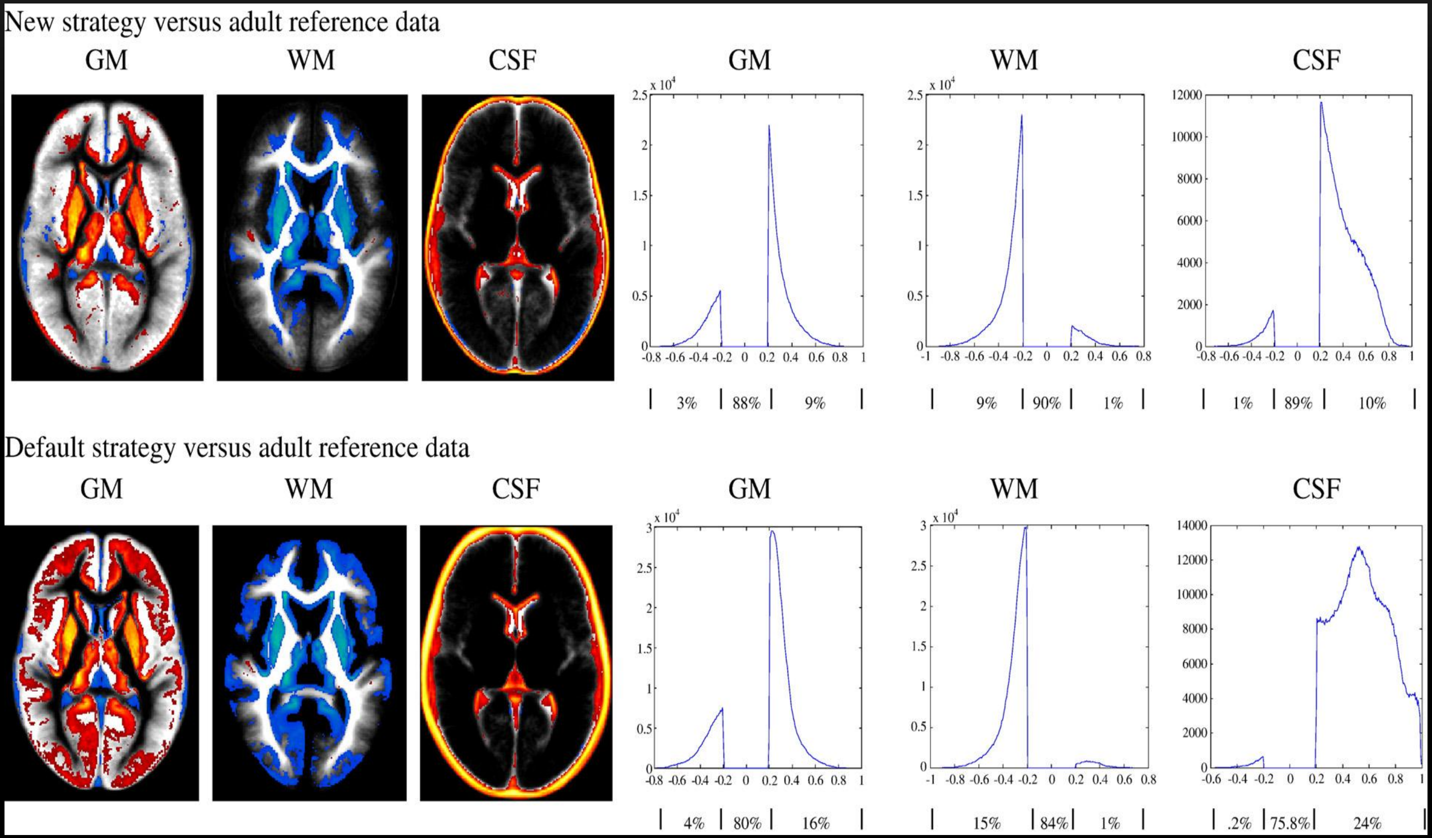




- Pediatric template

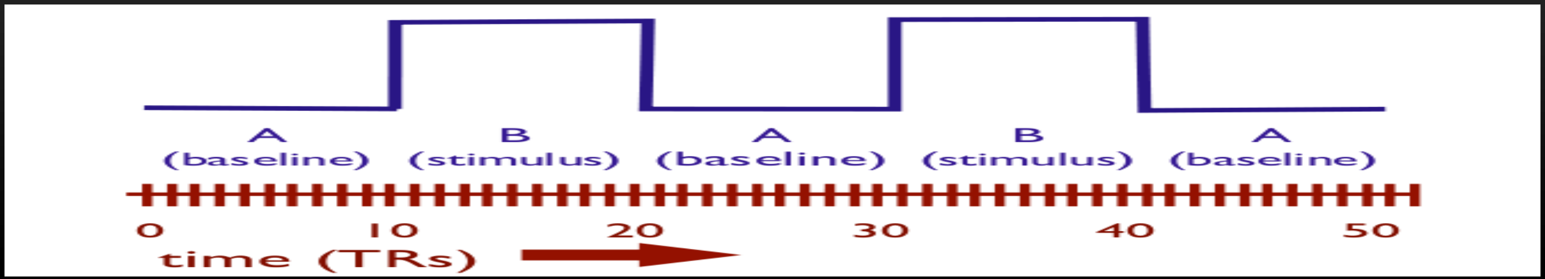
Age and gender appropriate templates for pediatrics

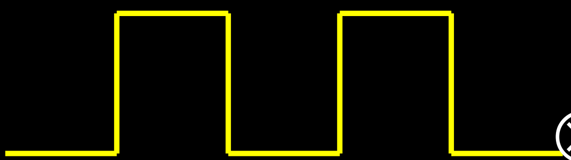
Wilke M, Holland SK, Altaye M, Gaser C. Template-o-matic: a toolbox for creating customized pediatric templates. NeuroImage (2008) 41(3): 903-13.


New strategy versus adult reference data
GM
WM
CSF
GM
WM
CSF
3%
88%
9%
9%
90%
1%
1%
89%
10%
Default strategy versus adult reference data
GM
WM
CSF
GM
WM
CSF
4%
80%
16%
15%
84%
1%
.2%
75.8%
24%

- Spatial smoothing
  - Use of Gaussian kernel
  - Increase SNR
  - Improve comparison across subjects
- Temporal filtering
  - Increase SNR
  - Reduce high frequency fluctuations and remove long term drift

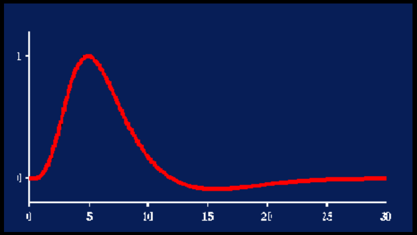
Design of typical fMRI studies: block design





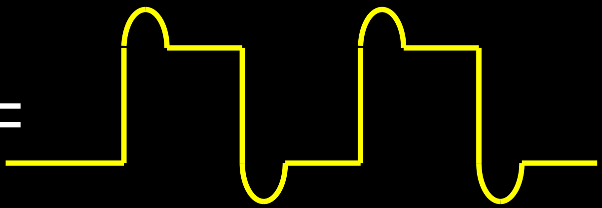
Boxcar function

$\otimes$

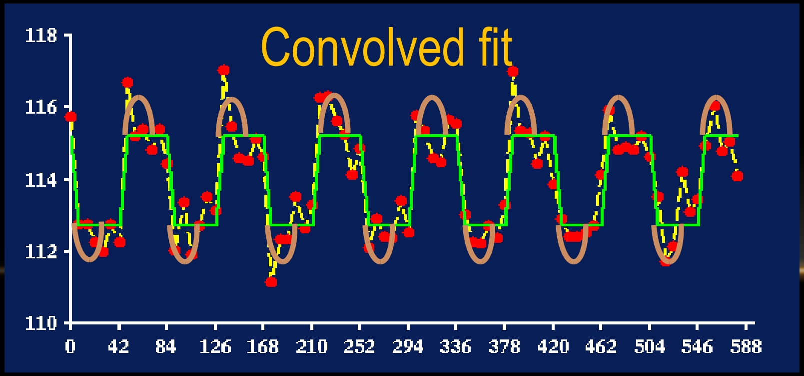


haemodynamic response

=



convolved with HRF



- Statistical methods for fMRI analysis
  - Activation analysis
    - Change between conditions, sessions, etc
    - Localized at a voxel or ROI level
      - GLM analysis
        - Two stage (summary statistics approach)
        - Full mixed model
  - Network analysis
    - Partitioning
      - Identify similarity in brain
        - PCA, ICA, Clustering
    - Functional connectivity
      - Correlations between remote neurophysiological events
        - Seed voxel approach, Spatial Bayesian Hierarchical model (BHM)
  - Prediction
    - Neural activity, group membership
        - BHM, SVM

- First level analysis
  - Fit a GLM for each subject at the voxel level
    - Address temporal AR correlations
      - Pre-coloring/temporal smoothing (Worsley, 1995)
      - Pre-whitning (Bullmore, 1996)
    - Then fit a GLM model where
      - $Y = X^{(1)}\theta^{(1)} + \varepsilon^{(1)}$
      - $Cov(\varepsilon^{(1)})=\sigma^2 V$
    - Estimation of $\theta^{(1)}$
      - Find W such that W V W'=I, then "whitened the model" by
        - $WY = WX^{(1)}\theta^{(1)}+W\varepsilon^{(1)}$
        - $Y^*=X^{(1)*}\theta^{(1)}+ \varepsilon^{(1)*}$
      - Use OLS on the whitened model
        - $\hat{\theta}^{(1)} = (X^{(1)*\prime}X^{(1)}*)^{-1}X^{(1)*\prime}Y$
        - $\widehat{Cov}(\hat{\theta}^{(1)})=\hat{\sigma}^2(X^{(1)*\prime}X^{(1)}*)^{-1}$
  - At this stage one can run a test (e.g. t-test) to see activated voxels for a given contrast for an individual

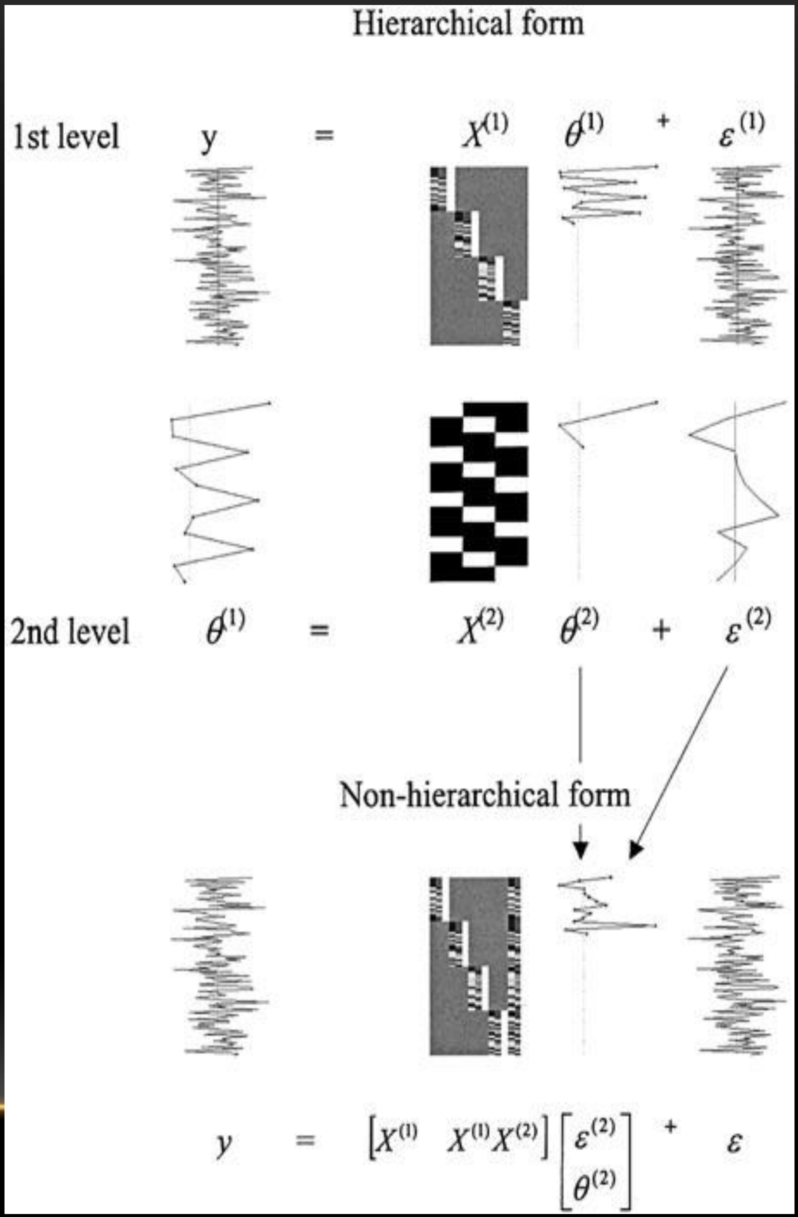
- Second level analysis at voxel level GLM
  - Fit a second level model that combines subject specific estimates
    - Second stage model
      - $\theta^{(1)} = X^{(2)}\theta^{(2)} + \varepsilon^{(2)}$
      - Cov($\varepsilon^{(2)}$)= $\begin{pmatrix} \sigma_1^2(X^{(1)*\prime}X^{(1)*})^{-1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_N^2(X^{(1)*\prime}X^{(1)*})^{-1} \end{pmatrix}$ $+\sigma^2{}_g\, I_N$
    - Estimation
        - $W\widehat{\theta^{(1)}} = WX^{(2)}\theta^{(2)}$+W$\varepsilon^{(2)}$
        - $\widehat{\theta^{(1)}}^{*}$=$X^{(2)*}\theta^{(2)}$+ $\varepsilon^{(2)*}$
      - Use OLS on the whitened model
        - $\hat{\theta}^{(2)} = (X^{(2)*\prime}X^{(2)}{}_{*})^{-1}X^{(2)*\prime}\widehat{\theta^{(1)}}^{*}$
        - $\widehat{Cov}(\hat{\theta}^{(2)})$=$(X^{(2)*\prime}X^{(2)^{*}})^{-1}$

Full mixed effects model

$$y = X^{(1)}\theta^{(1)} + \varepsilon^{(1)}$$
$$\theta^{(1)} = X^{(2)}\theta^{(2)} + \varepsilon^{(2)}$$

- $Cov(\varepsilon^{(1)}) = V$
- $Cov(\varepsilon^{(2)}) = V_g$

$$\begin{aligned} y &= X^{(1)}\left(X^{(2)}\theta^{(2)} + \varepsilon^{(2)}\right) + \varepsilon^{(1)} \\ &= \underbrace{X^{(1)}X^{(2)}\theta^{(2)}}_{\text{Fixed effect}} + \underbrace{X^{(1)}\varepsilon^{(2)} + \varepsilon^{(1)}}_{\text{Random effect}} \end{aligned}$$



- Estimation
  - Model is

$$y = X^{(1)}\left(X^{(2)}\theta^{(2)} + \varepsilon^{(2)}\right) + \varepsilon^{(1)}$$
$$= X^{(1)}X^{(2)}\theta^{(2)} + X^{(1)}\varepsilon^{(2)} + \varepsilon^{(1)}$$

$Cov(\varepsilon^{(1)}) = V$
Cov($\varepsilon^{(2)}$)=V_g

=$X^{(1)}X^{(2)}\theta^{(2)} + \omega$, where $\omega = X^{(1)}\varepsilon^{(2)} + \varepsilon^{(1)}$ & Cov$(\omega) = W$

Then

$$\widehat{\theta^{(2)}} = (X^{(2)\prime}X^{(1)\prime}W^{-1}X^{(1)}X^{(2)})^{-1}X^{(2)\prime}X^{(1)\prime}W^{-1}Y$$

$Cov(\widehat{\theta^{(2)}})$= $(X^{(2)\prime}X^{(1)\prime}W^{-1}X^{(1)}X^{(2)})^{-1}$

To get this the GLM need to be solved for the full vector Y and requires a large matrices and prohibitive computation burden and time

- Longitudinal (10yr) narrative comprehension in children and adolescent
  - Kids recurited 2000-2002 (n=28)
  - Hirerchical data (subject-year-time)

Let i=subject, j=voxel, k=year, t=time

1st level: $Y_{ijkt} = \beta_{0ijk} + \beta_{1ijk}X + \xi_{ijkt}$

2nd level: $\beta_{0ijk} = \lambda_{00ij} + \lambda_{01ij}(\text{year}) + \eta_{0ijk}$

$\beta_{1ijk} = \lambda_{10ij} + \lambda_{11ij}(\text{year}) + \eta_{1ijk}$

3rd level: $\lambda_{00ij} = \theta_{000j} + \theta_{001j}(IQ) + \varepsilon_{00ij}$

$\lambda_{01ij} = \theta_{010j} + \varepsilon_{01ij}$

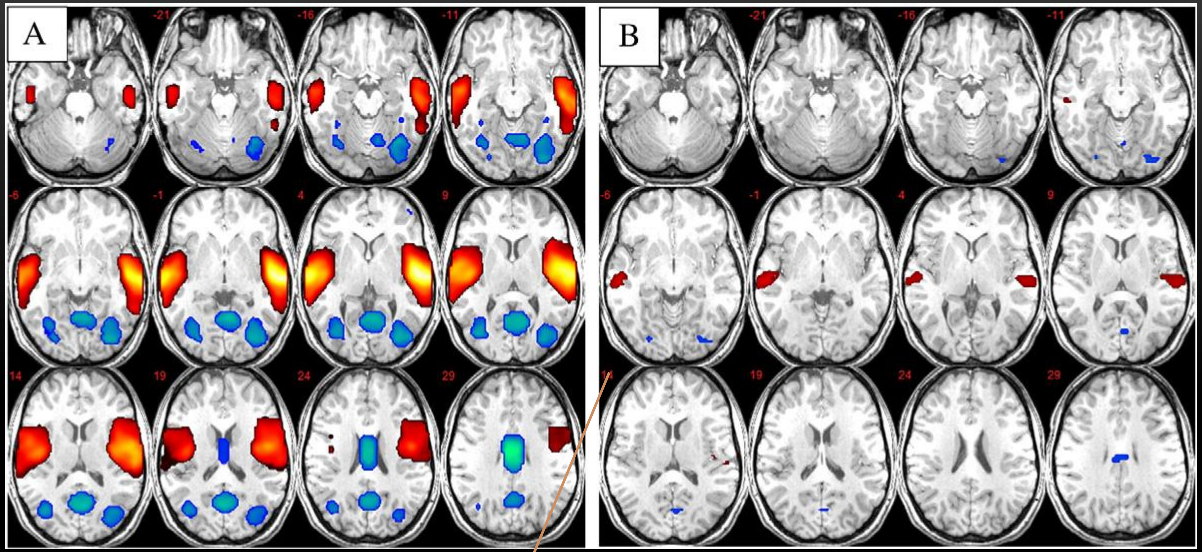
$\lambda_{10ij} = \theta_{100j} + \theta_{101j}(IQ) + \varepsilon_{10ij}$

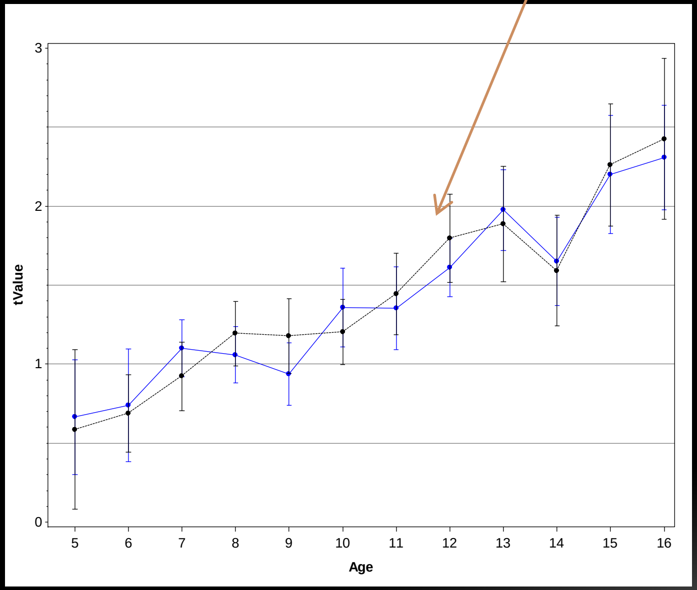
$\lambda_{11ij} = \theta_{110j} + \varepsilon_{11ij}$

- Putting them together

$$Y_{ijkt} = [\ \theta_{000j} + \theta_{001j}(IQ) + \theta_{010j}(\text{year}) + \theta_{100j}X + \theta_{101j}(IQ)(X) + \theta_{110j}(\text{year})(X) +$$
$$\varepsilon_{00ij} + \varepsilon_{10ij}(X) + \varepsilon_{01ij}(\text{year}) + \varepsilon_{11ij}(\text{year})(X) + \eta_{1ijk}X + \quad + \xi_{ijkt}$$

- Took 10 minutes per voxel, ~ 10,000 voxels ~ 1667 hrs
- Used two-stage modeling the time series first



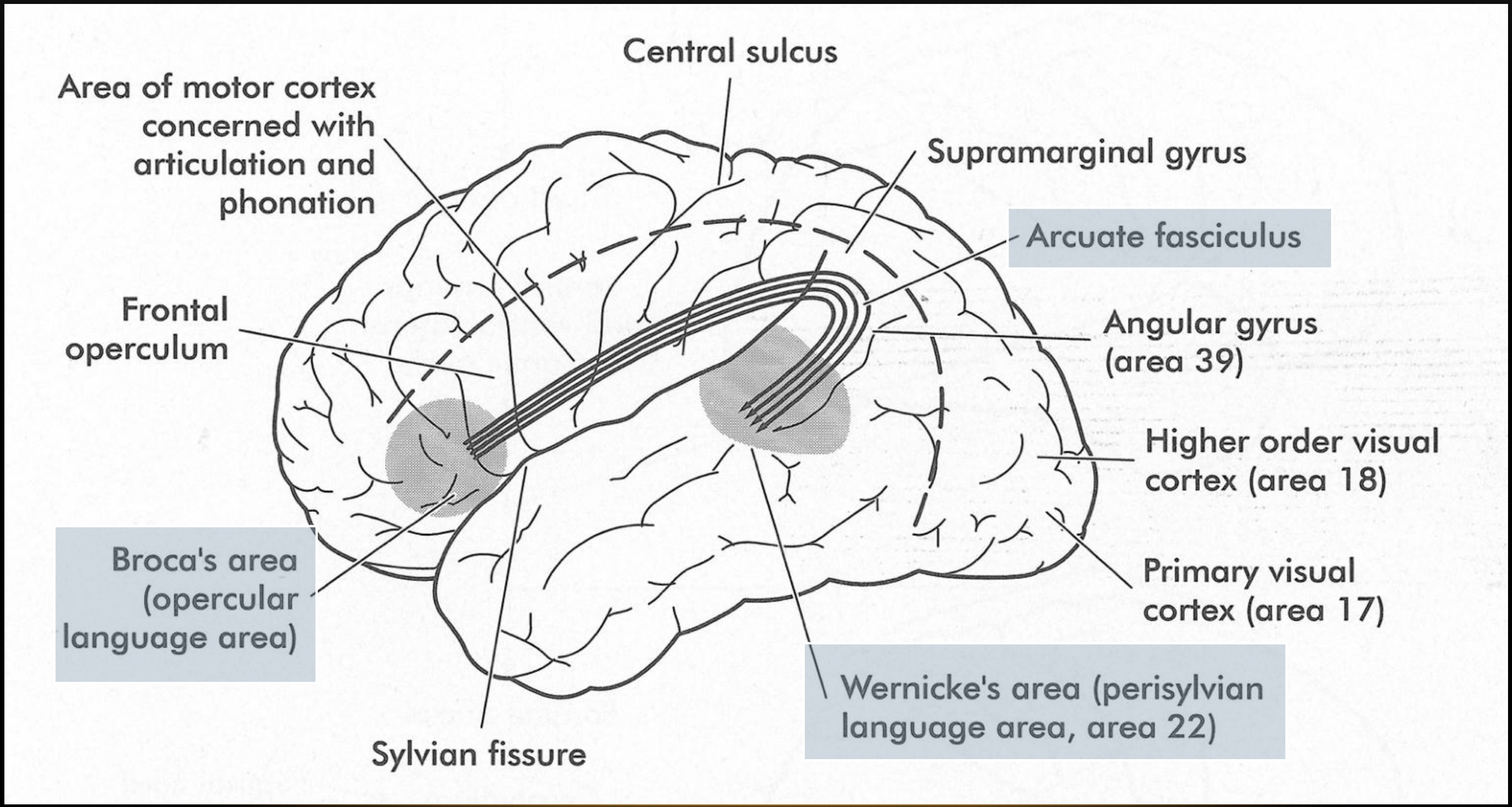


Szaflarski JP, Altaye M, Rajagopal A, Eaton K, Meng X, Plante E, Holland SK. A 10-year longitudinal fMRI study olf narrative comprehension in children and adolescents. NeuroImange (2012.

- Multiple comparison issues
  - ~100K voxels → 5K false positives!!
- Correction methods
  - FWE
    - Bonferroni
    - RFT
    - Permutation/randomization test
    - Wavelet methods
  - FDR
    - Different variant

- What about the spatial correlation?
  - RFT smoothing
  - Extent thresholding
  - Spatio-temporal model at second level for a defined anatomical areas (Derado et al 2010)

WERNICKE-GESHWIND MODEL OF HUMAN LANGUAGE



(Adapted from MA England, J Wakley: Color Atlas of Brain and Spinal Cord: An Introduction to Normal Neuroanatomy, St. Louis, 1991, Mosby)