

Martin Lindquist

Department of Biostatistics
Johns Hopkins
Bloomberg School of Public Health

Tor Wager

Department of Psychology and
Neuroscience and the
Institute for Cognitive Science
University of Colorado, Boulder

Model Building III: Filtering and nuisance covariates

GLM

A standard GLM can be written:

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

where

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & \cdots & X_{1p} \\ 1 & X_{21} & \cdots & X_{2p} \\ \vdots & \vdots & & \vdots \\ 1 & X_{n1} & \cdots & X_{np} \end{bmatrix} \times \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

↑ ↑ ↑ ↑

fMRI Data Design matrix Regression coefficients Noise

\mathbf{V} is the covariance matrix whose format depends on the noise model.

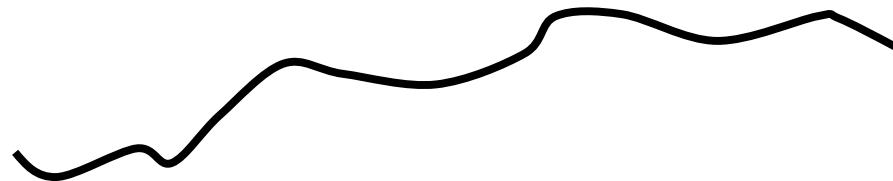
Nuisance Covariates

- Often model factors associated with known sources of variability, but that are not related to the experimental hypothesis, need to be included in the GLM.

- Examples of possible ‘nuisance regressors’:
 - Signal drift
 - Physiological (e.g., respiration) artifacts
 - Head motion, e.g. six regressors comprising of three translations and three rotations.
 - Sometimes transformations of the six regressors also included.

Drift

- Slow changes in voxel intensity over time (low-frequency noise) is present in the fMRI signal.
- Scanner instabilities and not motion or physiological noise may be the main cause of the drift, as drift has been seen in cadavers.
- Need to include drift parameters in our models.
 - Use splines, polynomial basis or discrete cosine basis



Model with Drift

$$Y = X \beta + \varepsilon$$

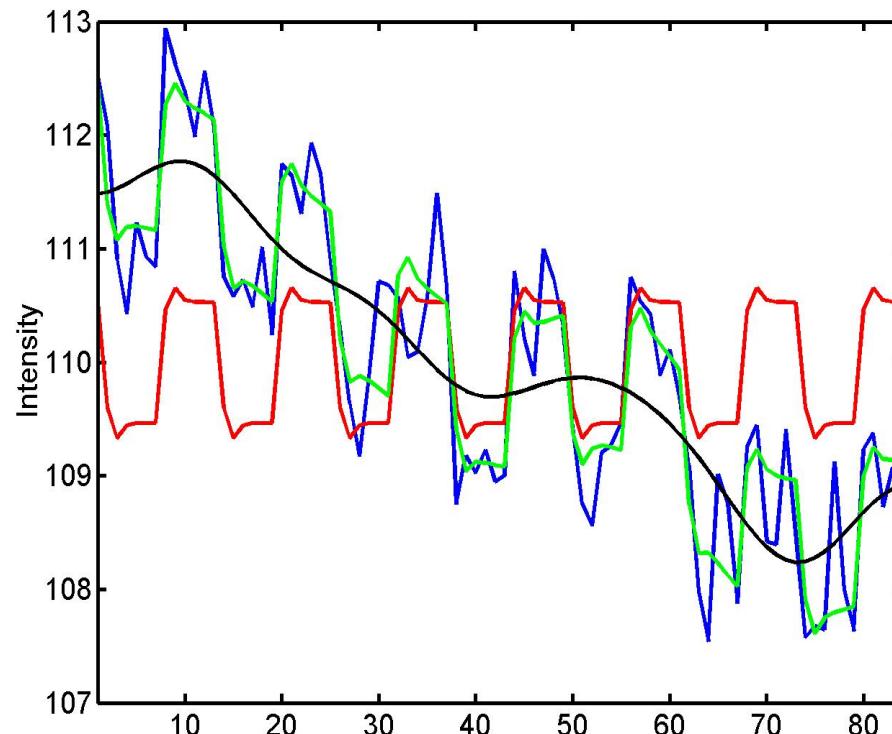
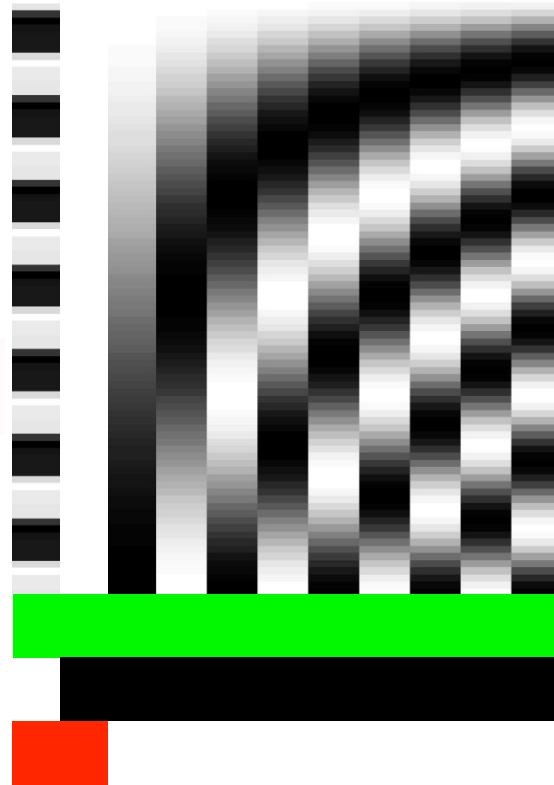
Diagram illustrating the linear regression model with drift:

- Y**: Response variable, represented by a vertical vector of observations.
- X**: Predictor variable, represented by a vertical matrix of observations. It includes a column of ones (intercept) followed by columns representing different predictors.
- β** : Coefficient vector, represented by a vertical vector of coefficients ($\beta_1, \beta_2, \dots, \beta_{11}$).
- ε** : Error term, represented by a vertical vector of error terms.

The equation $Y = X \beta + \varepsilon$ indicates that the observed response Y is equal to the product of the predictor matrix X and the coefficient vector β , plus the error term ε .

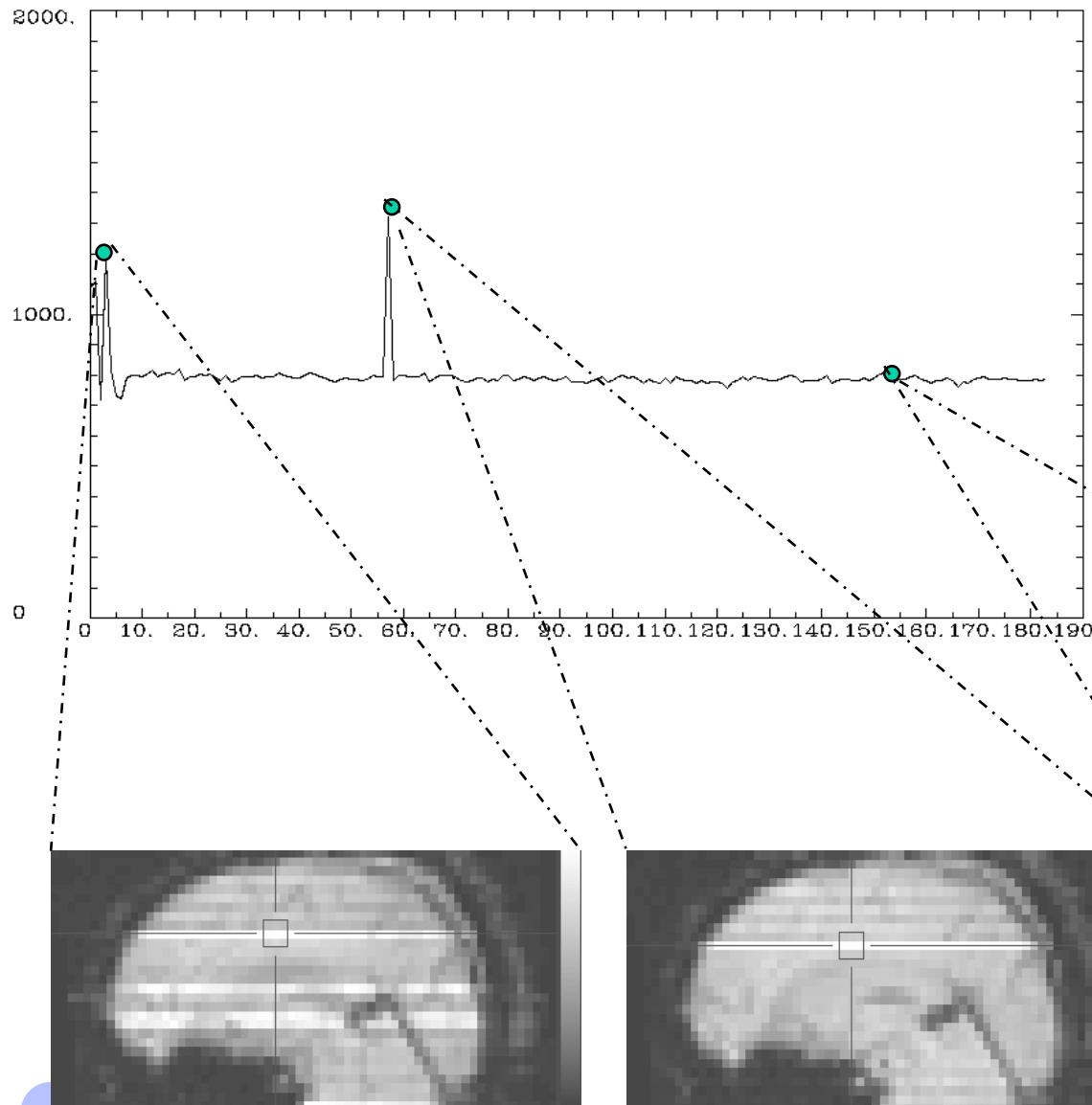


High Pass Filtering



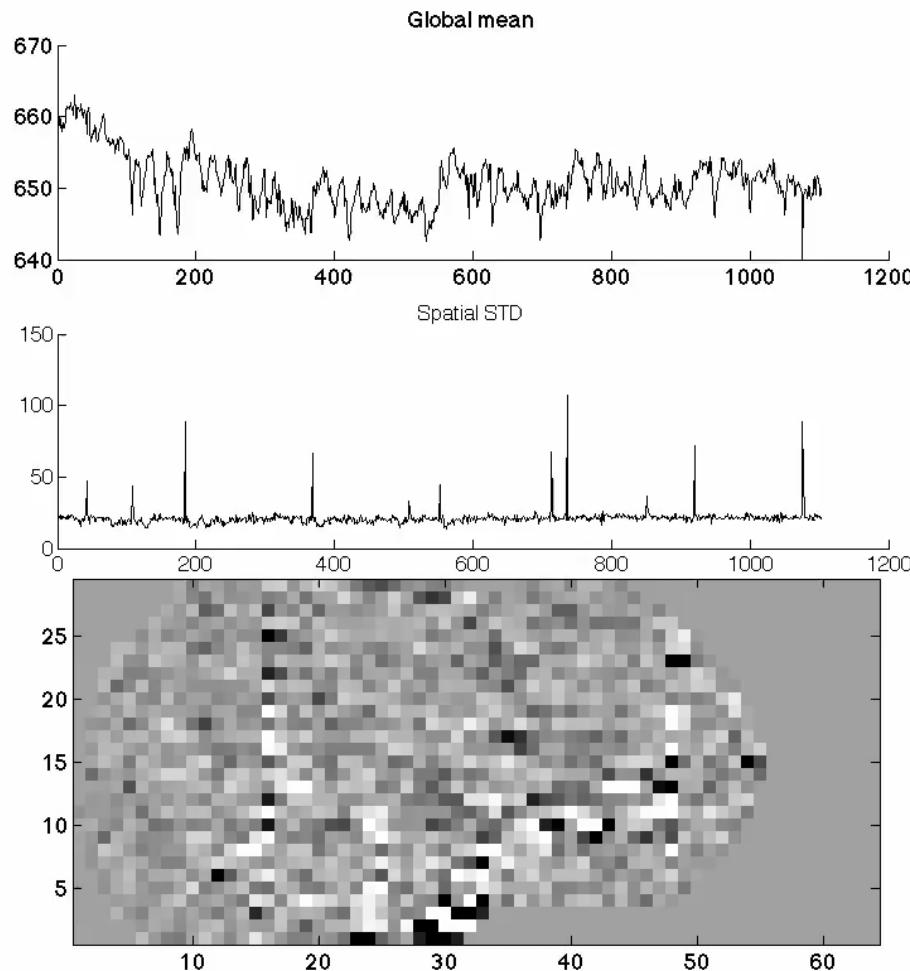
blue = data
black = mean + low-frequency drift
green = predicted response, taking into account low-frequency drift
red = predicted response (with low-frequency drift explained away)

Transient gradient artifacts



Transient spike artifacts in the data during isolated volume acquisitions are apparent in certain slices, as shown by the bright bands in the sagittal slices (bottom). This suggests that gradient performance was affected during acquisition of some echo-planar images, which were acquired slice-by-slice in interleaved order in this experiment.

Modeling transient gradient artifacts



Outlier detection

- Global mean
- Successive differences (RMSSD)
- Include as covariates in design matrix
- One regressor per “bad” image

Example nuisance
regressors in X

First images: T1
equilibrium issues

Spike

Spike



Physiological Noise

- Respiration and heart beat give rise to periodic noise, often aliased into task frequencies.
- It can potentially be modeled, but if the TR is too low there will be problems with aliasing.
 - Sampling rate must be at least twice as big as the frequency of the curve you seek to model.
- Hence, this type of noise is difficult to remove and is often left in the data giving rise to temporal autocorrelations.



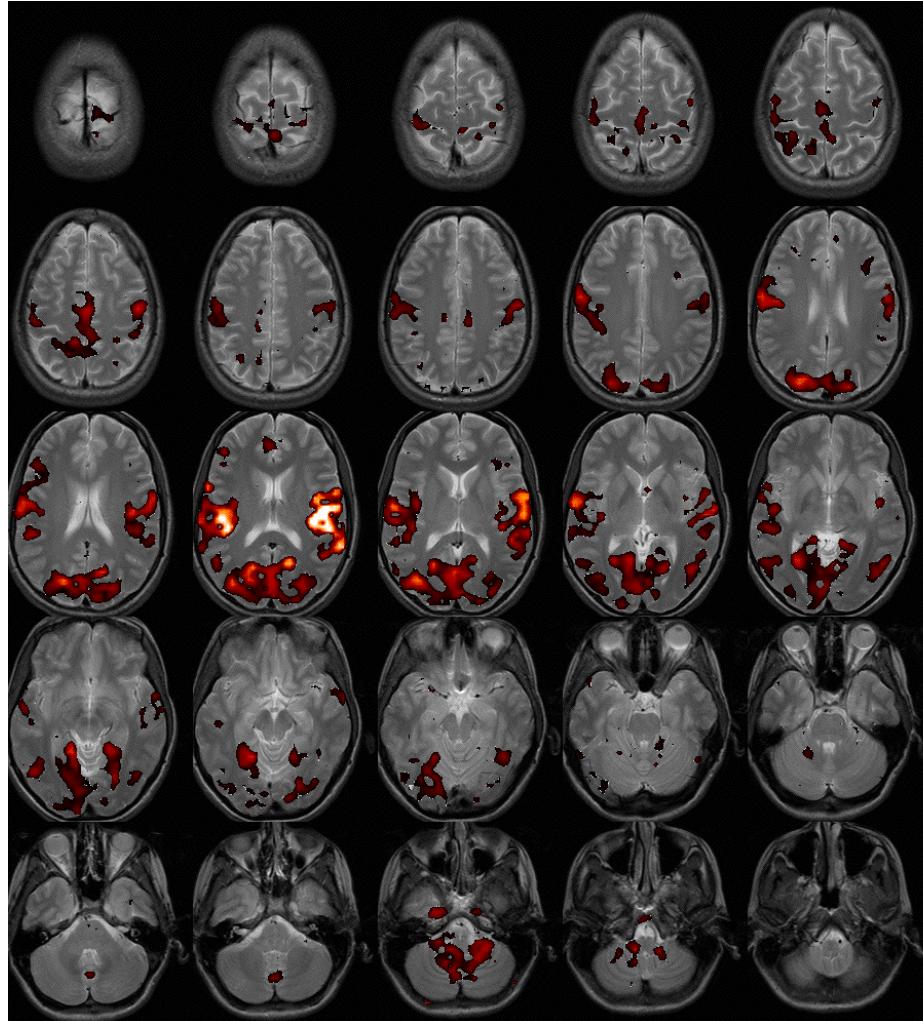
Physio noise reduction

- Neuronal activation
 - Respiration cycle
 - Cardiac cycle
 - Respiration volume ($\approx \text{CO}_2$)
 - Heart rate
- }
- RETROICOR
- }
- RVHRCOR

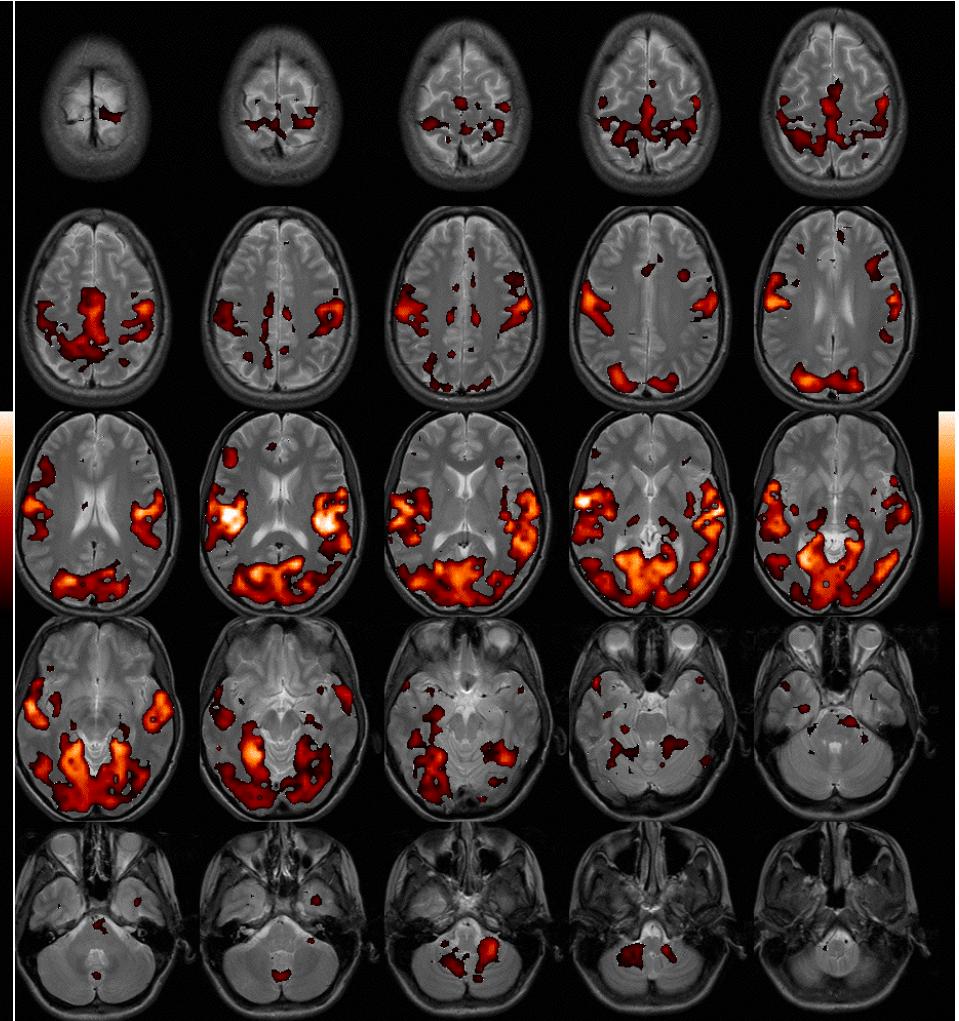


RETROICOR

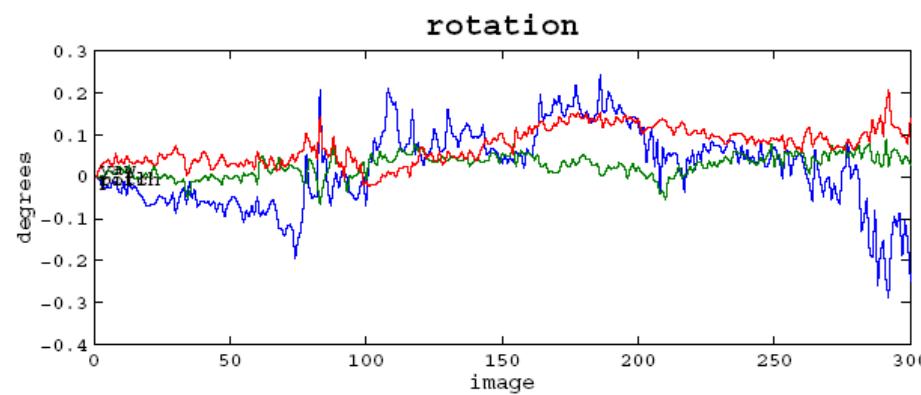
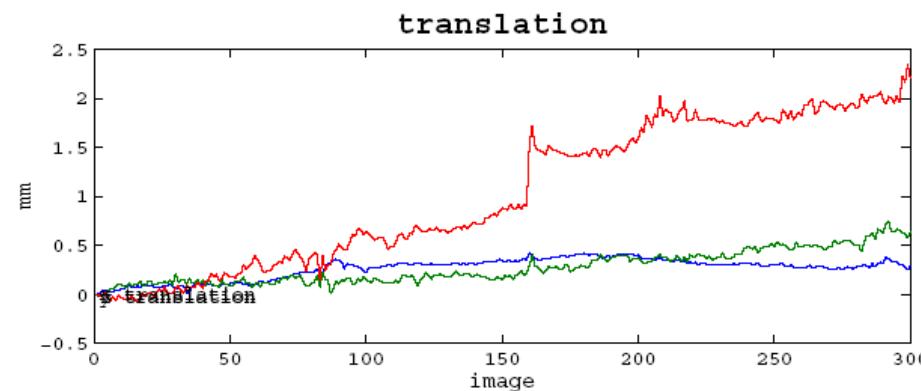
No retroicor



With retroicor



Head Motion



Motion

- Subject motion during the experiment can also give rise to serious problems.
- Basic motion correction (image realignment) is performed in the pre-processing stages of the analysis.
- However, motion induces complex changes in the magnetic field and ‘spin-history’ artifacts that cannot be removed.



Head movement-related artifacts

The influence of head motion on intrinsic functional connectivity MRI

Koene R.A. Van Dijk ^{a,b}, Mert R. Sabuncu ^{b,c}, Randy L. Buckner ^{a,b,d,e,*}

^a Harvard University Department of Psychology, Center for Brain Science, Cambridge, MA, USA

^b Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital, Charlestown, MA, USA

^c Computer Science and Artificial Intelligence Lab, Massachusetts Institute of Technology, Cambridge, MA, USA

^d Department of Psychiatry, Massachusetts General Hospital, Boston, MA, USA

^e Howard Hughes Medical Institute, Cambridge, MA, USA

Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion

Jonathan D. Power ^{a,*}, Kelly A. Barnes ^a, Abraham Z. Snyder ^{a,b},
Bradley L. Schlaggar ^{a,b,c,d}, Steven E. Petersen ^{a,b,d,e}

^a Department of Neurology, Washington University School of Medicine, St. Louis, MO, USA

^b Department of Radiology, Washington University School of Medicine, St. Louis, MO, USA

^c Department of Pediatrics, Washington University School of Medicine, St. Louis, MO, USA

^d Department of Anatomy & Neurobiology, Washington University School of Medicine, St. Louis, MO, USA

^e Department of Psychology, Washington University in Saint Louis, St. Louis, MO, USA

Head movement-related artifacts: Solutions



- **Nuisance regressors:** Movement- and global CSF-related covariates
- **'Scrubbing':** Drop images with high movement estimates (treat as missing data)

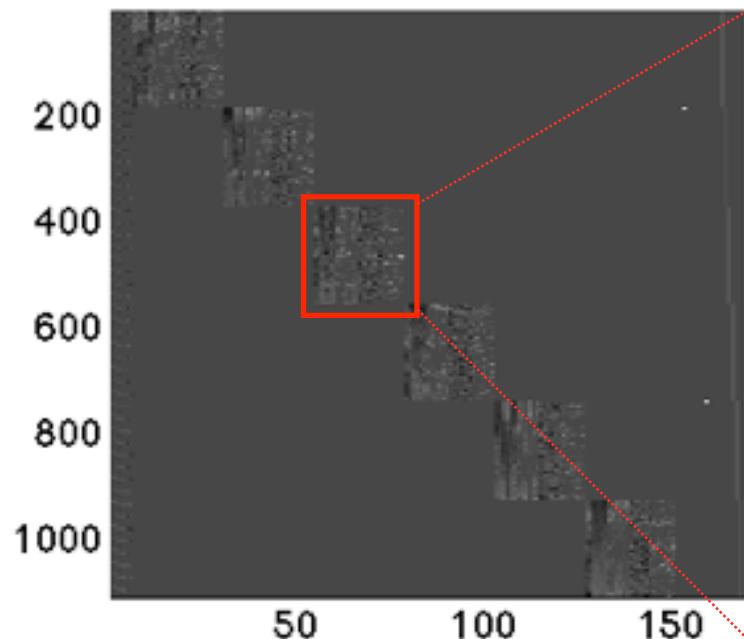


Modeling movement

Nuisance regressors, one run

Mvmt (centered), squares, successive differences, squared diffs [24 covs]

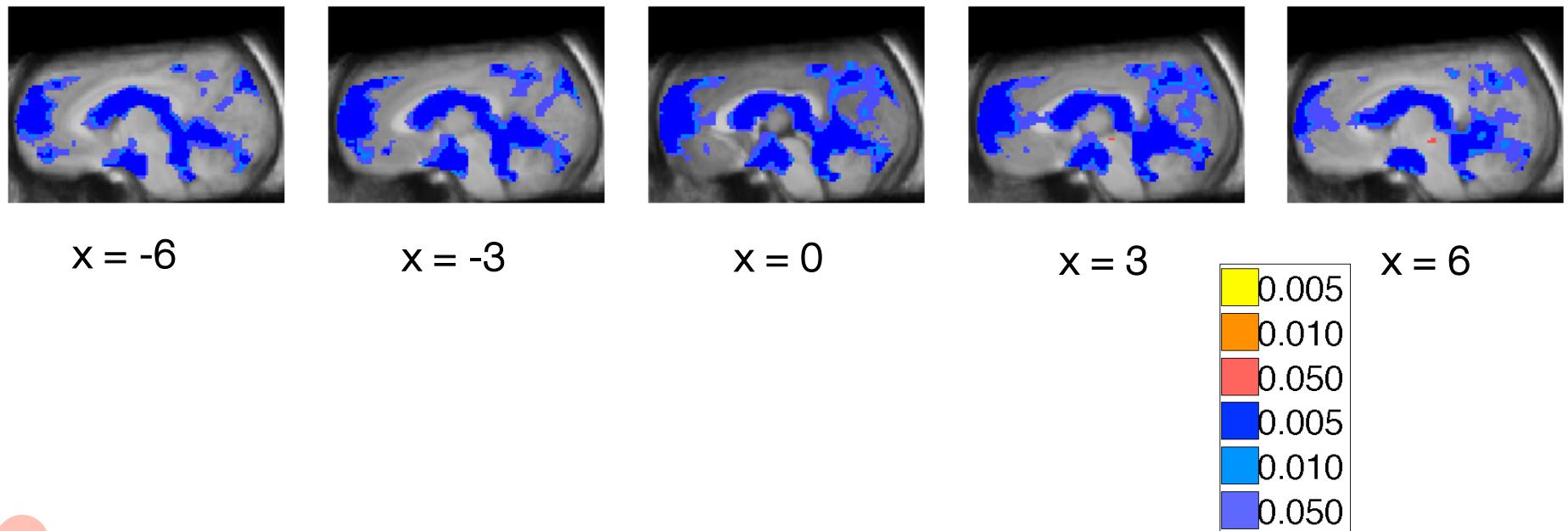
Full Design Matrix (zscored)



Lund et al. 2006¹⁷

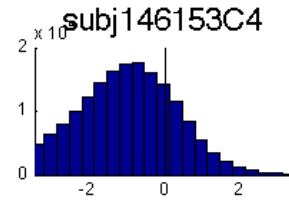
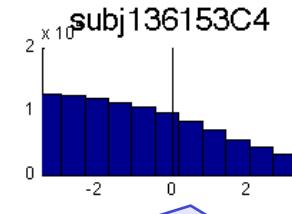
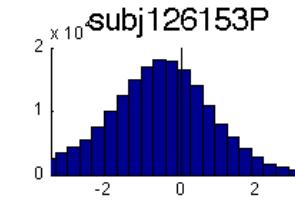
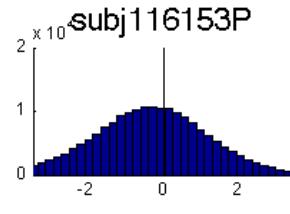
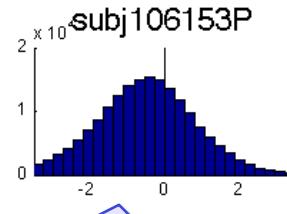
Trouble in paradise: Task-correlated movement and spike artifacts

- CS+ vs. CS- during Shock Conditioning
- Group analysis, N = 25. Problems should “average out”...right?

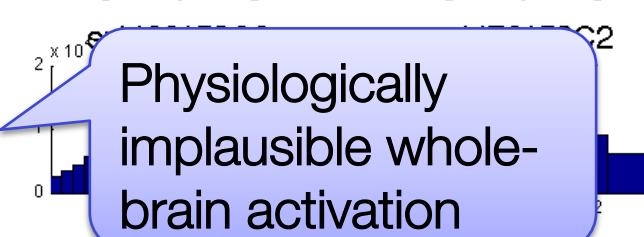
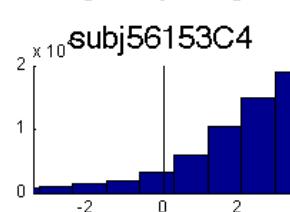
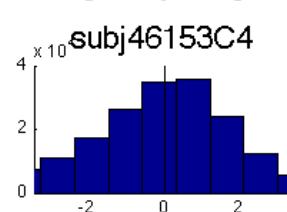
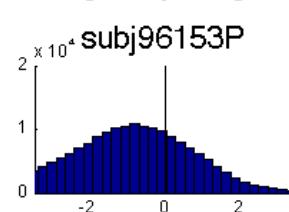
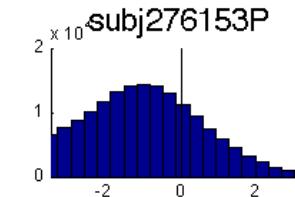
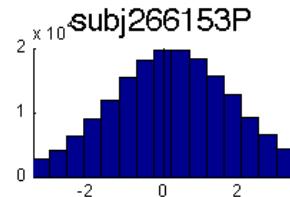
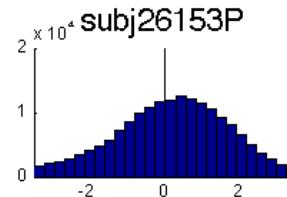
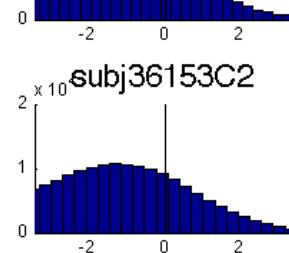
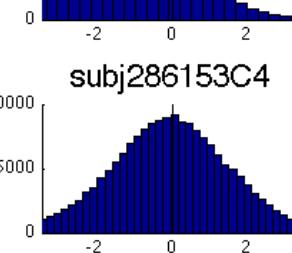
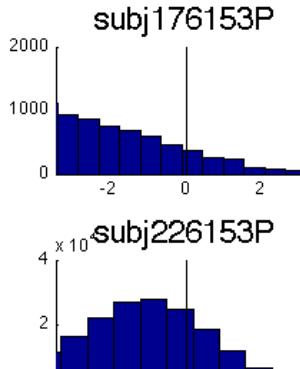
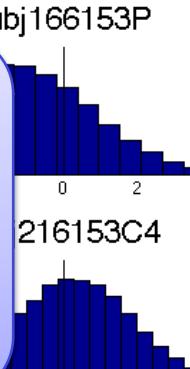


Detecting problems

Histograms of contrast values across the brain for each subject



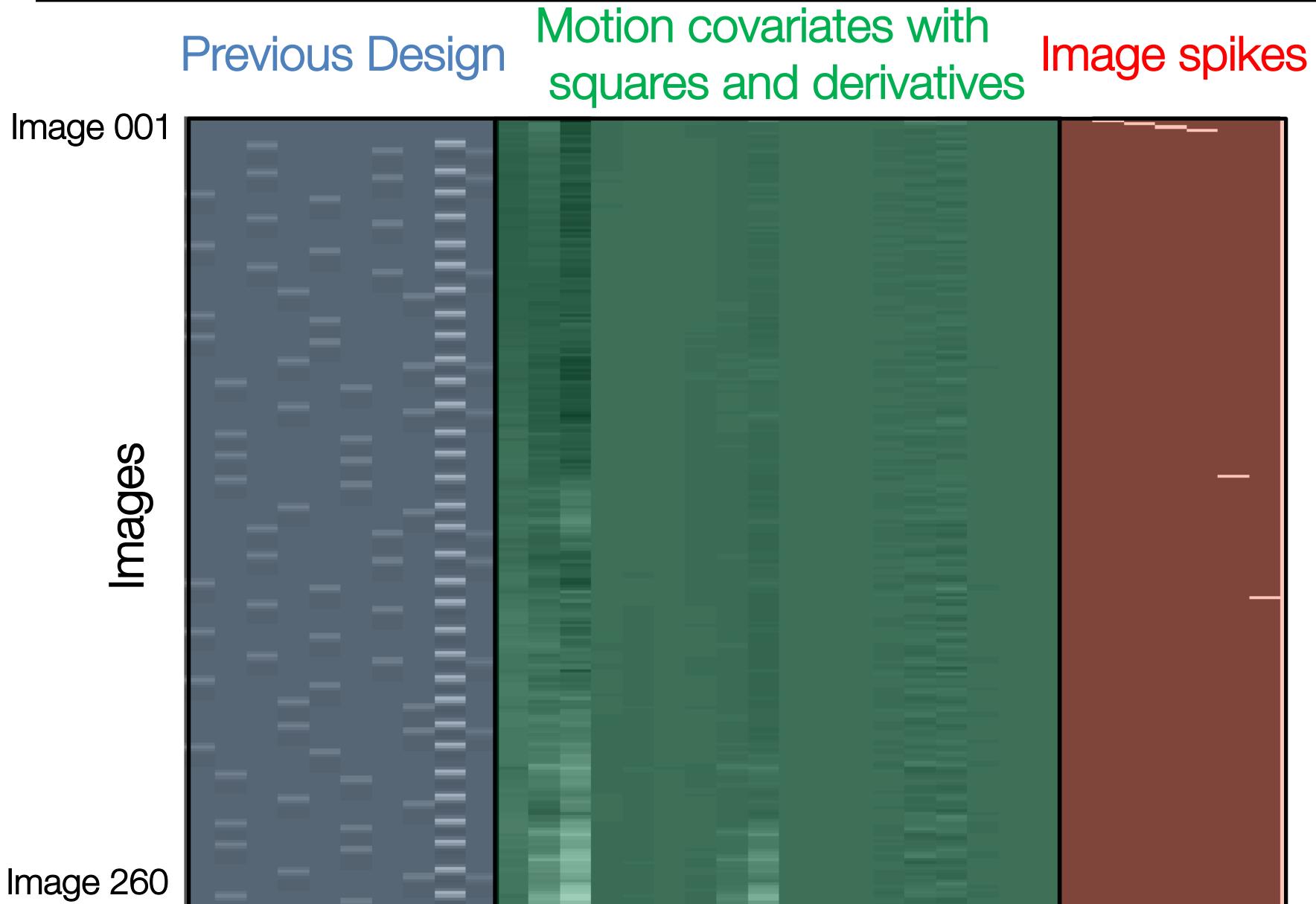
Distribution across entire brain.
Should be roughly mean zero and on same scale for all.



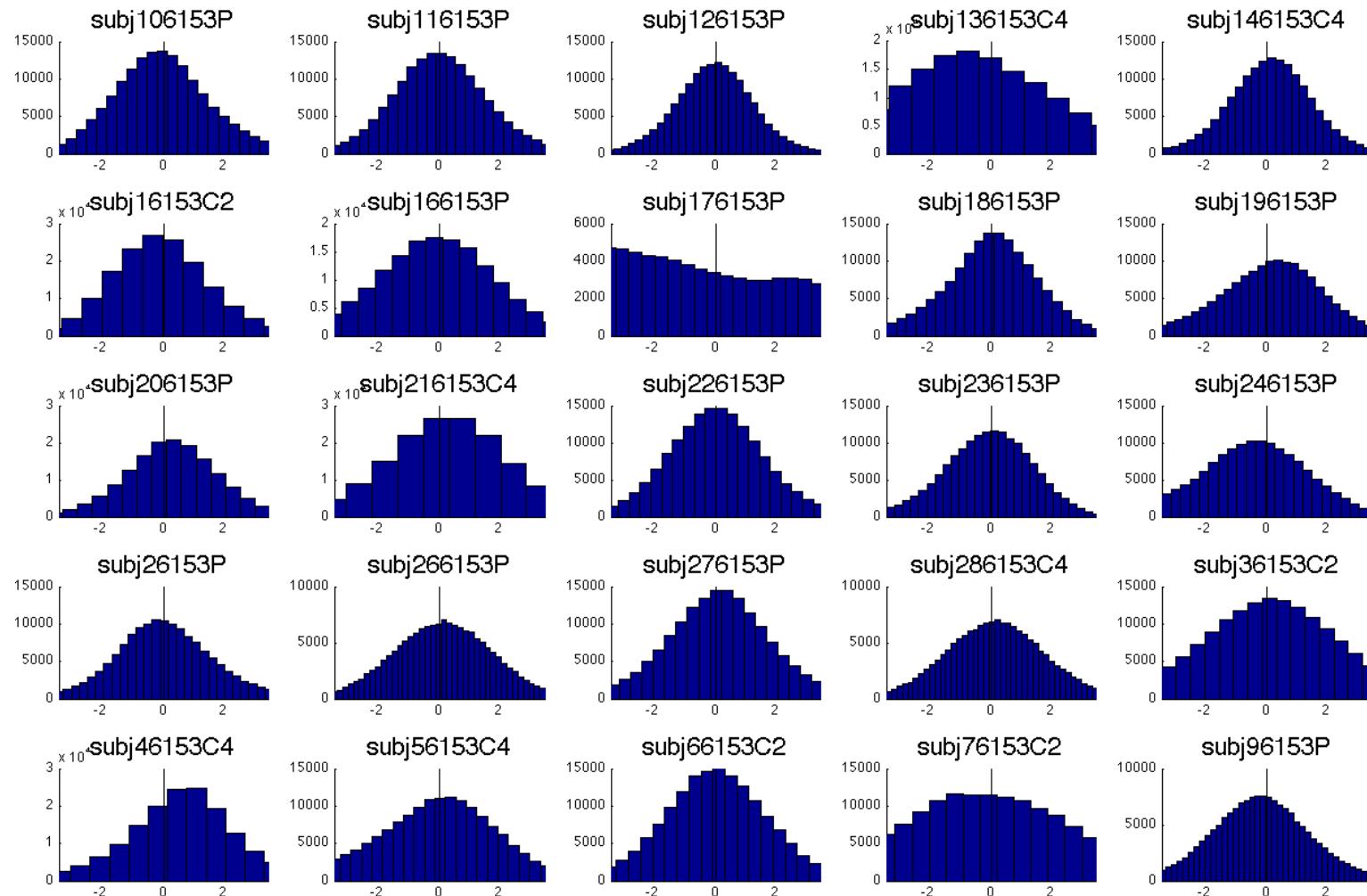
Physiologically implausible whole-brain deactivation

Physiologically implausible whole-brain activation

Additional nuisance covariates



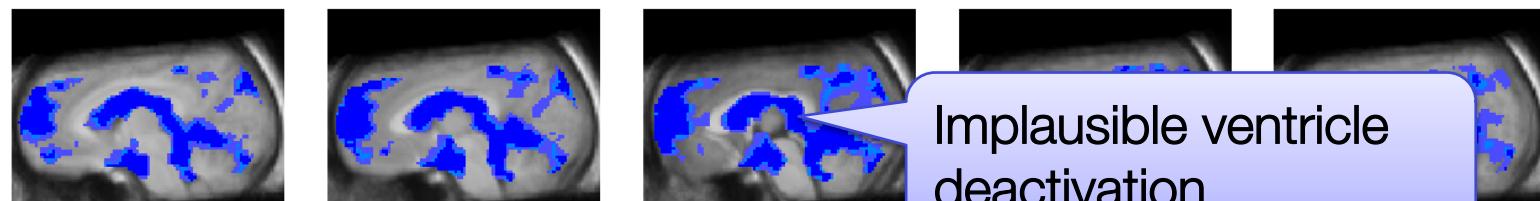
After nuisance regressors included: Much better!



Before vs. after nuisance covariates

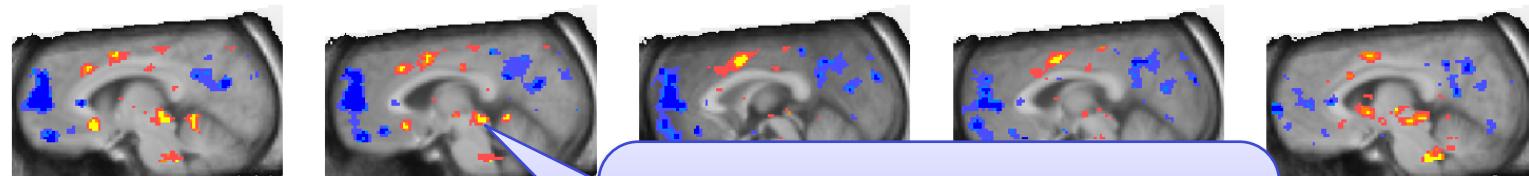
Contrast: CS+ vs. CS- during Shock Conditioning
Group analysis

Uncorrected



Implausible ventricle
deactivation

Corrected



Expected pattern of dorsal
cingulate increases and
“default-mode” decreases

End of Module



@fMRIstats

To scrub or not to scrub?

Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion

Jonathan D. Power ^{a,*}, Kelly A. Barnes ^a, Abraham Z. Snyder ^{a,b},
Bradley L. Schlaggar ^{a,b,c,d}, Steven E. Petersen ^{a,b,d,e}

^a Department of Neurology, Washington University School of Medicine, St. Louis, MO, USA

^b Department of Radiology, Washington University School of Medicine, St. Louis, MO, USA

^c Department of Pediatrics, Washington University School of Medicine, St. Louis, MO, USA

^d Department of Anatomy & Neurobiology, Washington University School of Medicine, St. Louis, MO, USA

^e Department of Psychology, Washington University in Saint Louis, St. Louis, MO, USA

1. Missing not at random

- Movement scans are not independent of "treatment assignment:" experimental conditions and subject characteristics.

2. Artificially alters error distribution and autocorrelation

3. interpolation not accounted for in stats, degrees of freedom, p-values

4. what is the ground truth for validating the scrubbing?

5. Ad hoc user decisions - hard to replicate choices, arbitrary threshold

* A matter of degree