

---

Computational  
Neuroscience  
Laboratory

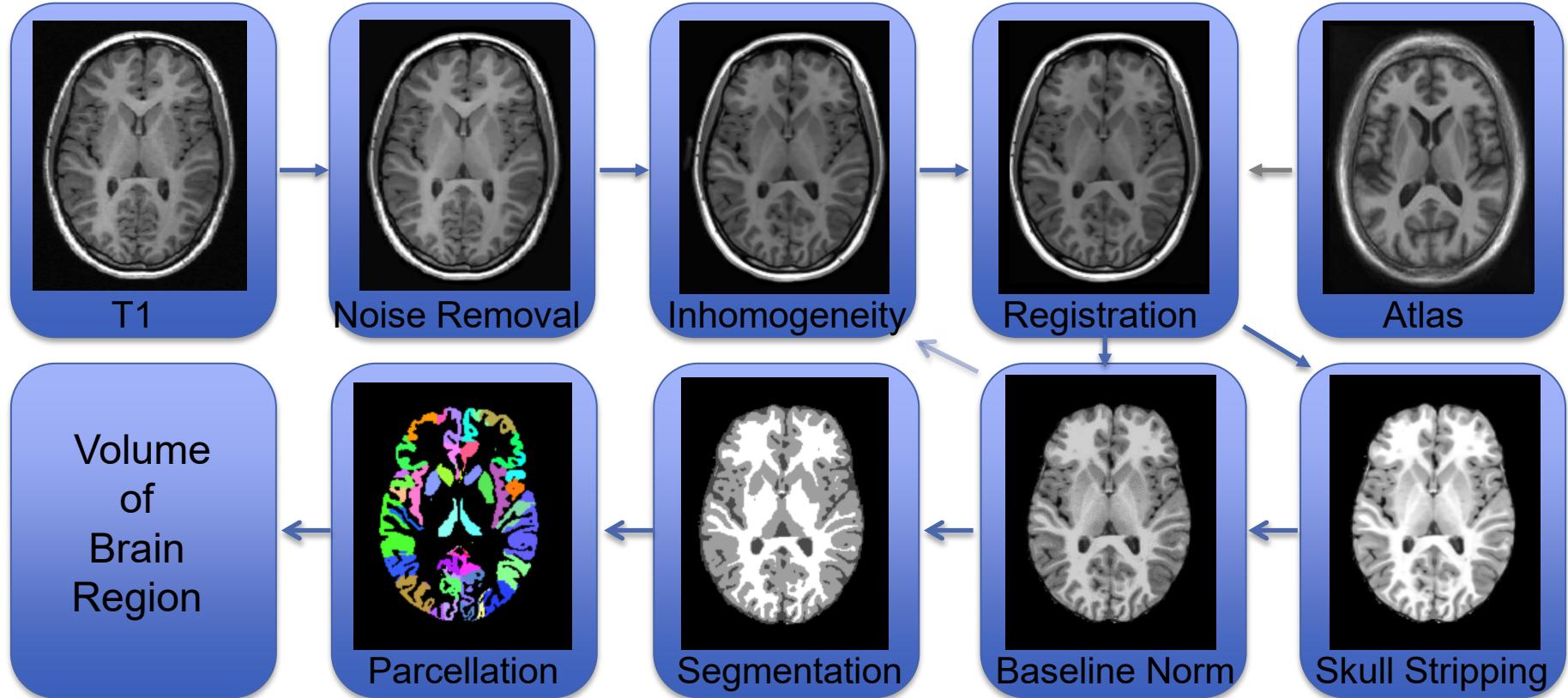
# Structural Group Analysis

Autumn 2023

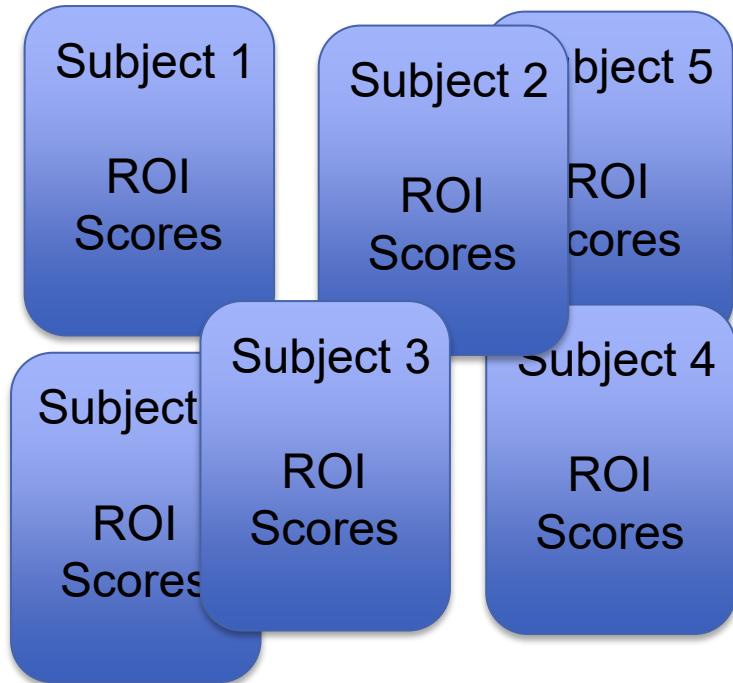
Session 9 – 10/24/2023



# Processing of Structural MRI



# Perform Group Analysis



Expand knowledge  
about disease

# Types of Group Analysis

## Correlation



a known score

Where in the brain are there associations between brain volume and test score?

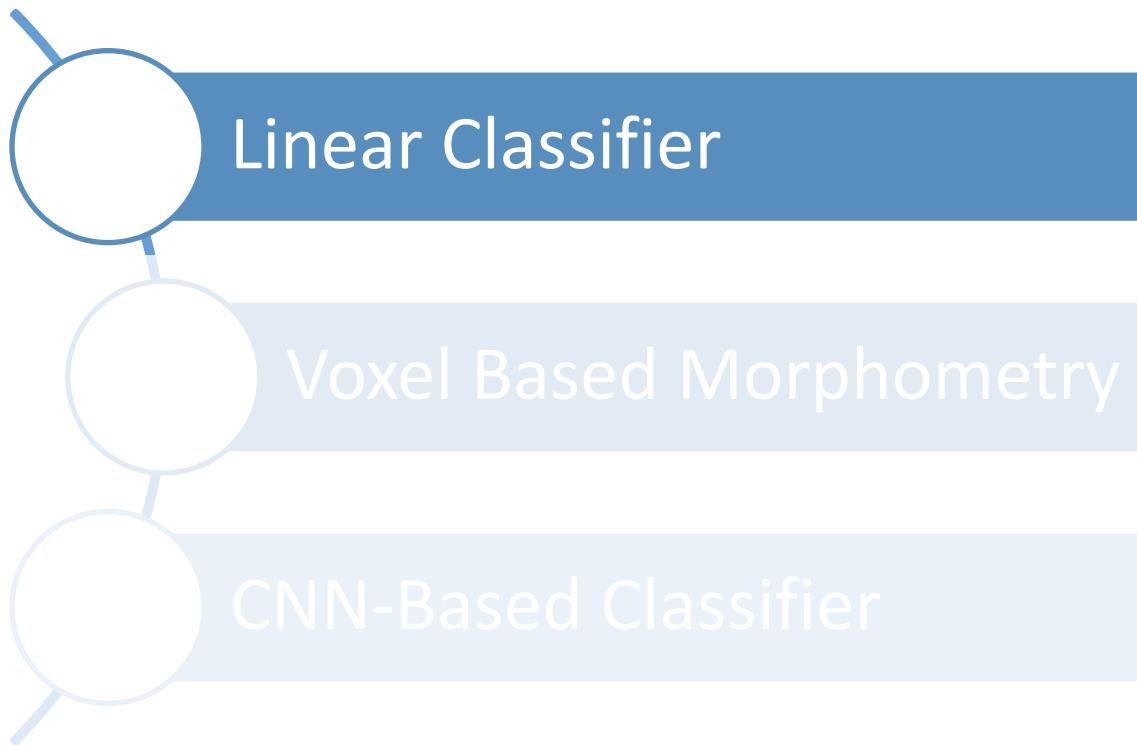
## Group comparison



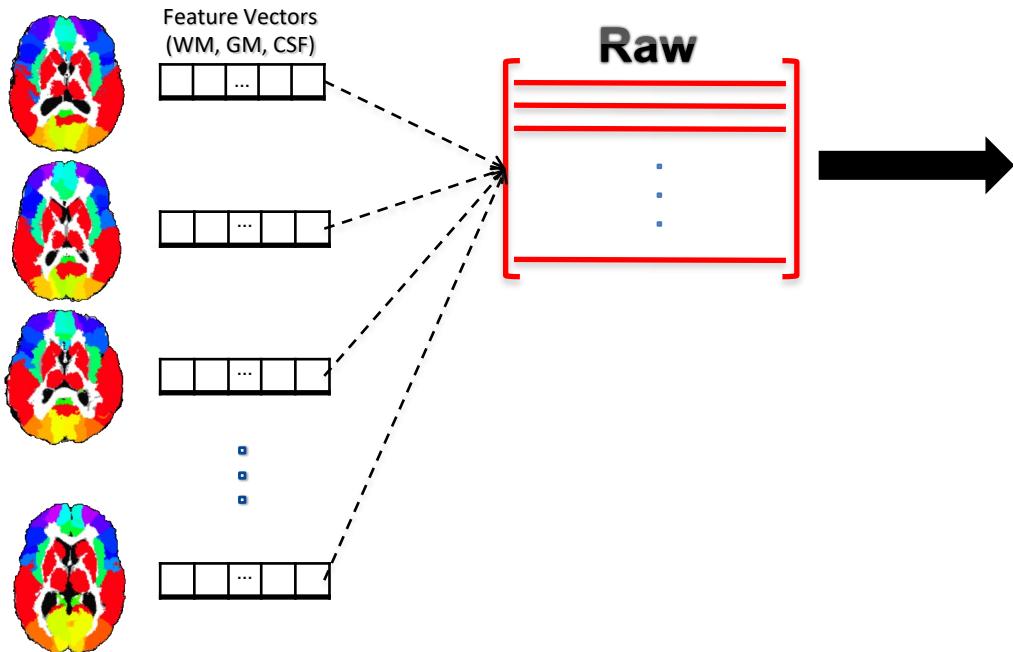
Where in the brain do the Simpsons and the Griffins have differences in brain volume?

Su & van Duin, Voxel Based Morphometry, UCL

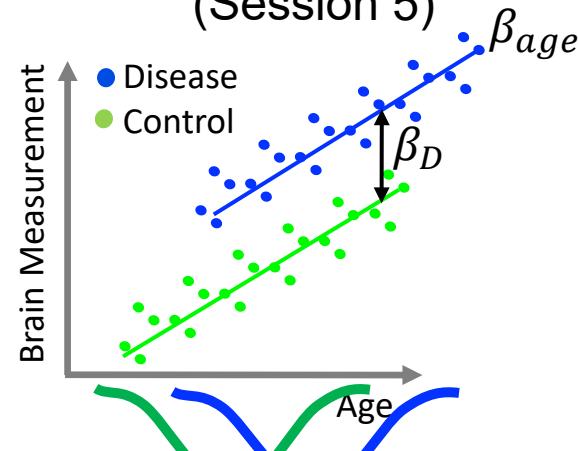
# Today...



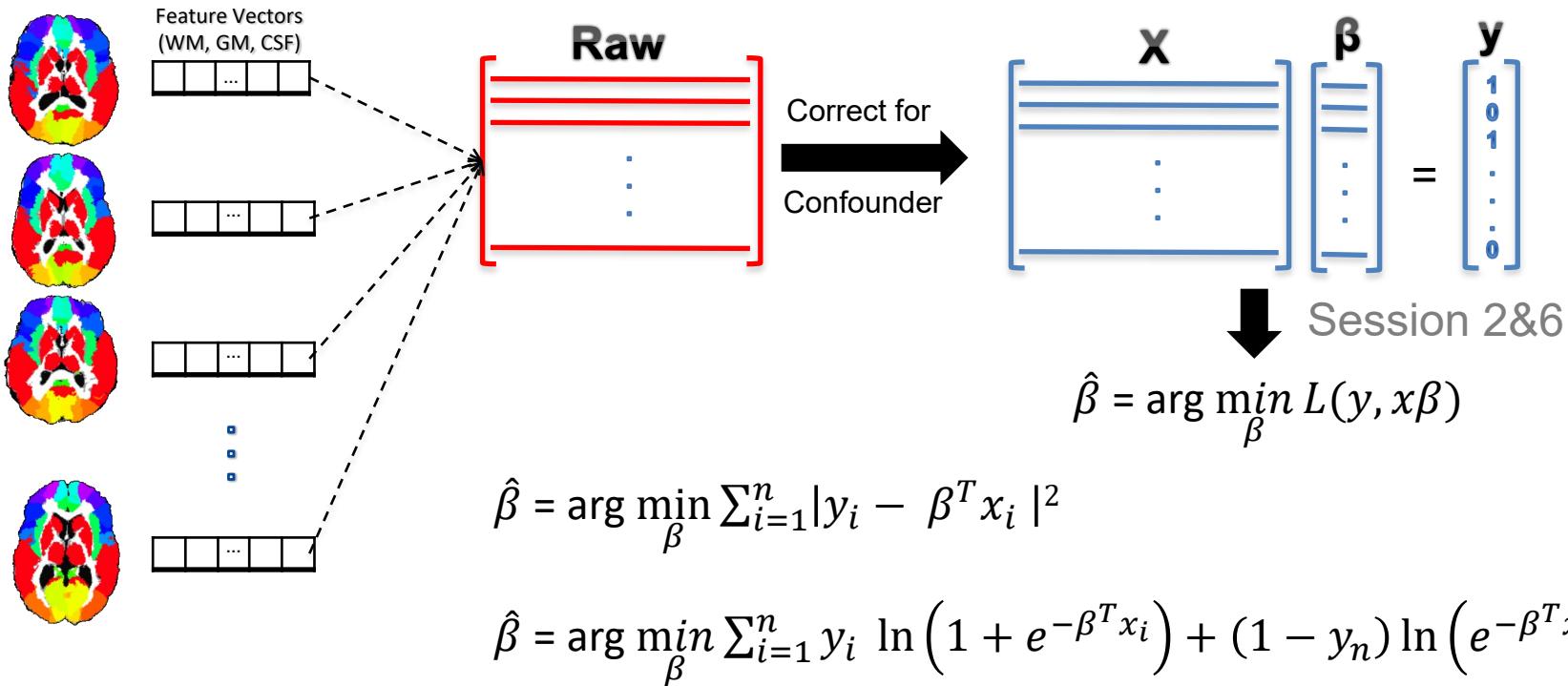
# ROI-Based Analysis



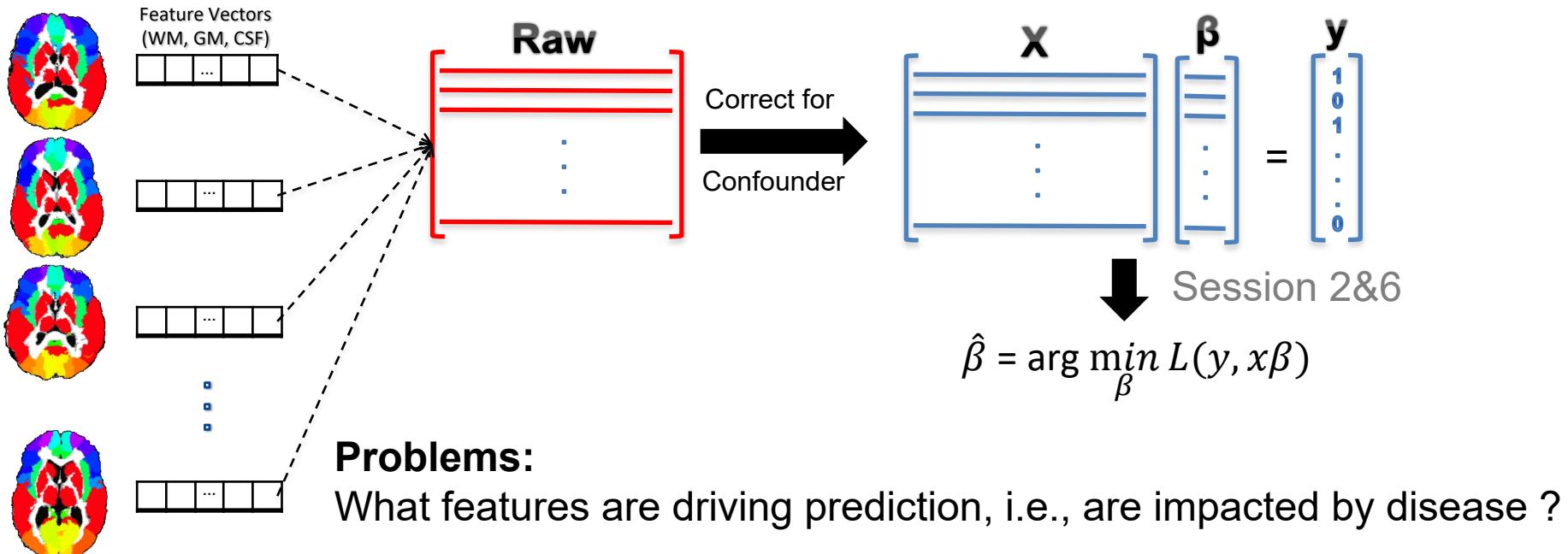
Correct for Confounder  
(Session 5)



# ROI-Based Analysis



# ROI-Based Analysis



# Enforce Sparsity via Regularization

$$\begin{bmatrix} \mathbf{x} \\ \vdots \\ \vdots \\ \mathbf{x} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta} \\ \vdots \\ \vdots \\ \boldsymbol{\beta} \end{bmatrix} = \begin{bmatrix} \mathbf{y} \\ 1 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} \rightarrow \hat{\boldsymbol{\beta}} = \arg \min_{\boldsymbol{\beta}} L(\mathbf{y}, \mathbf{x}\boldsymbol{\beta}) + \lambda R(\boldsymbol{\beta})$$

$R(\boldsymbol{\beta})$  should enforce sparsity of features  
(only select a few)

select  $\lambda$  during training via parameter exploration

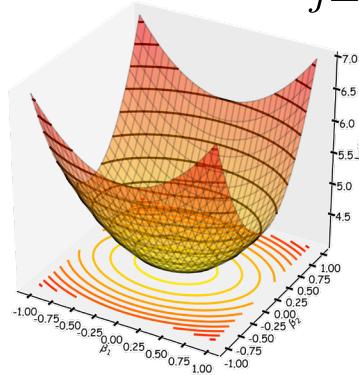
# Ridge Regularization

$$\begin{bmatrix} \mathbf{x} \\ \vdots \\ \vdots \\ \mathbf{x} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta} \\ \vdots \\ \vdots \\ \boldsymbol{\beta} \end{bmatrix} = \begin{bmatrix} \mathbf{y} \\ 1 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$



$$\hat{\boldsymbol{\beta}} = \arg \min_{\boldsymbol{\beta}} L(\mathbf{y}, \mathbf{x}\boldsymbol{\beta}) + \lambda R(\boldsymbol{\beta})$$

$$L_2 \text{ norm of the vector } \boldsymbol{\beta} : R(\boldsymbol{\beta}) = \sum_{j=1}^J \beta_j^2 = \|\boldsymbol{\beta}\|_2^2$$



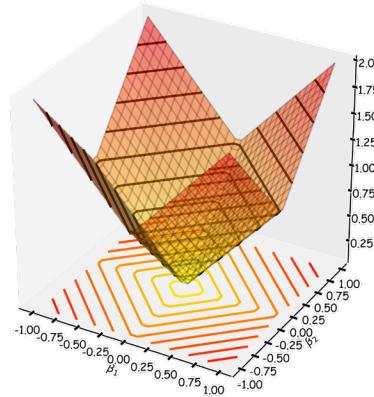
# Least Absolute Shrinkage & Selection Operator (LASSO)

$$\begin{bmatrix} \mathbf{x} \\ \vdots \\ \vdots \\ \mathbf{x} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta} \\ \vdots \\ \vdots \\ \boldsymbol{\beta} \end{bmatrix} = \begin{bmatrix} \mathbf{y} \\ 1 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$



$$\hat{\boldsymbol{\beta}} = \arg \min_{\boldsymbol{\beta}} L(\mathbf{y}, \mathbf{x}\boldsymbol{\beta}) + \lambda R(\boldsymbol{\beta})$$

$$l_1 \text{ norm of the vector } \boldsymbol{\beta} : R(\boldsymbol{\beta}) = \sum_{j=1}^J |\beta_j| = \|\boldsymbol{\beta}\|_1$$



# Least Absolute Shrinkage & Selection Operator (LASSO)

$$\begin{bmatrix} \mathbf{x} \\ \vdots \\ \vdots \\ \mathbf{x} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta} \\ \vdots \\ \vdots \\ \boldsymbol{\beta} \end{bmatrix} = \begin{bmatrix} \mathbf{y} \\ 1 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} \rightarrow \hat{\boldsymbol{\beta}} = \arg \min_{\boldsymbol{\beta}} L(\mathbf{y}, \mathbf{x}\boldsymbol{\beta}) + \lambda R(\boldsymbol{\beta})$$

$$l_1 \text{ norm of the vector } \boldsymbol{\beta} : R(\boldsymbol{\beta}) = \sum_{j=1}^J |\beta_j| = \|\boldsymbol{\beta}\|_1$$

Problem: How many features are important ?

- Number of selected features depends on training data

Possible solutions:

- Only select features with weights above a threshold:

Problem: classification is based on additional features  
(Haufe et al., 2014; Sabuncu, 2014)

- Set  $\lambda$  so only fixed number of measures are identified

Problem:  $\lambda$  is data dependent so that patterns across folds cannot be compared

## "L<sub>0</sub> Norm"

$$\begin{bmatrix} \mathbf{x} \\ \vdots \\ \vdots \\ \mathbf{x} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta} \\ \vdots \\ \vdots \\ \boldsymbol{\beta} \end{bmatrix} = \begin{bmatrix} \mathbf{y} \\ 1 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$



$$\hat{\boldsymbol{\beta}} = \arg \min_{\boldsymbol{\beta}} L(\mathbf{y}, \mathbf{x}\boldsymbol{\beta}) + \lambda R(\boldsymbol{\beta})$$

$L_1$  norm and  $L_2$  norm are just approximation of sparsity that do not solve the problem

solve original sparsity problem using the " $L_0$  norm", i.e.

$$\hat{\boldsymbol{\beta}} = \arg \min_{\boldsymbol{\beta}} L(\mathbf{y}, \mathbf{x}\boldsymbol{\beta}) \text{ s.t. } \|\boldsymbol{\beta}\|_0 \leq k$$

where  $\|\boldsymbol{\beta}\|_0$  counts the number of non-zero entries in  $\boldsymbol{\beta}$

Bonus Question: Why is  $\|\cdot\|_0$  not a norm ?

$$\|\alpha x\| \neq \alpha \|x\|$$

## “L<sub>0</sub> Norm”

$$\begin{bmatrix} \mathbf{x} \\ \vdots \\ \vdots \\ \mathbf{x} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta} \\ \vdots \\ \vdots \\ \boldsymbol{\beta} \end{bmatrix} = \begin{bmatrix} \mathbf{y} \\ 1 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$



$$\hat{\boldsymbol{\beta}} = \arg \min_{\boldsymbol{\beta}} L(\mathbf{y}, \mathbf{x}\boldsymbol{\beta}) + \lambda R(\boldsymbol{\beta})$$

$L_1$  norm and  $L_2$  norm are just approximation of sparsity that do not solve the problem

solve original sparsity problem using the “ $L_0$  norm”, i.e.

$$\hat{\boldsymbol{\beta}} = \arg \min_{\boldsymbol{\beta}} L(\mathbf{y}, \mathbf{x}\boldsymbol{\beta}) \text{ s.t. } \|\boldsymbol{\beta}\|_0 \leq k$$

where  $\|\boldsymbol{\beta}\|_0$  counts the number of non-zero entries in  $\boldsymbol{\beta}$

**How to solve this non-continuous minimization problem ?**

# Penalty Decomposition (Yong & Pohl, 2015)

$$\begin{bmatrix} \mathbf{x} \\ \vdots \\ \vdots \\ \mathbf{x} \end{bmatrix} \begin{bmatrix} \beta \\ \vdots \\ \vdots \\ \beta \end{bmatrix} = \begin{bmatrix} \mathbf{y} \\ 1 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} \rightarrow \hat{\beta} = \arg \min_{\beta} L(y, x\beta) \text{ s.t. } \|\beta\|_0 \leq k$$
$$\hat{\beta} = \arg \min_{\beta, w} L(y, xw) \text{ s.t. } \|\beta\|_0 \leq k \text{ and } w - \beta = 0$$

Solve  $w$  with fixed  $\beta'$

$$\arg \min_w L(y, xw) + \lambda \|\beta' - w\|_2^2$$

Gradient Descent

Solve  $\beta$  with fixed  $w'$

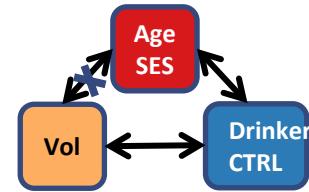
$$\arg \min_{\beta} \|\beta - w'\|_2^2 \text{ s.t. } \|\beta\|_0 \leq k$$

Closed Form Solution

# Example: NCANDA

Park et al. Scientific Reports, 2017

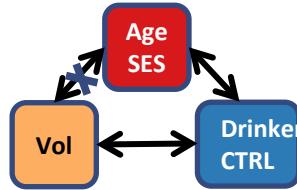
- 671 no alcohol drinking adolescent vs. 34 regular alcohol drinking adolescent
- 32 Freesurfer Structural Scores and 112 regional mean TBSS Fa, L1,Lt, MD scores
- Groups are matched with respect to sex, race, manufacturer, supratentorial volume  
**but not age and socio-economic status (SES)**



# Findings

Park et al. Scientific Reports, 2017

Matched Accuracy = Accuracy computed on 34 CTRLS  
and 34 regular drinkers matched  
with respect to age and SES.



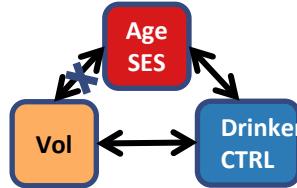
Method	Sens	Spec	Accuracy
Raw	20.6	94.6	57.4
Sequential	32.4	94.0	62.9

Labels predicted by sparse classifier based on “L<sub>0</sub>-norm”

# Findings

Park et al. Scientific Reports, 2017

Age Test = Accuracy of predictions in distinguishing young from old participants is insignificantly better than random assignment



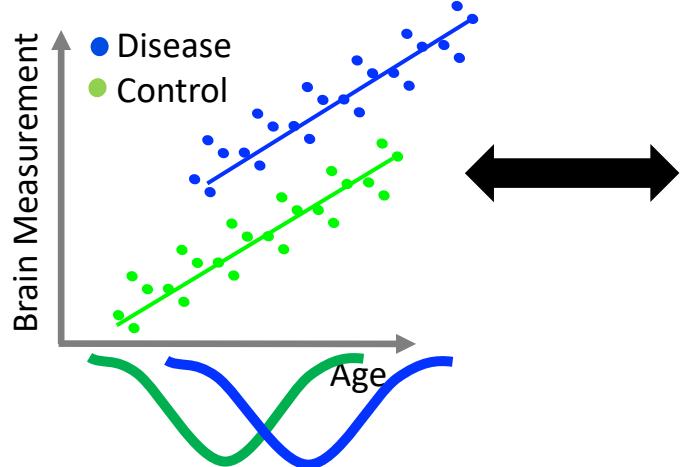
Method	Sens	Spec	Accuracy	Matched Accuracy	Age Test
Raw	20.6	94.6	57.4	57.4	Failed
Sequential	32.4	94.0	62.9	58.8	Failed

**Predictions are still confounded**

Labels predicted by sparse classifier based on “L<sub>0</sub>-norm”

# Joint Confounder Correction & Classification

Correct for Confounder  
(Session 5)

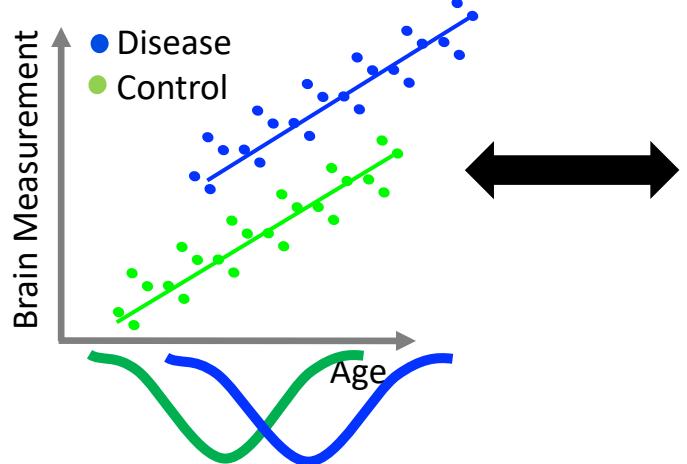


Let  $m$  be the meta data (such as age of subjects)

$$\begin{matrix} \mathbf{x} \\ \vdots \\ \mathbf{x} \end{matrix} \quad \begin{matrix} \boldsymbol{\beta} \\ \vdots \\ \boldsymbol{\beta} \end{matrix} = \begin{matrix} \mathbf{y} \\ \vdots \\ \mathbf{y} \end{matrix}$$

# Joint Confounder Correction & Classification

Correct for Confounder  
(Session 5)



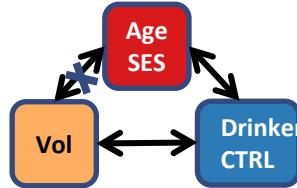
Let  $m$  be the meta data (such as age of subjects)

$$\text{GLM} \quad [x, m] \beta = y$$

# Findings

Park et al. Scientific Reports, 2017

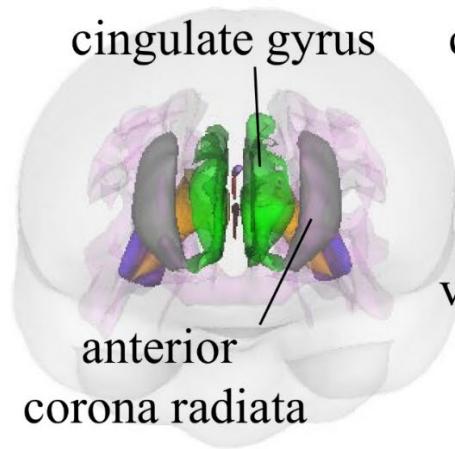
Age Test = Accuracy of predictions in distinguishing young from old participants is insignificantly better than random assignment



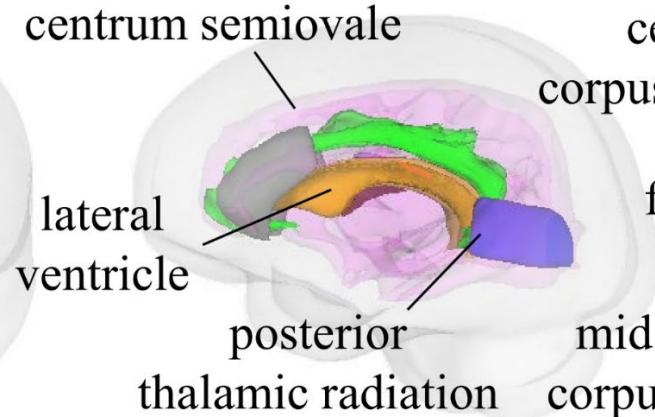
Method	Sens	Spec	Accuracy	Matched Accuracy	Age Test
Raw	20.6	94.6	57.4	57.4	Failed
Sequential	32.4	94.0	62.9	58.8	Failed
Joint	67.6	84.2	75.9	76.5	Passed

Labels predicted by sparse classifier based on “L<sub>0</sub>-norm”

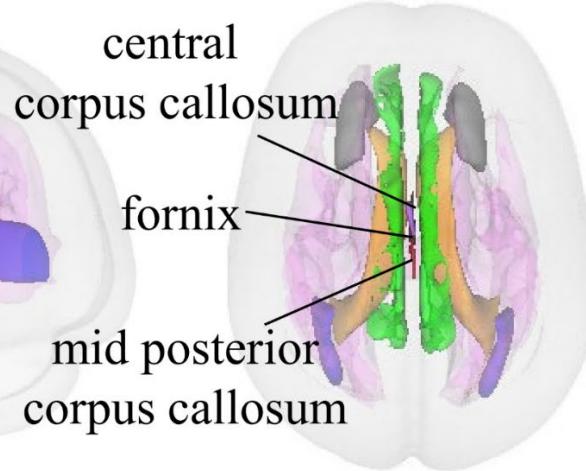
# 8 Regions Distinguishing No-low from Regular Drinkers



Anterior



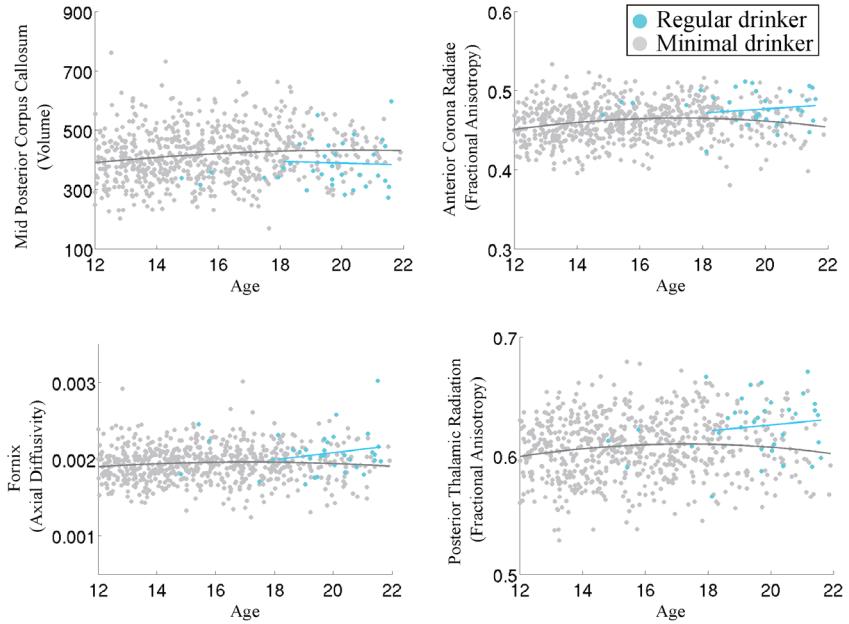
Left



Superior

Important to check output of regression-based approaches for the effects of confounding factors

# Regional Scores



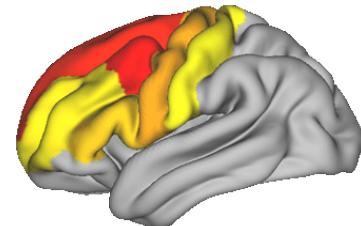
Machine learning requires analyzing MRI metrics as a whole to gain knowledge about the effect of alcohol on individuals

# Problem: Imbalanced Data Set

	Total	sex		Age (years)	svol ( $\times 10^6$ )
		F	M		
CTRL	245	122	123	45.59 $\pm$ 17.17	1.27 $\pm$ 0.13
ALC	226	67	159	48.49 $\pm$ 10.04	1.27 $\pm$ 0.11
HIV	65	20	45	51.81 $\pm$ 8.44	1.27 $\pm$ 0.15
ALC+HIV	66	23	43	50.97 $\pm$ 8.12	1.23 $\pm$ 0.14

298 MRI measurements based on SRI24 atlas and Freesurfer sex, age, and supratentorial volume (svol) were regressed out

**How to select informative features and deal with imbalance ?**



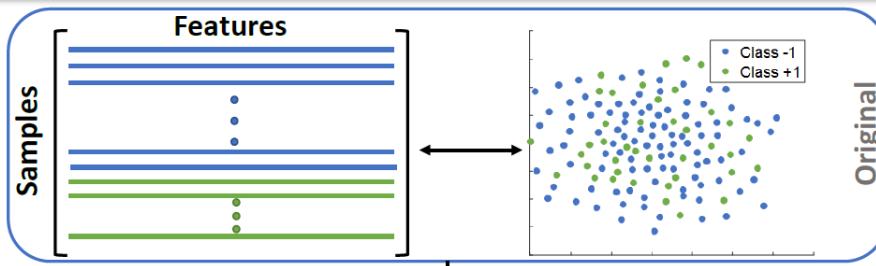
# Select Features and Samples

Adeli et al., IEEE PAMI 2019

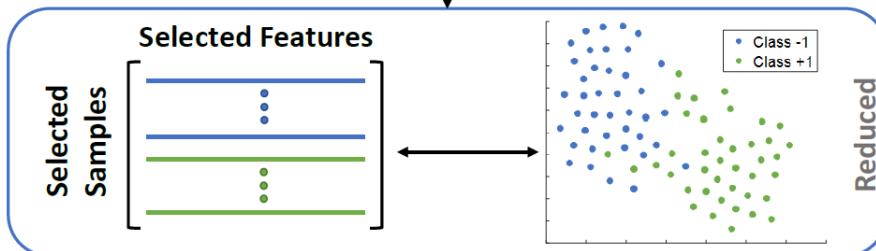
Adeli et al., Biological Psychiatry: CNNI, 2019

## Challenges

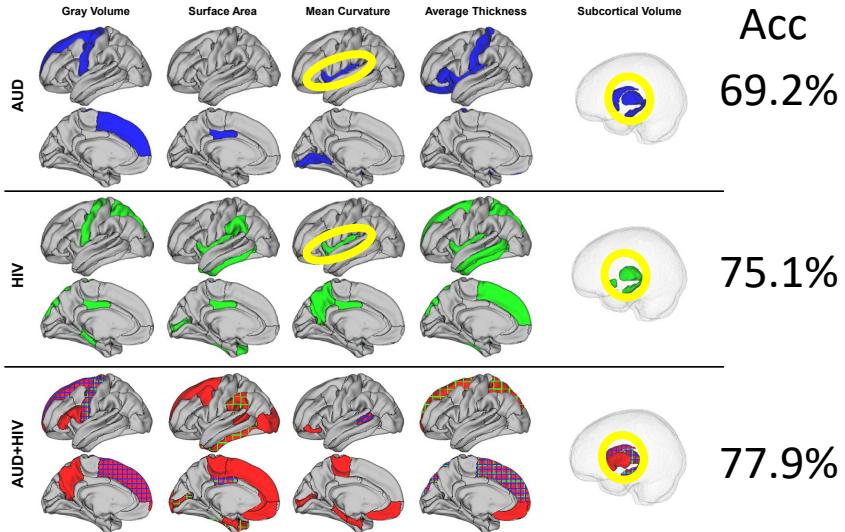
- Redundant Features
- Noisy data
- Skewed Class Distributions (imbalanced classification)



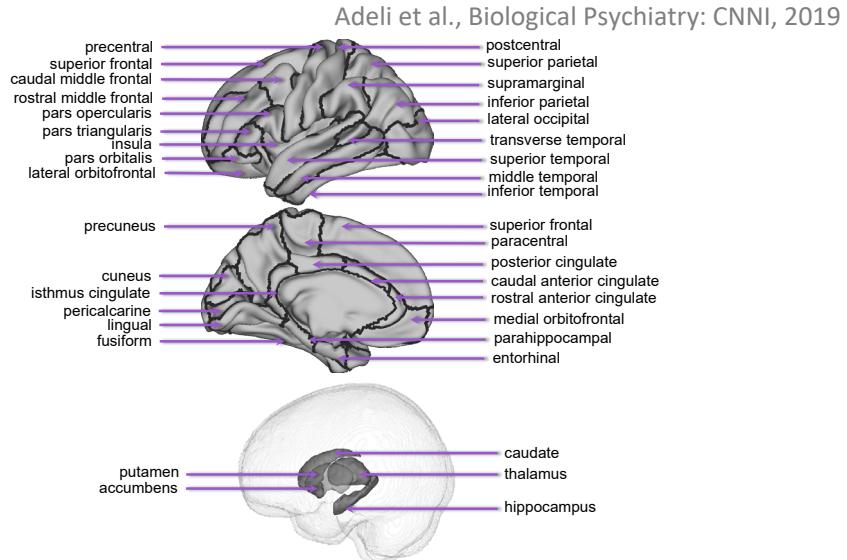
Logistic Regression Confined by Sample and Feature Selection



# Diagnostic Patterns



Study supports compounding effect of AUD and HIV as the pattern for AUD&HIV is the largest and lead to the most accurate predictions



# Diagnostic Score

Controls

AUD

Controls

HIV

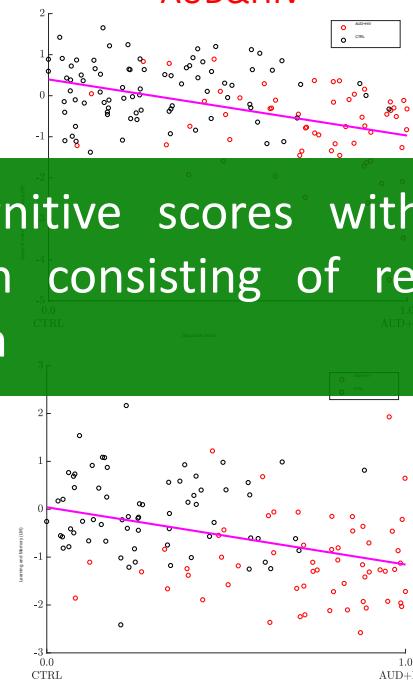
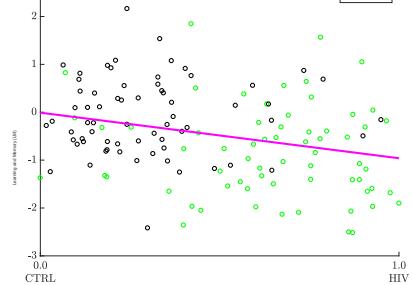
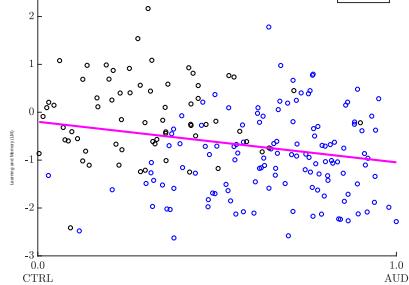
Adeli et al., Biological Psychiatry: CNNI, 2019

HIV

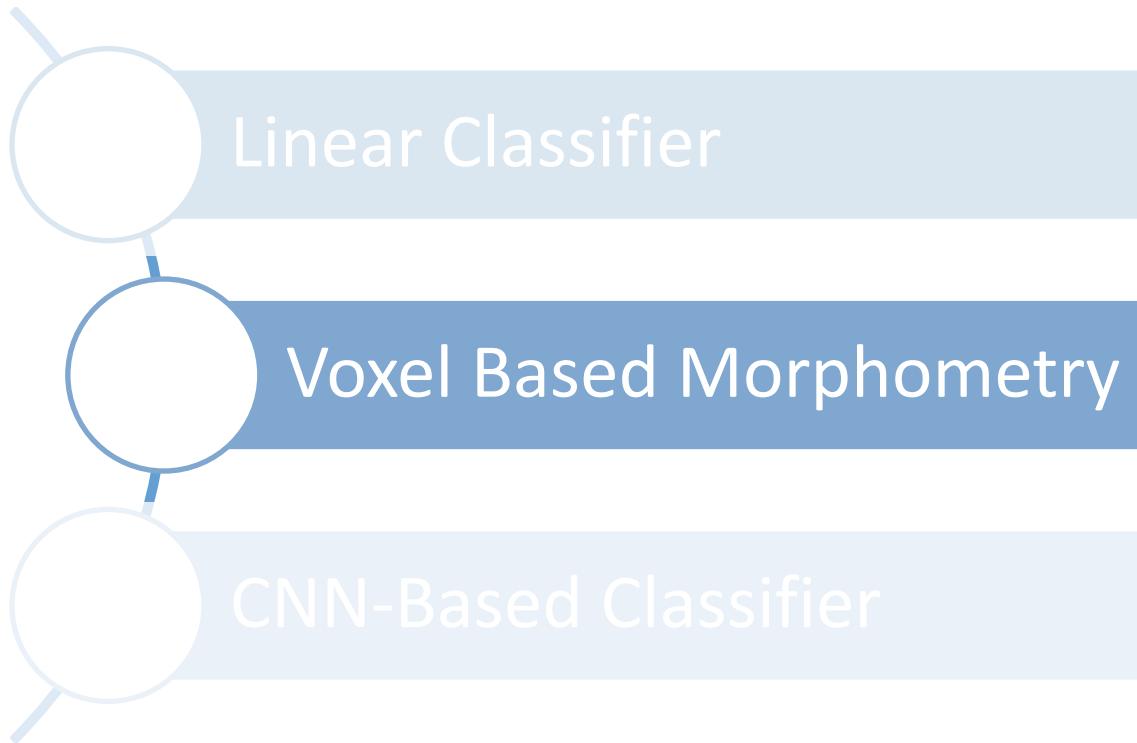
AUD&HIV



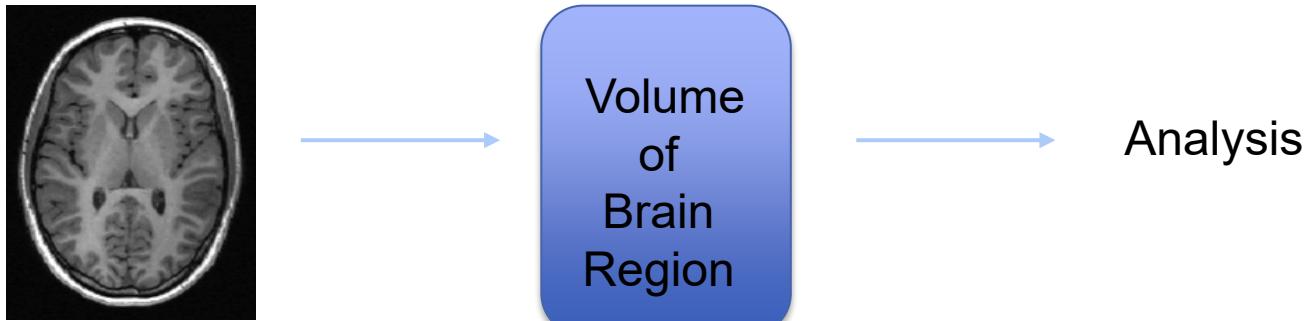
Diagnostic scores predicted cognitive scores with the corresponding diagnostic pattern consisting of regions closely linked to this brain function



# Today...



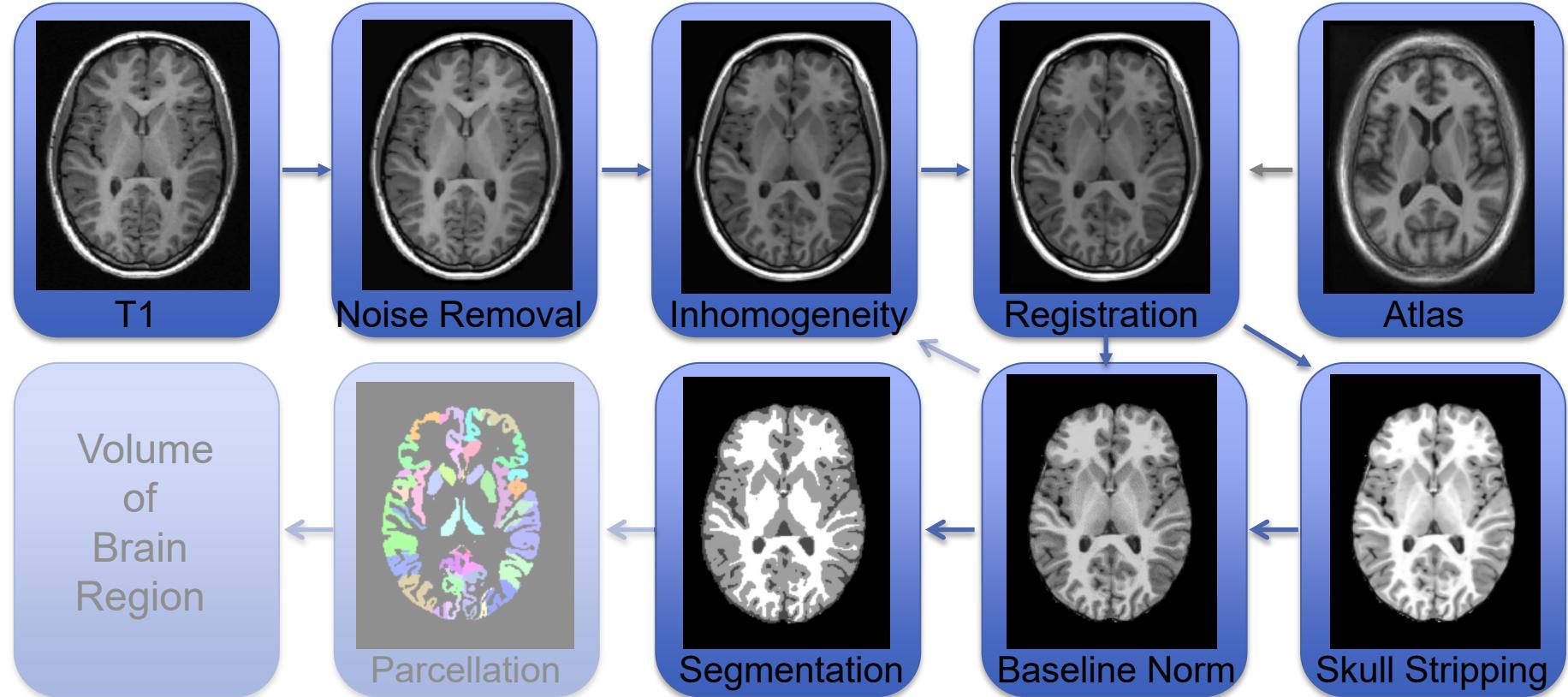
# How to identify group differences from MRIs



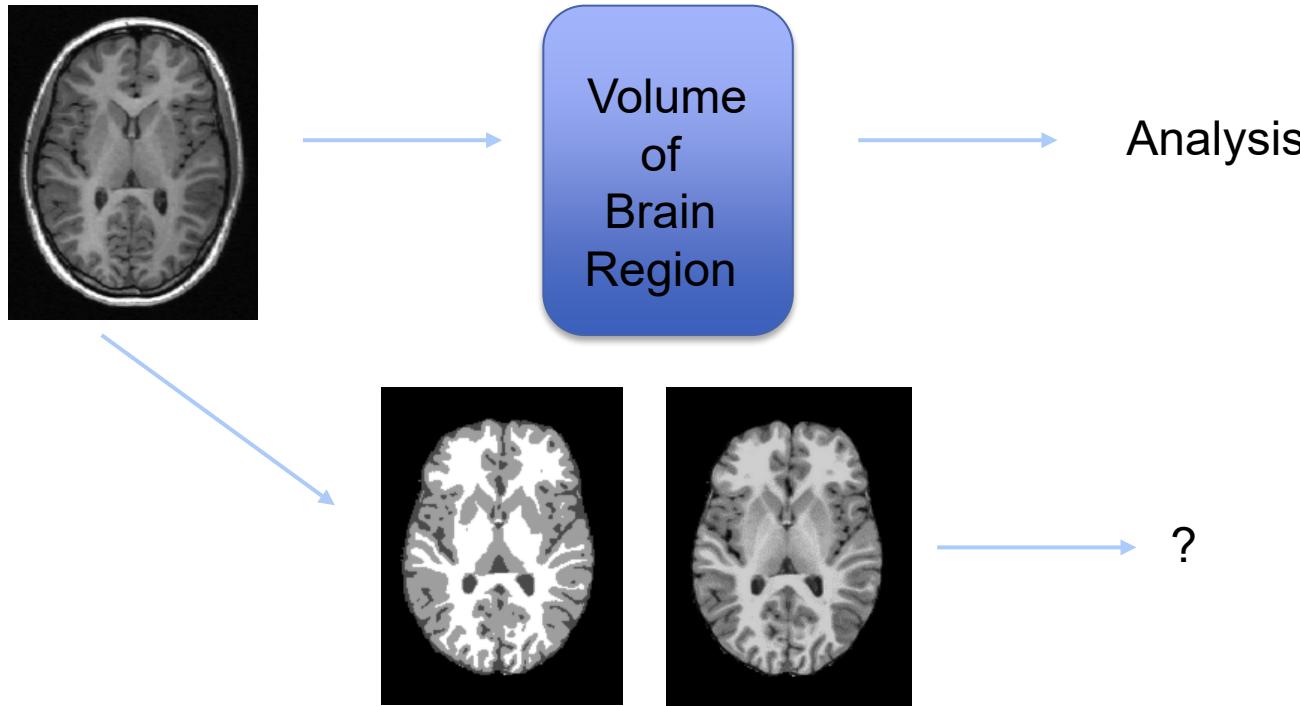
Problem:



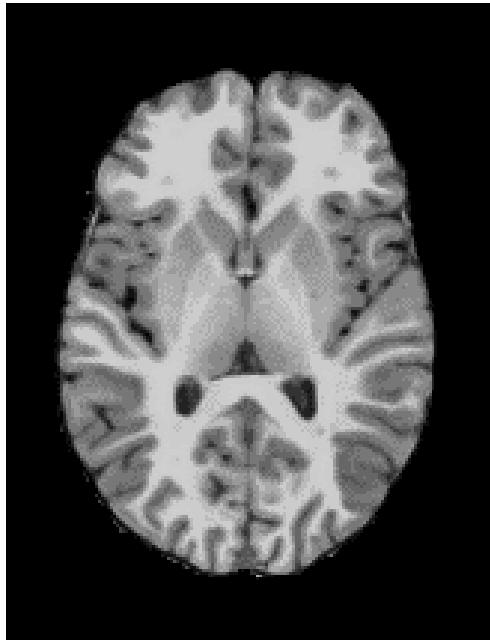
# Processing of Structural MRI



# How to identify group differences from MRIs



# Remember



MRI intensity are relative values that are only meaningful in within neighborhoods

# Align to Template

Individual T1



Group Template  
(Target)

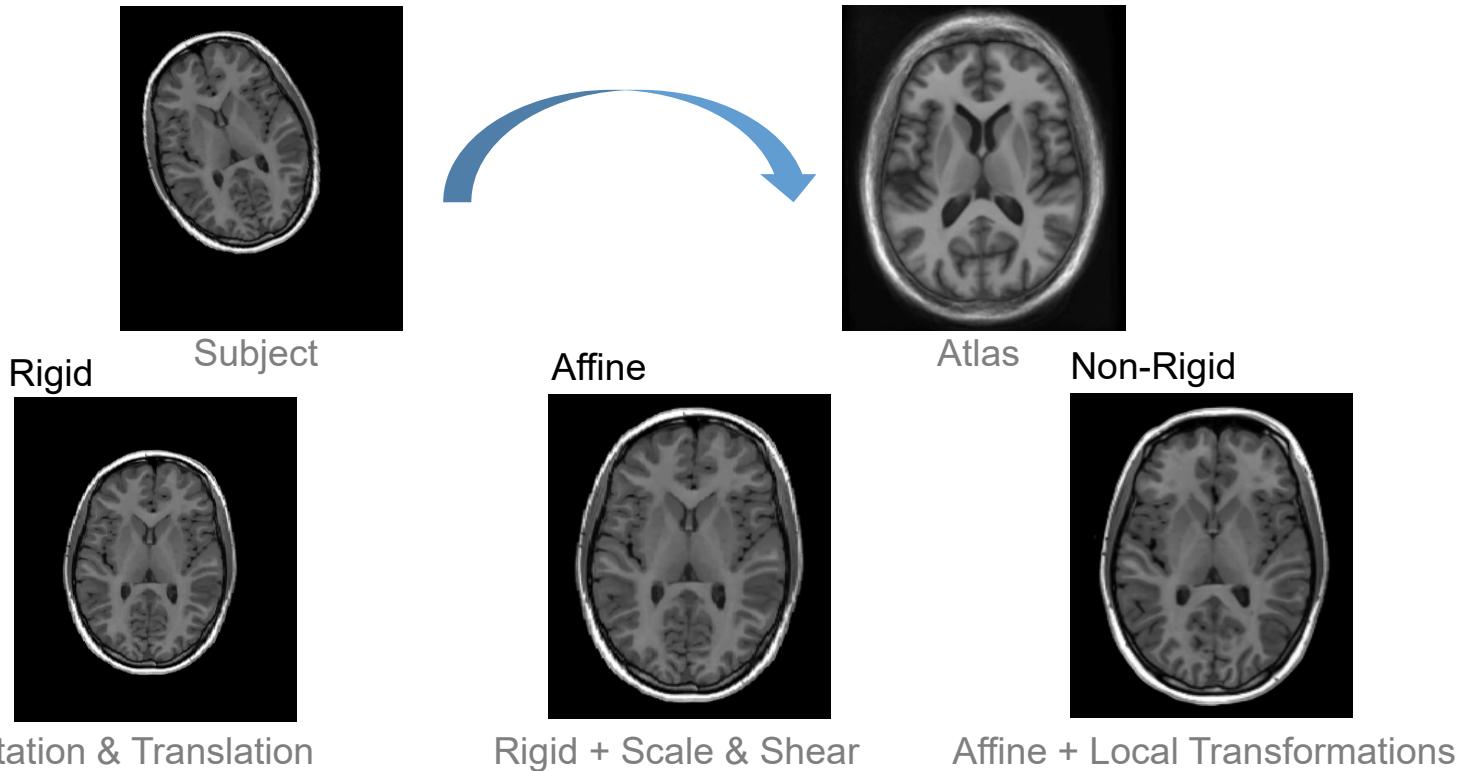


Non-Rigid  
Registration

**“Optimized VBM” Good, et al, NI, 2001. Douaud, et al, Brain. 2007.**

Slide from Voxel-based Morphometric Analysis, Martinos Center for Biomedical Imaging, MGH

# Align Atlas to Subject (or vice versa)



<http://stnava.github.io/ANTs>

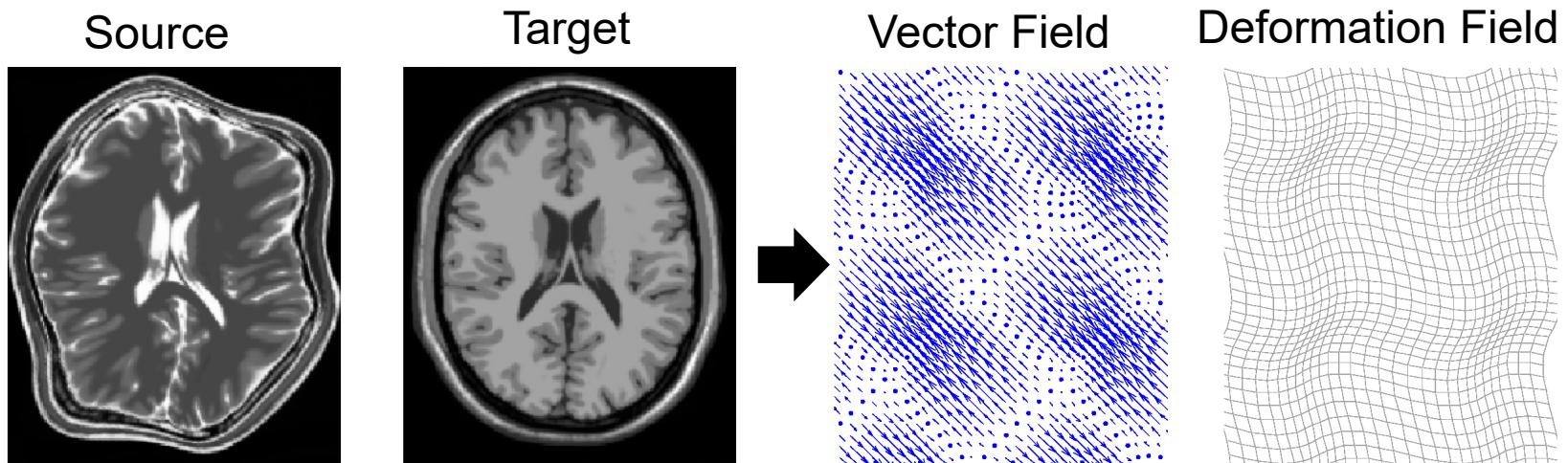
35

Machine Learning for Neuroimaging - Autumn 2023 – Session 9



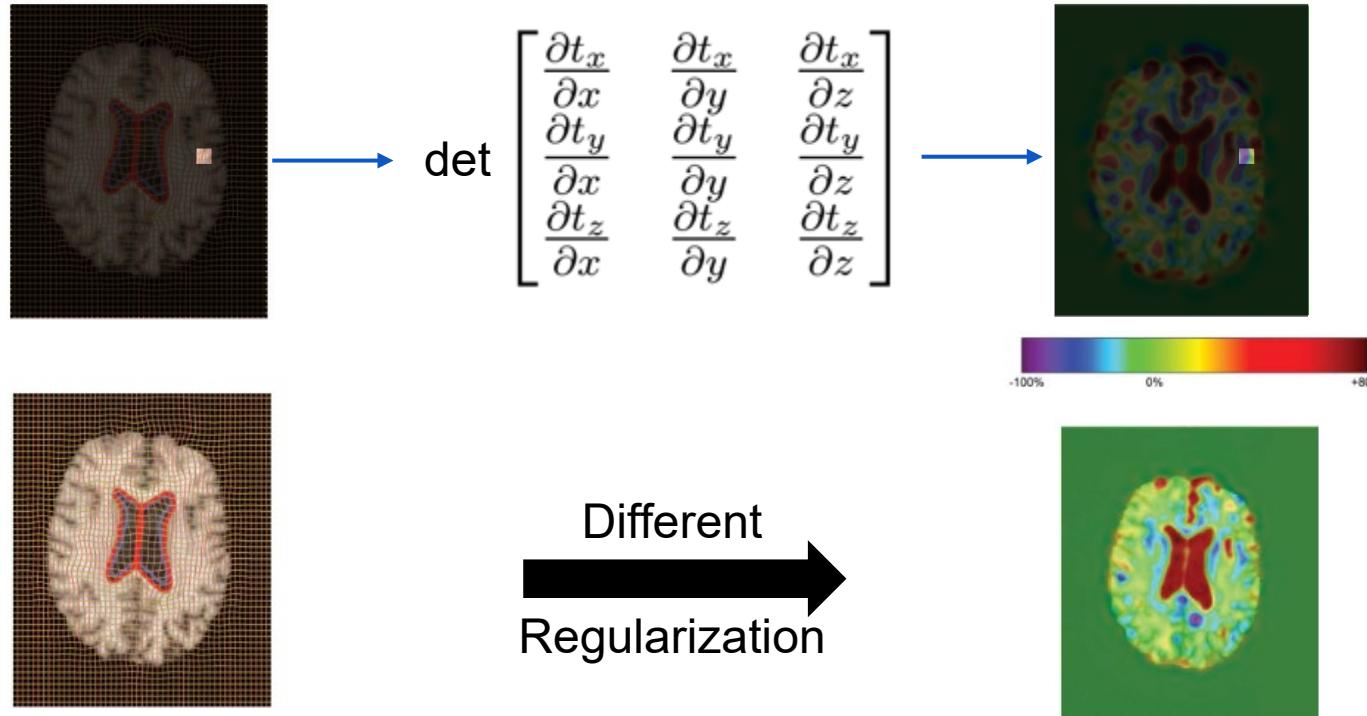
Computational Neuroscience  
Laboratory - Stanford

# Non-rigid Registration



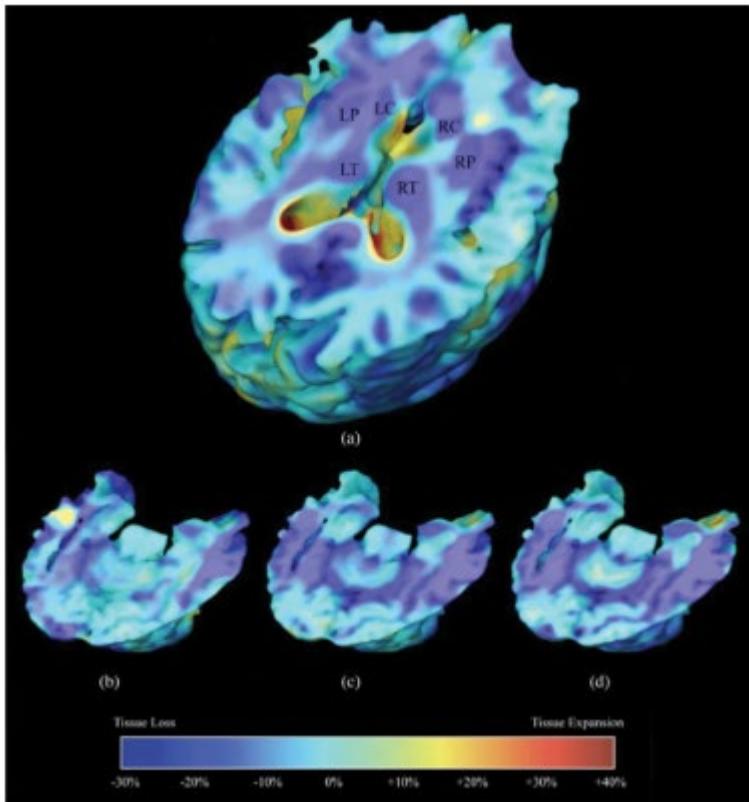
Li et al. *Entropy* 2019, 21, 189. <https://doi.org/10.3390/e2102018>

# Computing Jacobian Determinant



Leow et al. , IEEE TMI, 2007

# Jacobian Determinant



Leow et al. , IEEE TMI, 2007

# Align to Template

Individual T1



Group Template  
(Target)

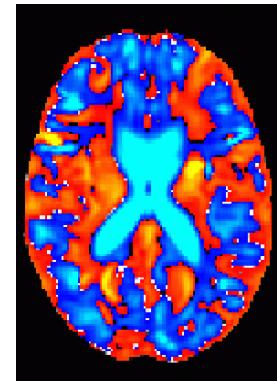


Non-Rigid  
Registration

Individual T1  
(Template Space)



Jacobian

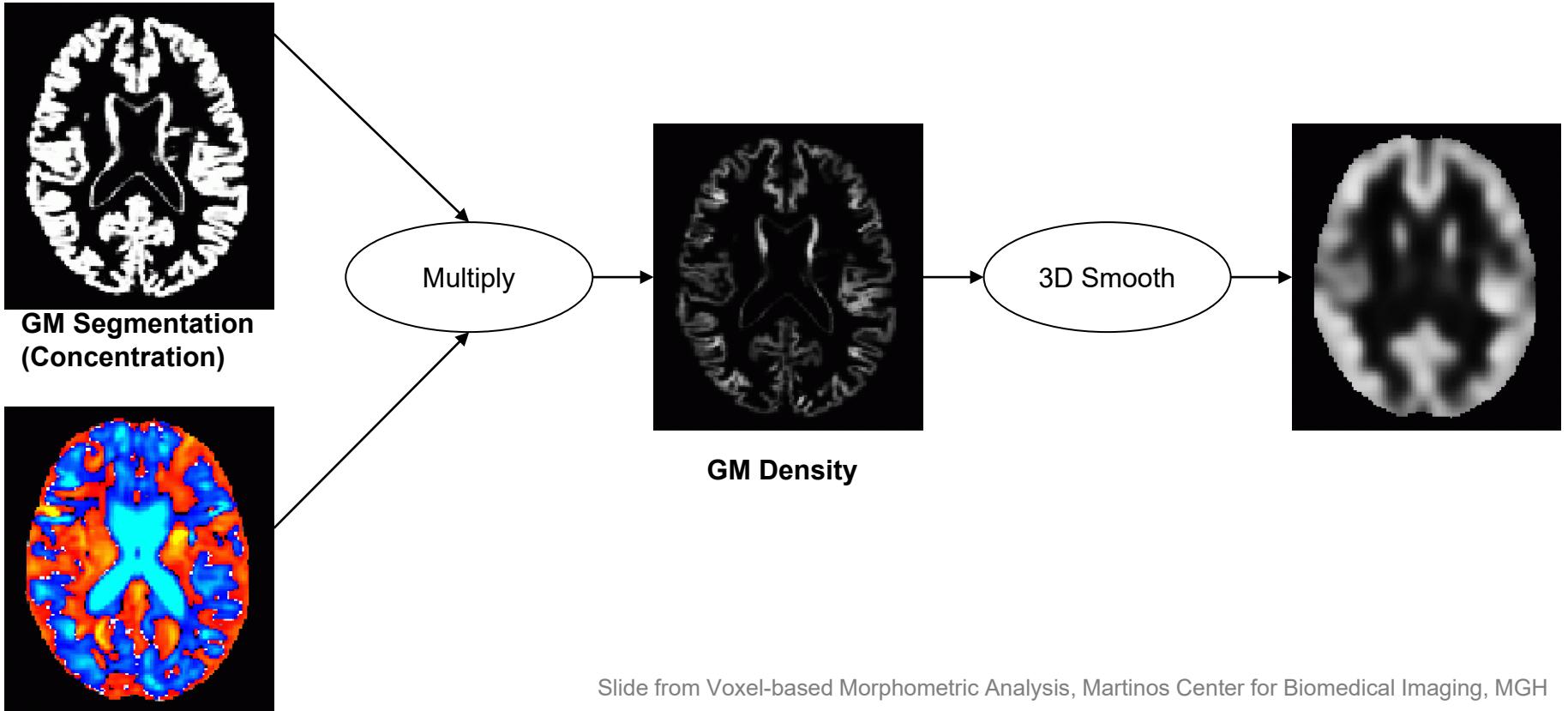


Expansion  
Compression

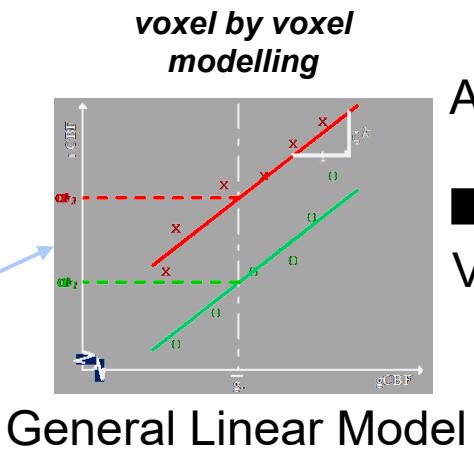
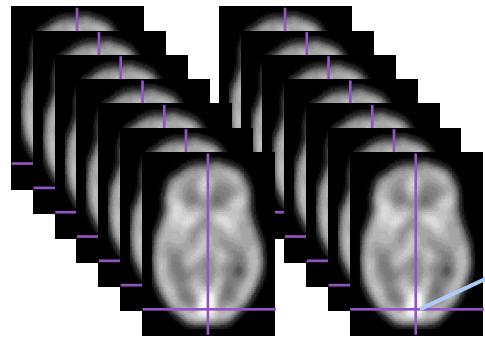
**“Optimized VBM” Good, et al, NI, 2001. Douaud, et al, Brain. 2007.**

Slide from Voxel-based Morphometric Analysis, Martinos Center for Biomedical Imaging, MGH

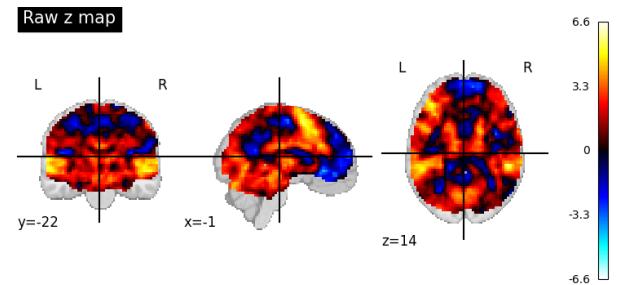
# Modulation and Smoothing



# General Linear Model at a voxel level



Across  
all  
Voxels



Statistical Parametric Mapping (SPM)

Example:

[https://nilearn.github.io/dev/auto\\_examples/05\\_glm\\_second\\_level/plot\\_thresholding.html](https://nilearn.github.io/dev/auto_examples/05_glm_second_level/plot_thresholding.html)

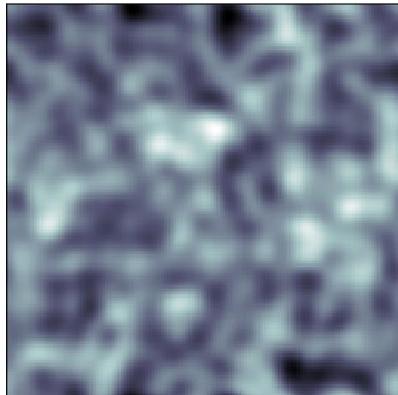
# False Positive

- One t-test with  $p < .05$ 
  - a 5% chance of (at least) one false positive
- 3 t-tests, all at  $p < .05$ 
  - All have 5% chance of a false positive
  - So actually you have  $3 \times 5\%$  chance of a false positive  
= 15% chance of introducing a false positive
- In VBM, depending on your resolution
  - 1000000 voxels
  - 1000000 statistical tests
- do the maths at  $p < .05!$ 
  - 50000 false positives
- So what to do?
  - Bonferroni Correction
  - Random Field Theory/ Family-wise error (used in SPM)

# Multiple Comparison Correction

**Bonferroni-Correction** (controls false positives at individual voxel level):

- divide desired p value by number of comparisons
- $.05/1000000 = p < 0.0000005$  at every single voxel
- Not a brilliant solution (false negatives)!
- Added problem of spatial correlation
  - data from one voxel will tend to be similar to data from nearby voxels



Z-Scores

# Multiple Comparison Correction

**Bonferroni-Correction** (controls false positives at individual voxel level):

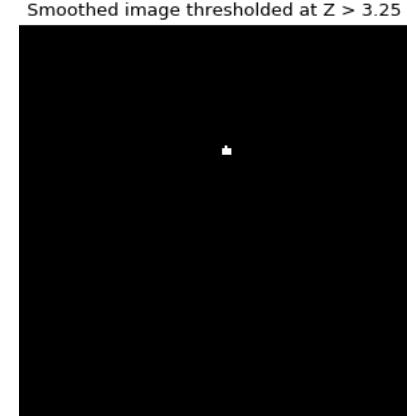
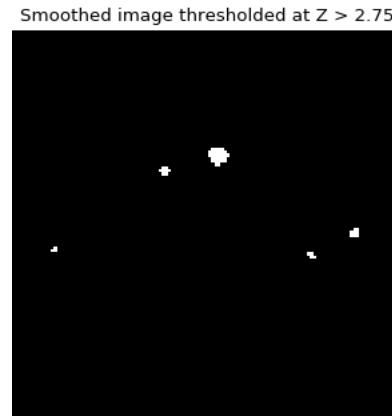
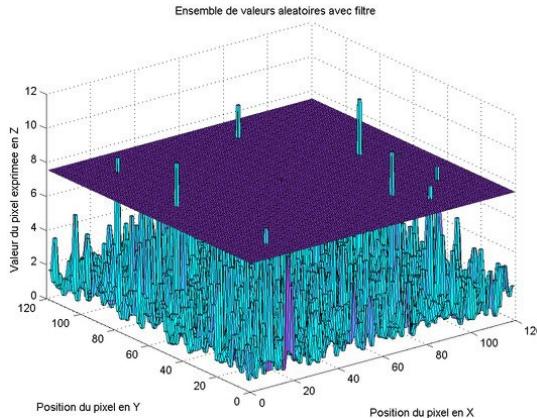
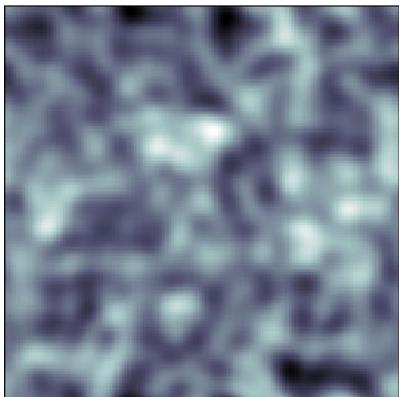
Added problem of spatial correlation: data from one voxel will tend to be similar to data from nearby voxels

## Family Wise Error (FWE) / Spatial Extent Methods

Use **Gaussian Random Field** theory to determine right threshold for a smooth statistical map which gives the required **FWE**, i.e., the probability that one or more of the significance tests results is a false positive within the volume of interest.

- effectively controls the number of false positive regions rather than voxels

<https://www.ehu.eus/ccwintco/uploads/4/4c/Spm-rft-slides-poirrier06.pdf>



# Multiple Comparison Correction

**Bonferroni-Correction** (controls false positives at individual voxel level):

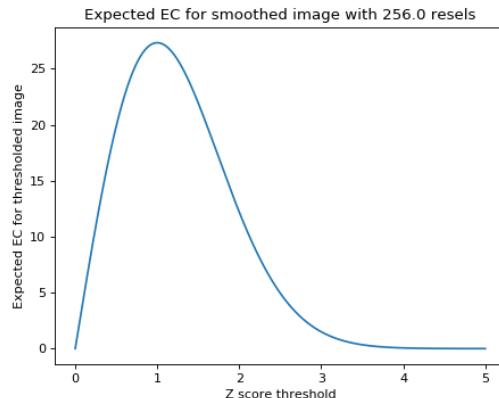
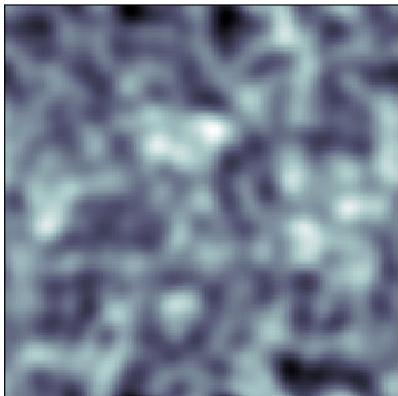
Added problem of spatial correlation: data from one voxel will tend to be similar to data from nearby voxels

## Family Wise Error (FWE) / Spatial Extent Methods

Use **Gaussian Random Field** theory to determine right threshold for a smooth statistical map which gives the required **FWE**, i.e., the probability that one or more of the significance tests results is a false positive within the volume of interest.

- effectively controls the number of false positive regions rather than voxels

<https://www.ehu.eus/ccwintco/uploads/4/4c/Spm-rft-slides-poirrier06.pdf>



Euler characteristic (EC):  
expected number of islands generated by chance  
 $EC=0.05 \Rightarrow Z=4.05$

Nothing would be significant in example

# Multiple Comparison Correction

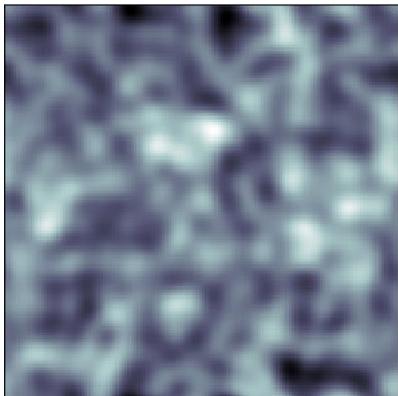
**Bonferroni-Correction** (controls false positives at individual voxel level):

Added problem of spatial correlation: data from one voxel will tend to be similar to data from nearby voxels

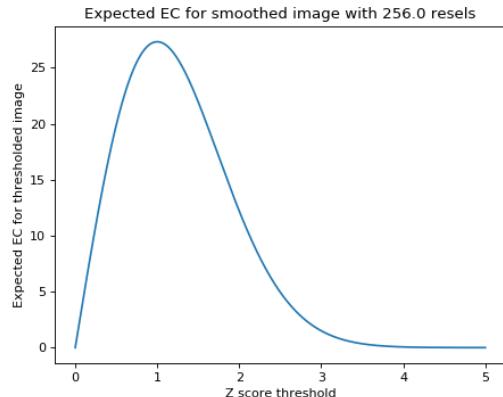
## Family Wise Error (FWE) / Spatial Extent Methods

Use **Gaussian Random Field** theory to determine right threshold for a smooth statistical map which gives the required **FWE**, i.e., the probability that one or more of the significance tests results is a false positive within the volume of interest.

- effectively controls the number of false positive regions rather than voxels
- relates to minimum size of cluster (see <https://www.ehu.eus/ccwintco/uploads/4/4c/Spm-rft-slides-poirrier06.pdf>)
- Good: a “safe” way to correct;  
Bad: probably missing a lot of true positives – smoothing requirement



Z-Scores



Euler characteristic (EC):  
expected number of islands generated by chance  
 $EC=0.05 \Rightarrow Z=4.05$

Nothing would be significant in example

# Multiple Comparison Correction

**Bonferroni-Correction** (controls false positives at individual voxel level):

Added problem of spatial correlation: data from one voxel will tend to be similar to data from nearby voxels

## Family Wise Error (FWE) / Spatial Extent Methods

Use **Gaussian Random Field** theory to determine right threshold for a smooth statistical map which gives the required **FWE**, i.e., the probability that one or more of the significance tests results is a false positive within the volume of interest.

- effectively controls the number of false positive regions rather than voxels
- relates to minimum size of cluster (see <https://www.ehu.eus/ccwintco/uploads/4/4c/Spm-rft-slides-poirrier06.pdf>)
- Good: a “safe” way to correct;  
Bad: probably missing a lot of true positives – smoothing requirement

## False Discovery Rate (FDR)

Controls the expected proportion of false positives among suprathreshold voxels (i.e., labeled as significant) only

- Using FDR,  $q < 0.05$ : we expect 5% of those voxels for each SPM to be false positives (1,000 voxels)
- Bad: less stringent than FWE so more false positives  
Good: fewer false negatives (i.e., more true positives), no smoothness requirement

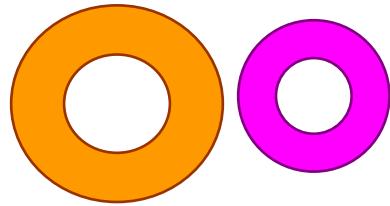
More lenient may be better for smaller studies – often used for fMRI studies

## Small Volume Correction (a.k.a. spotlight)

- Select regions based on hypothesis that is ideally motivated by previous work
- Reduces the number of comparisons
- Increases the chance of identifying significant voxels in a ROI

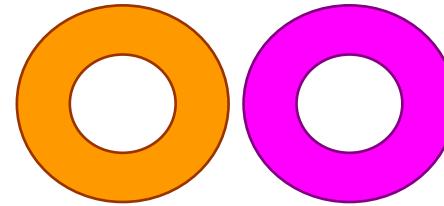
# Other issues with VBM

- Controlling for total intracranial volume (TIV)
- Uniformly bigger brains may have uniformly more GM/ WM



brain A      brain B

Differences without  
accounting for TIV

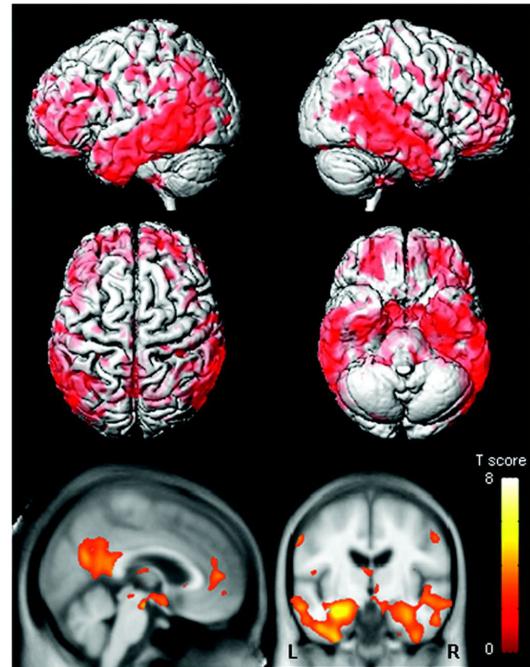


brain A      brain B

differences after TIV has been “covaried  
out” (Differences uniformly distributed with hardly  
any impact at local level)

Including total GM or WM volume as a  
covariate adjusts for global atrophy and looks  
for regionally-specific changes

# Atrophy associated with progression to AD in amnestic mild cognitive impairment



Whitwell et al., Neurology 2007

# Summary

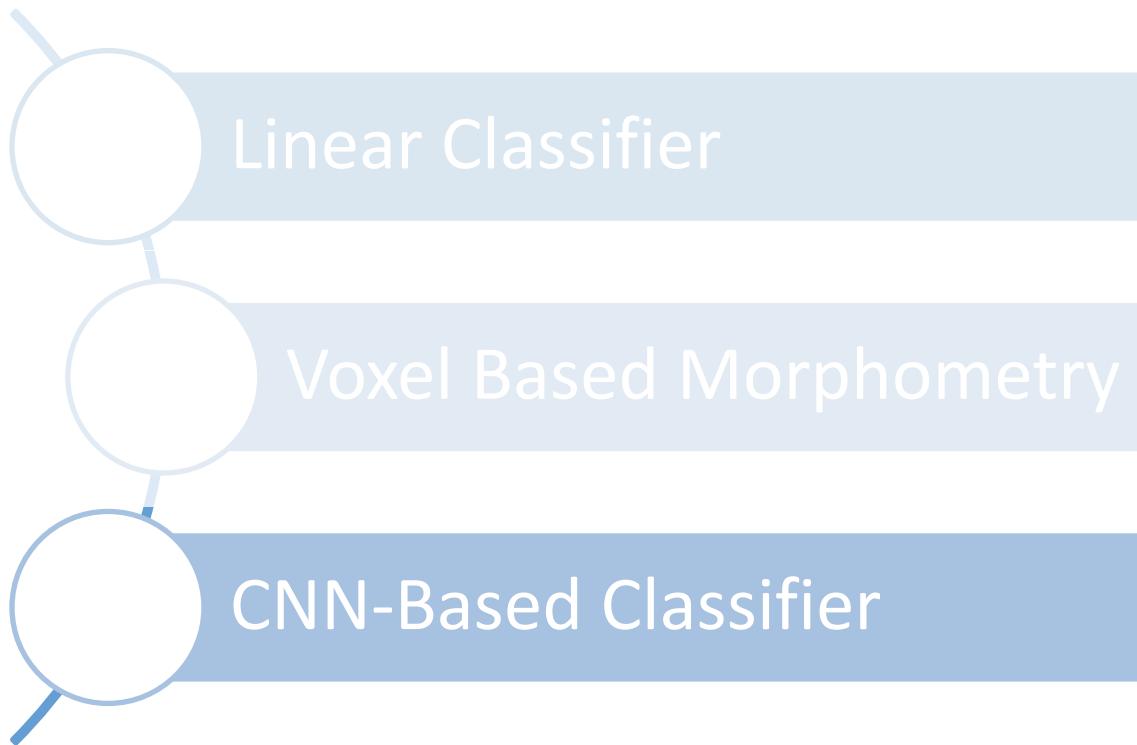
## Advantages

- Fully automated: quick and not susceptible to human error and inconsistencies
- Unbiased and objective
- Not based on regions of interests; more exploratory
- Picks up on differences/ changes at a global and local scale
- Has highlighted structural differences and changes between groups of people as well as over time

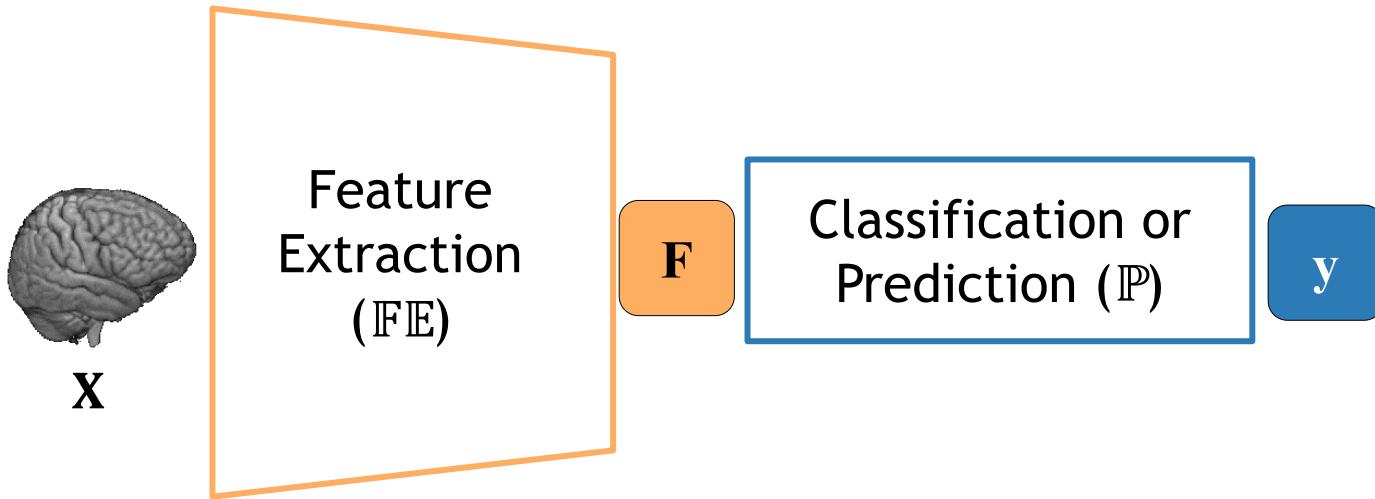
## Disadvantages

- Data collection constraints (exactly the same way)
- Statistical challenges
- Results may be flawed by preprocessing steps
- Underlying cause of difference unknown
- Interpretation of data- what are these changes when they are not volumetric?
- **Only group level findings**

# Today...



# End-to-end Training



Features  $F$  learned on-the-fly by feature extractor  $\text{FE}$  (e.g. Conv layers)

# Identifying HAND Brain Pattern

Zhang, Zhao et al., Medical Image Anlaysis, 2022

UCSF MRI Data Set (PI: Valcour)

Control  
N=156

HIV  
N=37

Cognitive Impaired  
HIV Negative  
N=335

HAND  
N=145

**How to apply CNN-based classifier to this data set ?**

Omitting samples to match cohorts is not an option

# Identifying HAND Brain Pattern

Zhang, Zhao et al., Medical Image Anlaysis, 2022

UCSF MRI Data Set (PI: Valcour)

Control  
N=156

HIV  
N=37

Cognitive Impaired  
HIV Negative  
N=335

HAND  
N=145

Insufficient Number of HIV Samples

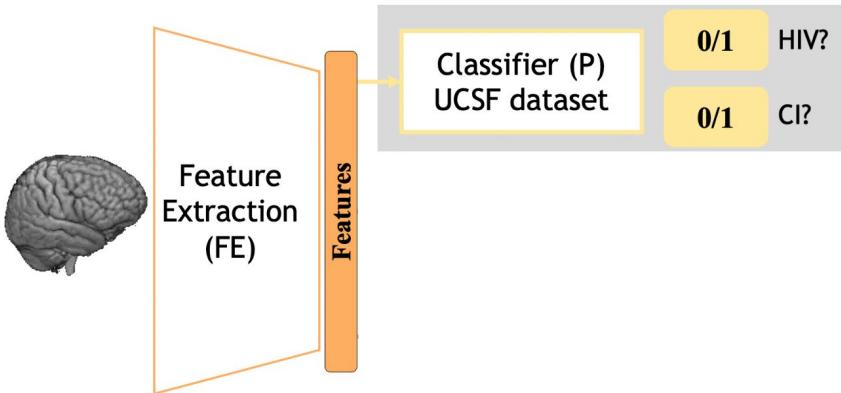
No publicly available data that contain all four cohorts

Add data set with only two cohorts

How to distinguish effect of HAND from study effect

# Combining Multiple Datasets

Zhang, Zhao et al., Medical Image Analysis, 2022



Dataset	Class	# of subjects	Age	Sex
UCSF	Ctrl	156	$70.1 \pm 5.9$	146 M / 10 F
	HIV	37	$64.9 \pm 3.8$	36 M / 1 F
	CI	335	$67.3 \pm 7.3$	165 M / 170 F
	HAND	145	$64.1 \pm 5.0$	136 M / 9 F

\* brain size is defined by the supratentorial volum

# Impact on Accuracy

<i>Multi-class (All)</i>	
Control	<b>55.7±8.0%</b>
CI-only	65.3±6.1% †
HIV-only	24.6±12.3% †*
HAND	49.6±8.6%
bAcc	48.8±3.6% †
Std	17.5%

\* Accuracy not significantly higher than chance (two-tailed  $p > 0.05$ , permutation test)

† Accuracy significantly lower than the three-domain model (two-tailed  $p < 0.05$ , Hardin-Shumway test).

# Impact on Accuracy

	Multi-class (All)	Multi-Label	
		One-Domain (UCSF)	
Control	<b>55.7±8.0%</b>	51.8±10.2%	
CI-only	65.3±6.1% †	73.4±4.8%	
HIV-only	24.6±12.3% †*	29.6±4.7% †*	
HAND	49.6±8.6%	42.8±9.3% †	
bAcc	48.8±3.6% †	49.4±1.7% †	
Std	17.5%	18.4%	

\* Accuracy not significantly higher than chance (two-tailed  $p > 0.05$ , permutation test)

† Accuracy significantly lower than the three-domain model (two-tailed  $p < 0.05$ , Hardin-Shumway test).

# Impact on Accuracy

	Multi-class (All)	Multi-Label	
		One-Domain (UCSF)	Single-Predictor (ALL)
Control	<b>55.7±8.0%</b>	51.8±10.2%	43.5±13.4% †
CI-only	65.3±6.1% †	73.4±4.8%	71.6±5.3%
HIV-only	24.6±12.3% †*	29.6±4.7% †*	43.6±16.4%
HAND	49.6±8.6%	42.8±9.3% †	49.6±5.8%
bAcc	48.8±3.6% †	49.4±1.7% †	52.1%±3.7
Std	17.5%	18.4%	13.4%

\* Accuracy not significantly higher than chance (two-tailed  $p > 0.05$ , permutation test)

† Accuracy significantly lower than the three-domain model (two-tailed  $p < 0.05$ , Hardin-Shumway test).

# Impact on Accuracy

	Multi-class (All)	Multi-Label			Ours
		One-Domain (UCSF)	Single-Predictor (ALL)	Two-Domain (UCSF+SRI)	
Control	<b>55.7±8.0%</b>	51.8±10.2%	43.5±13.4% †	51.9±9.1%	
CI-only	65.3±6.1% †	73.4±4.8%	71.6±5.3%	<b>76.1±6.7%</b>	
HIV-only	24.6±12.3% †*	29.6±4.7% †*	43.6±16.4%	32.9±13.7% *	
HAND	49.6±8.6%	42.8±9.3% †	49.6±5.8%	49.6±12.6%	
bAcc	48.8±3.6% †	49.4±1.7% †	52.1%±3.7	52.6%±7.6	
Std	17.5%	18.4%	13.4%	18.0%	

\* Accuracy not significantly higher than chance (two-tailed  $p > 0.05$ , permutation test)

† Accuracy significantly lower than the three-domain model (two-tailed  $p < 0.05$ , Hardin-Shumway test).

# Impact on Accuracy

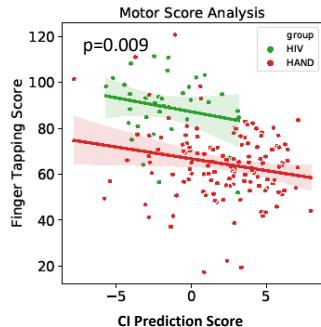
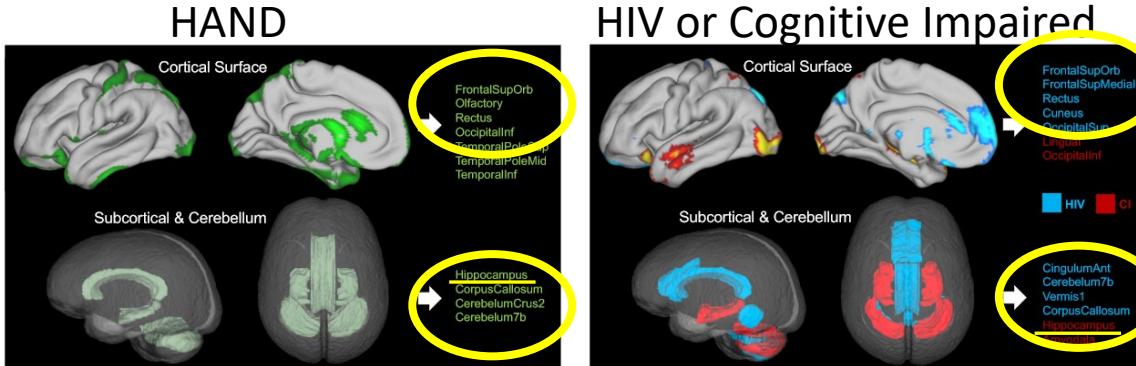
	Multi-class (All)	Multi-Label			
				Ours	
		One-Domain (UCSF)	Single-Predictor (ALL)	Two-Domain (UCSF+SRI)	Three-Domain (All)
Control	<b>55.7±8.0%</b>	51.8±10.2%	43.5±13.4% †	51.9±9.1%	51.8±7.5%
CI-only	65.3±6.1% †	73.4±4.8%	71.6±5.3%	<b>76.1±6.7%</b>	74.0±3.4%
HIV-only	24.6±12.3% †*	29.6±4.7% †*	43.6±16.4%	32.9±13.7% *	43.9±13.6%
HAND	49.6±8.6%	42.8±9.3% †	49.6±5.8%	49.6±12.6%	51.0±13.6%
bAcc	48.8±3.6% †	49.4±1.7% †	52.1%±3.7	52.6%±7.6	55.2±4.7%
Std	17.5%	18.4%	13.4%	18.0%	13.2%

\* Accuracy not significantly higher than chance (two-tailed  $p > 0.05$ , permutation test)

† Accuracy significantly lower than the three-domain model (two-tailed  $p < 0.05$ , Hardin-Shumway test).

# HIV-associated Neurocognitive Disorder

Zhang, Zhao et al., Medical Image Analysis, 2022



# Announcements

**Thursday: Proposal and Midterm**

# Thank you!

- <https://ml4n.Stanford.edu/>
- [psyc221-aut2324-staff@staff.stanford.edu](mailto:psyc221-aut2324-staff@staff.stanford.edu)