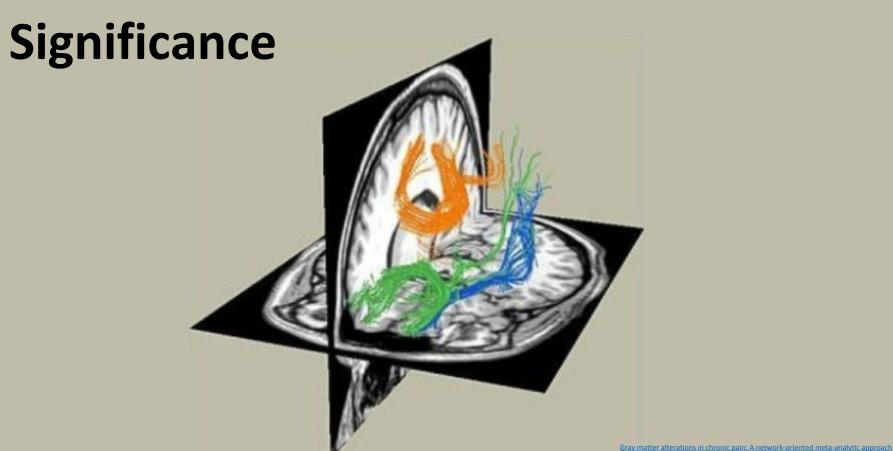
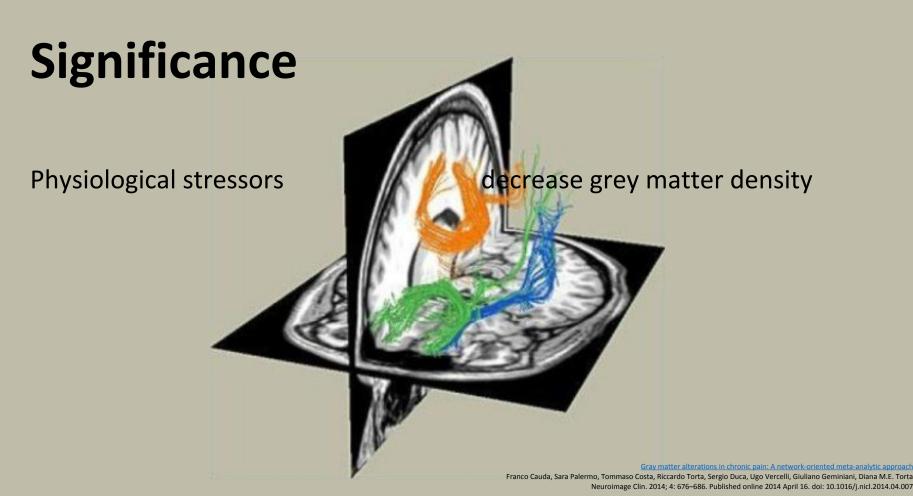
Brain Synapse Density

Jay Miller
Akash Ray
Emily Marshall
Bijan Varjavand

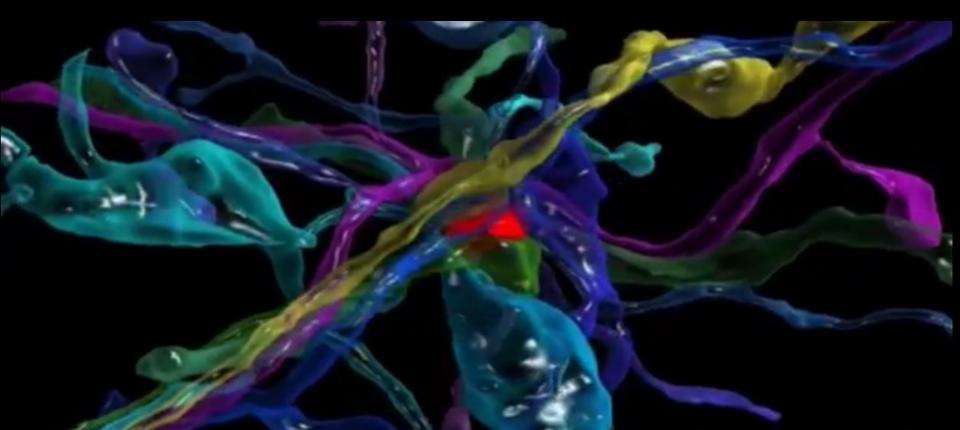
Significance



Franco Cauda, Sara Palermo, Tommaso Costa, Riccardo Torta, Sergio Duca, Ugo Vercelli, Giuliano Geminiani, Diana M.E. Torta Neuroimage Clin. 2014; 4: 676–686. Published online 2014 April 16. doi: 10.1016/j.nicl.2014.04.007



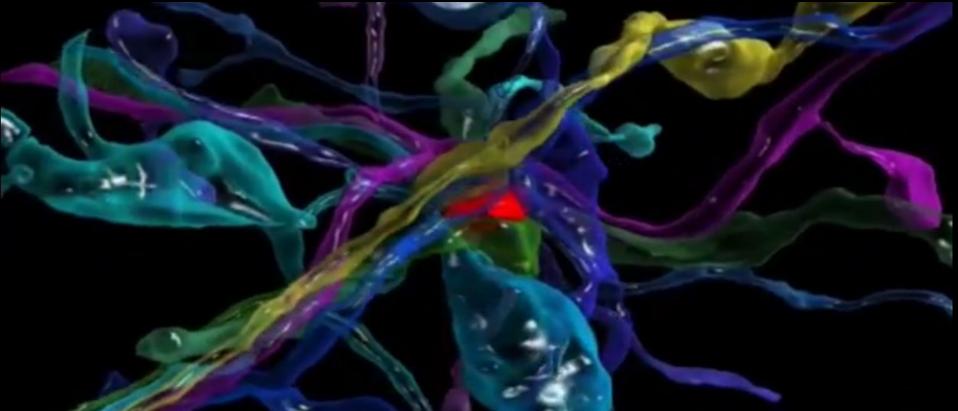
Synapse Density

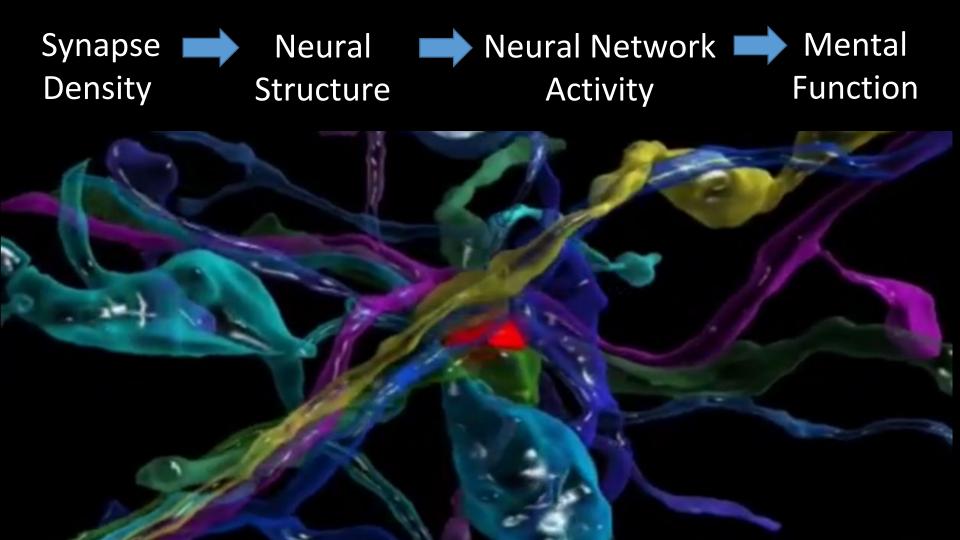




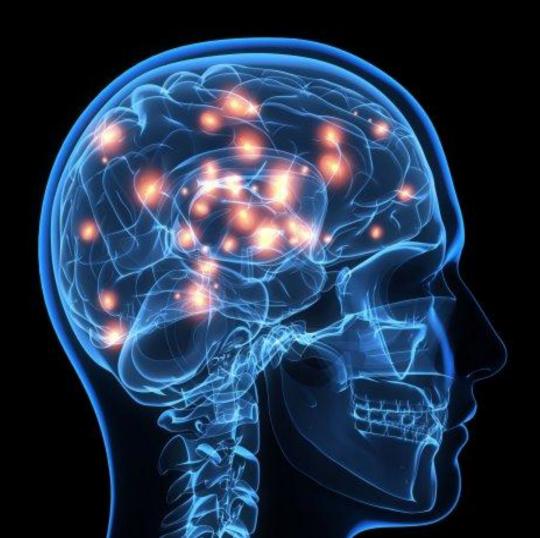








Gap



Gap

Incomplete understanding of neural networks



Gap

Incomplete understanding of neural networks

Incomplete modeling of distributed synapses



Challenge

Large data set

Lack of clarity & labels

Only do regression

Difficult to interpret & understand

Formal Statement of Problem

Independent variables: $X \in \mathbb{R}^4 \text{ s.t. } x_i = (c_x, c_y, c_z, s)$

Dependent variables: $Y \in \mathbb{R}$

Unknown parameters: $\theta \in \Theta$

Regression: find a conditional distribution for Y|X, parameterized by θ i.e. $y_i = f(x_i; \theta)$

For estimated $\hat{\theta}$, loss is function of y_i, \hat{y}_i s.t $\hat{y}_i = f(x_i; \hat{\theta})$

In other words,

 $\hat{\theta} = \operatorname{argmin}_{\theta} l(\hat{y}_i, \hat{y}_i)$

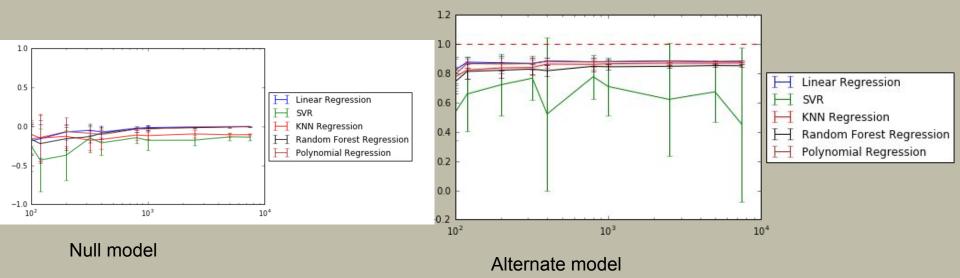
Model assumptions

- (X, Y) are iid
- Y is conditionally dependent on X

Algorithms

- Linear Regression
- -Support Vector Regression (SVR)
- -K-Nearest Neighbor Regression (KNN)
- -Random Forest Regression (RF)
- -Polynomial Regression (degree = 2)

Simulated regressions (10 fold cross validation)



Regressions on actual data (10 fold cross validation)

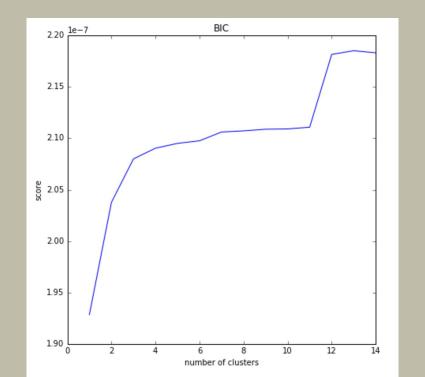
```
R^2 of Linear Regression: 0.62 (+/- 0.40)
R^2 of SVR: 0.57 (+/- 0.18)
R^2 of KNN Regression: 0.25 (+/- 2.54)
R^2 of Random Forest Regression: 0.79 (+/- 0.51)
R^2 of Polynomial Regression: 0.85 (+/- 0.27)
```

Model checking

Independence? Yes

le9 Covariance of Synapse Density dataset 2.85 0 2.70 20 2.55 40 2.40 60 2.25 2.10 80 1.95 20 40 60 80 1.80

Identical distributions? No



What's next?

- Interpret regression results

```
R^2 of Linear Regression: 0.62 (+/- 0.40)
R^2 of SVR: 0.57 (+/- 0.18)
R^2 of KNN Regression: 0.25 (+/- 2.54)
R^2 of Random Forest Regression: 0.79 (+/- 0.51)
R^2 of Polynomial Regression: 0.85 (+/- 0.27)
```

- More regressions? Is the success of our regressions mainly due to strong dependence between synapses and unmasked?
 - I.e. what's the marginal distribution for synapses?