

# Connectomic Heritability: a Causal Modeling Analysis

**Jaewon Chung**

(he/him) - NeuroData lab

*Johns Hopkins University - Biomedical Engineering*

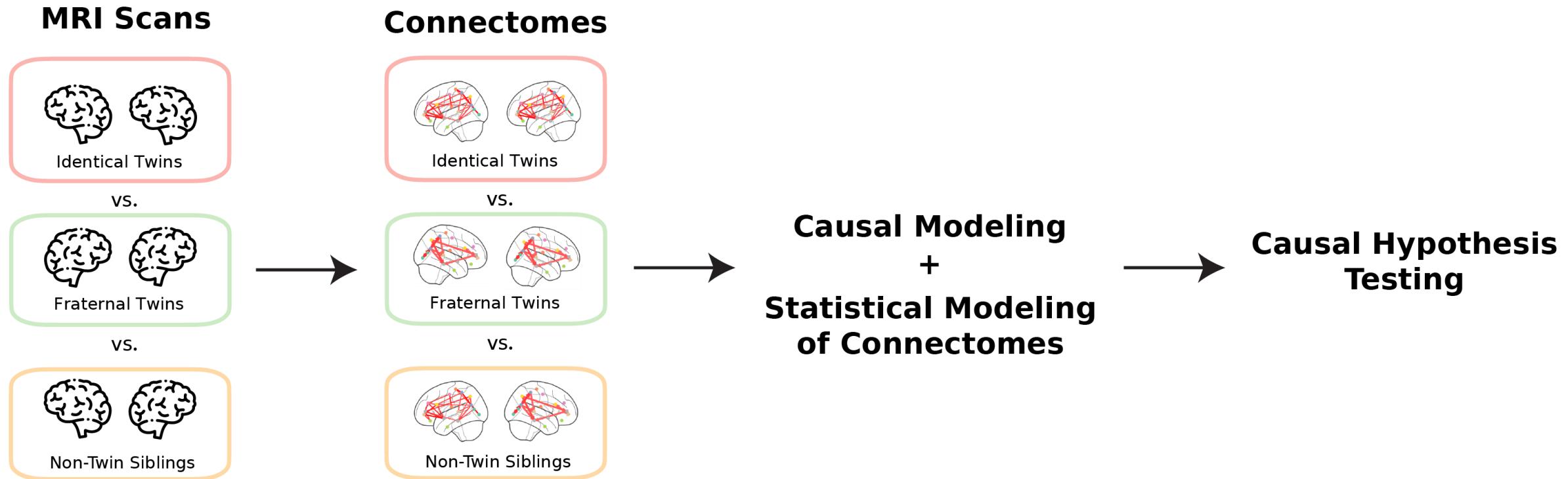
✉ [j1c@jhu.edu](mailto:j1c@jhu.edu)

/github/ [@j1c \(Github\)](#)

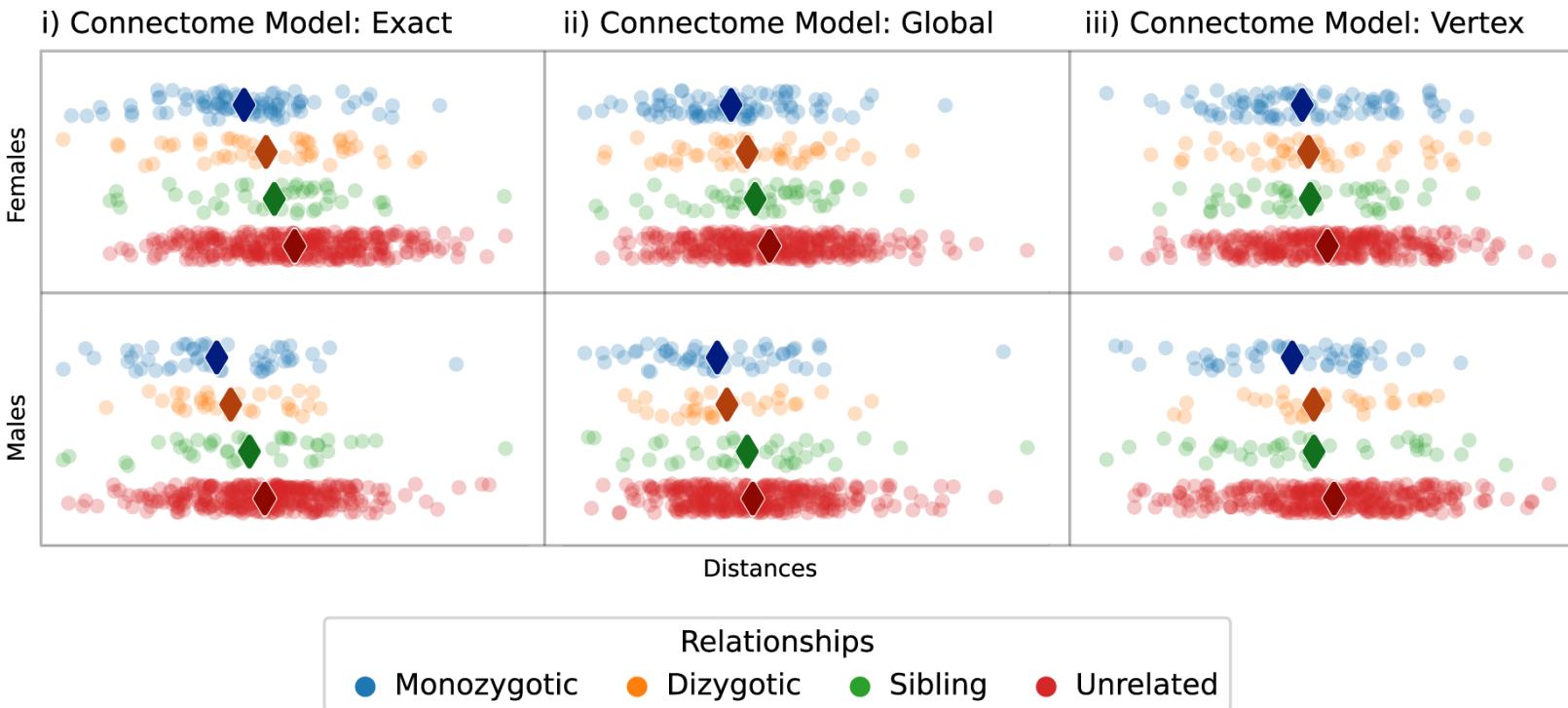
/twitter/ [@j1c \(Twitter\)](#)



# How do we study connectomic heritability?

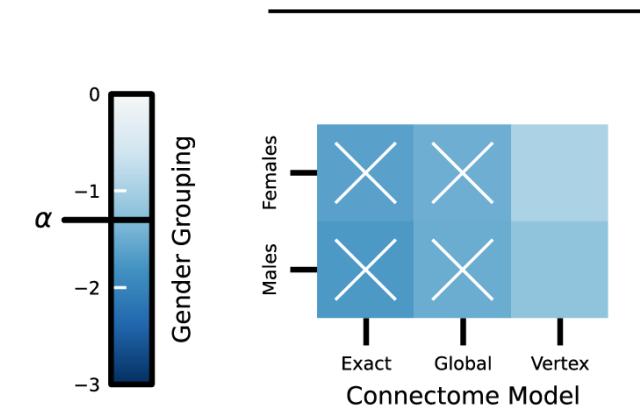


### A) Pairwise Distances Separated by Familial Relationships



### B) P-values From Tests for Heritability

#### Causal Hypothesis Test



# Acknowledgements

JHU



Eric  
Bridgeford



Mike  
Powell



Derek  
Pisner



Carey  
Priebe



Joshua T.  
Vogelstein

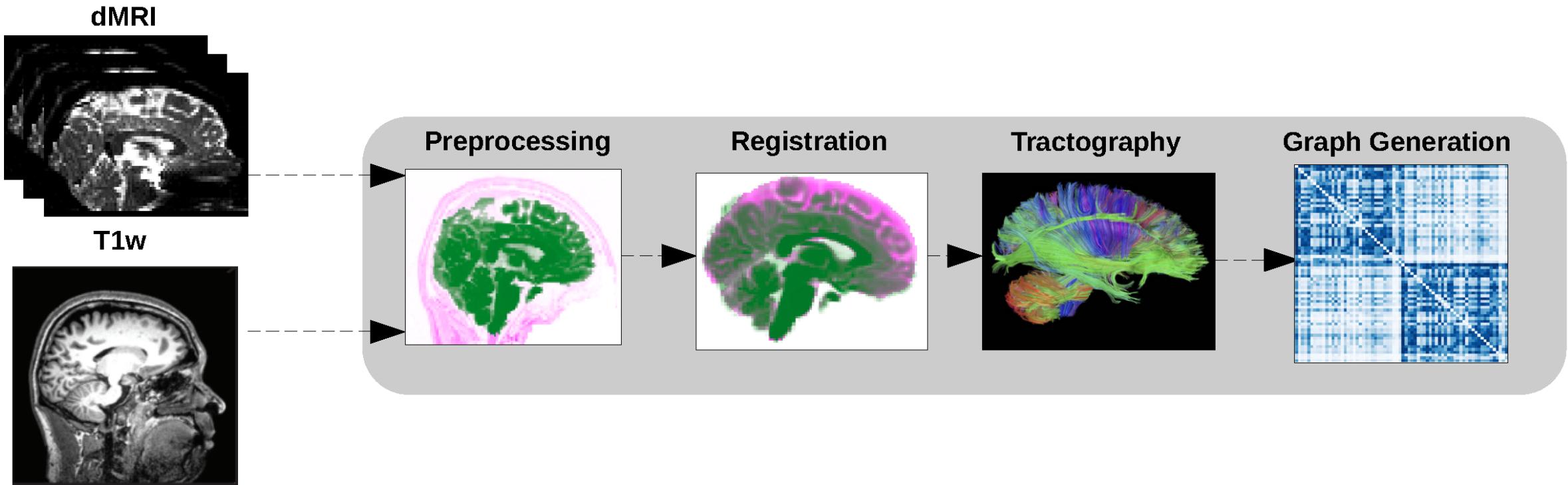
**Microsoft Research**

Dax Pryce, Bryan Tower, Nick Caurvina, Patrick Bourke,  
Jonathan McLean, Carolyn Buractaon, Amber Hoak.

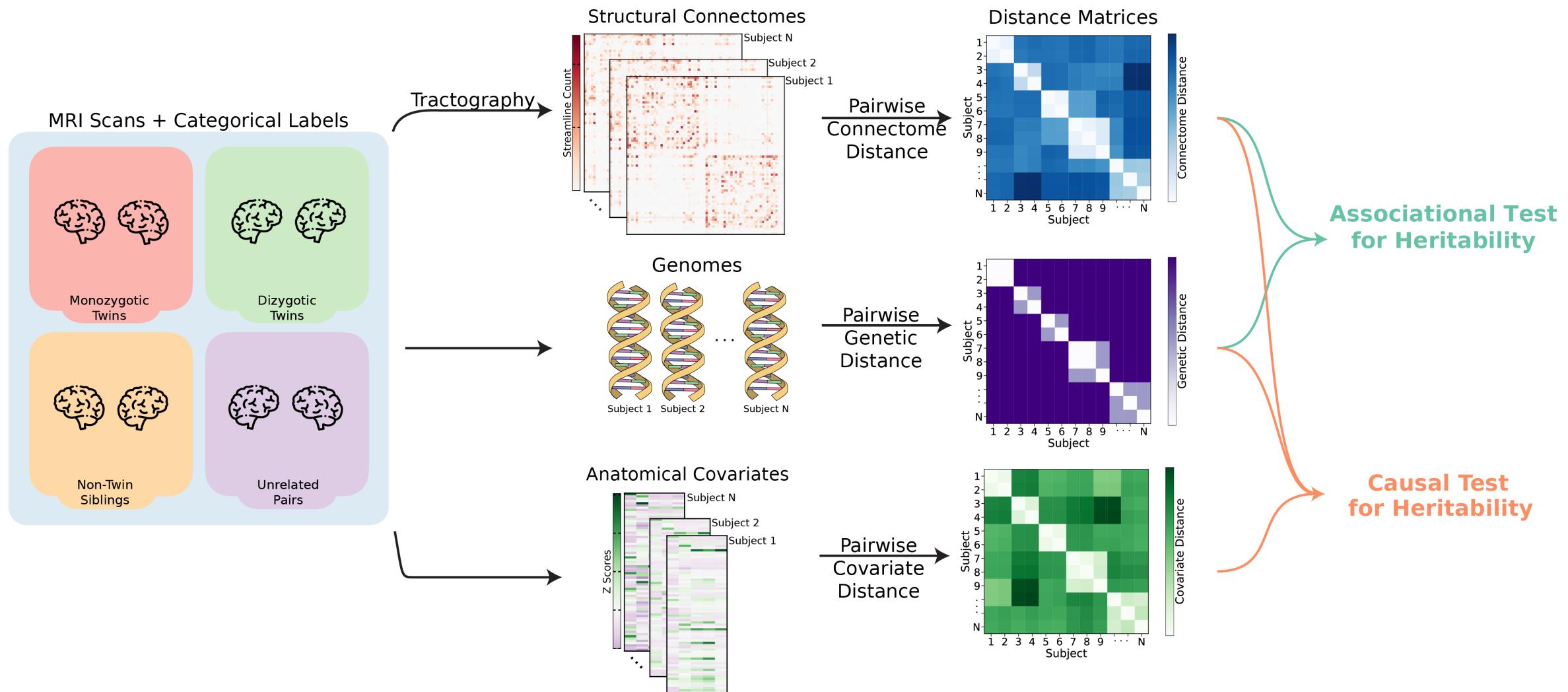
**Questions?**

# **Additional Slides**

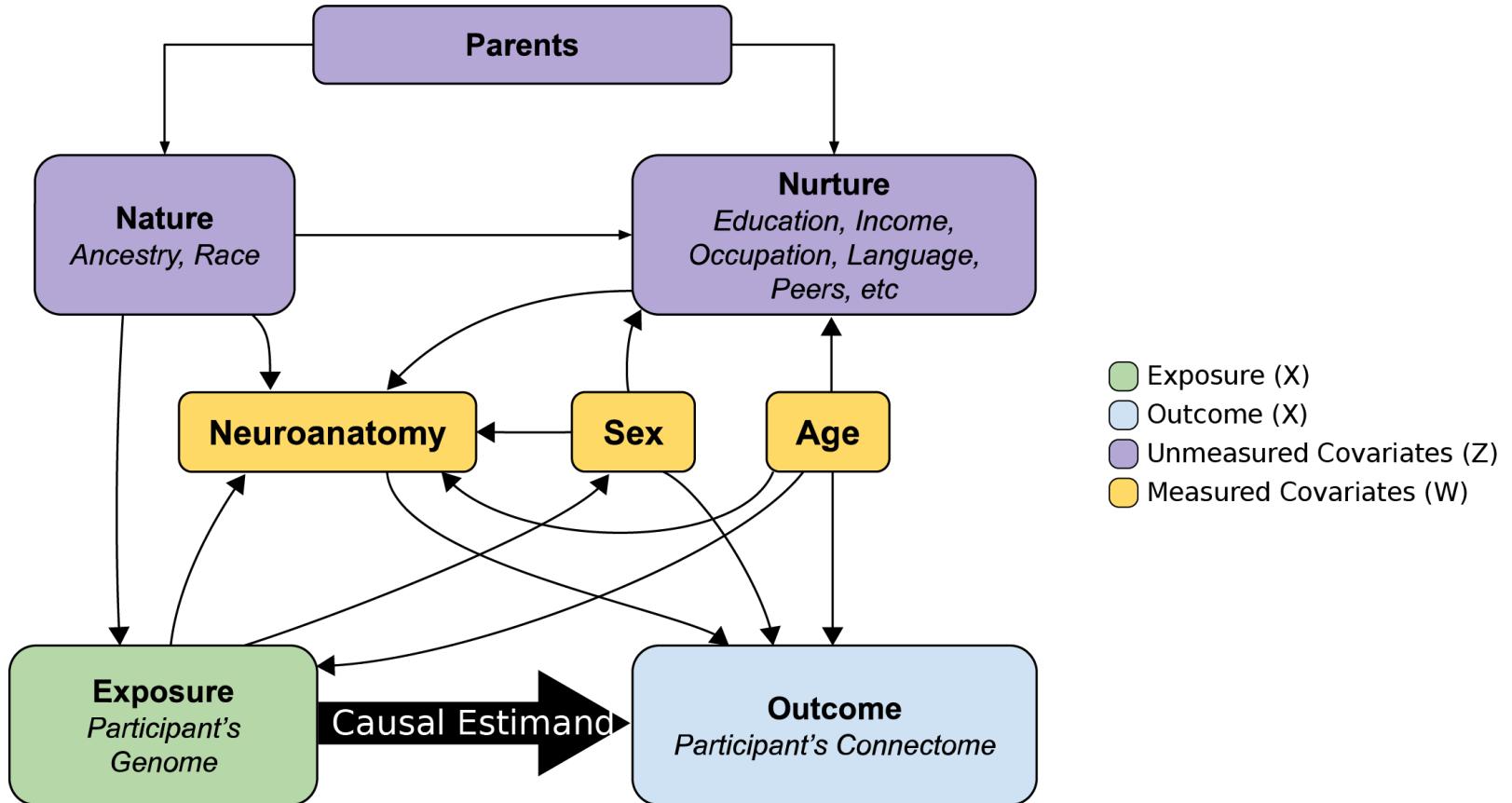
# MRI to Connectomes



# Analysis Framework



# Causal Directed Acyclic Graph



# Causal Estimand

- $X$  denote exposure,  $Y$  denote outcome,  $W$  denote measured covariates,  $Z$  denote unmeasured covariates.
- Want to estimate the effect of different exposures on the outcome, which is quantified using the backdoor formula if  $W$  and  $Z$  close all backdoor paths.

$$f_{w,z}(y|x) = \int_{\mathcal{W} \times \mathcal{Z}} f(y|x, w, z) f(w, z) d(w, z)$$

- Above integrates over *all* measured and unmeasured covariates.

$$f(y|x) = \int_{-\mathcal{W} \times \mathcal{Z}} f(y|x, w, z) f(w, z|x) d(w, z)$$

- Averages the true outcome distribution over the *conditional* distribution of the measured and unmeasured covariates.

# Hypothesis Testing

## Associational Heritability

- We observe the pairs  $(x_i, y_i)$  for  $i \in [n]$ .
- Only be able to estimate the functions of  $(X, Y)$
- The corresponding hypothesis test is:

$$H_0 : f(y|x) = f(y) \quad \text{vs} \quad H_A : f(y|x) \neq f(y).$$

## Causal Heritability

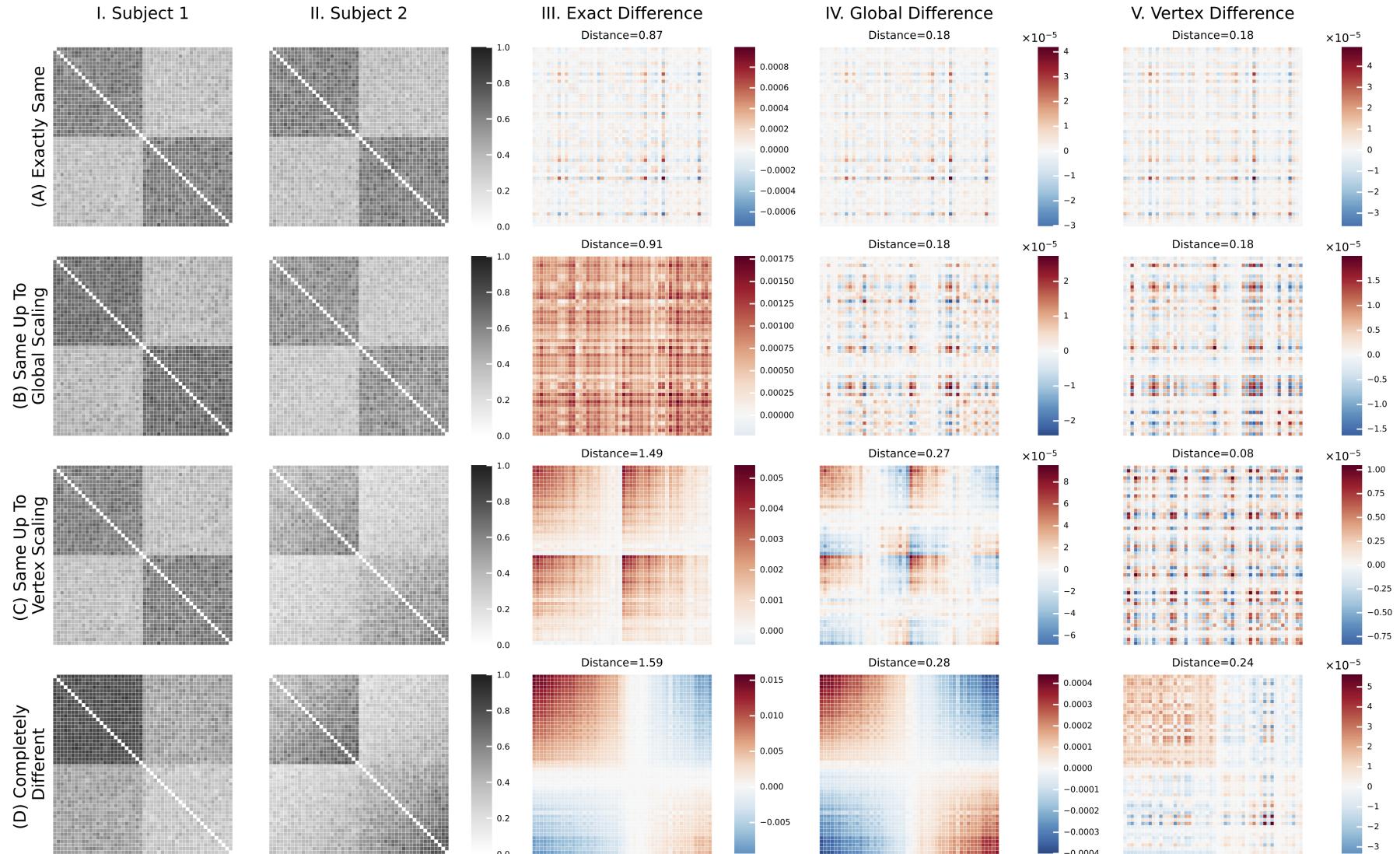
- We observe the triples  $(x_i, y_i, w_i)$  for  $i \in [n]$ .
- Only be able to estimate the functions of  $(X, Y, W)$
- The corresponding hypothesis test is:

$$H_0 : f(y|x, w) = f(y|w) \quad \text{vs} \quad H_A : f(y|x, w) \neq f(y|w).$$

# Connectome Models

- Exact model:** This model measures all differences between latent positions, with differences in the latent positions implying differences in the connectomes themselves.
- Global model:** This model examines whether the latent positions of one connectome are a scaled version of the other. For example, if the number of edges in male connectomes are consistently larger than those in females, we have no way of differentiating whether significant findings from the exact model are a result of differences in scaling or differences in the fundamental structure of the connectomes themselves.
- Vertex model:** This model is similar to the global model, but it allows for each vertex to be scaled differently. The idea behind this approach is that some vertices may have a greater impact on the overall network than others, so scaling them differently can provide a more accurate representation of the network. Consider the examples of how brain regions connect with each other. Regions in the same hemisphere are more likely to be connected than across hemispheres. Even within the same hemisphere, different regions may have distinct preferences for forming connections with other specific regions.

# Connectome Model Simulations



# Shortcomings - Model assumptions

- No interaction between genome and environment
- No epistasis
  - Effect of one gene is dependent on another
  - Ex: black hair and baldness
- No dominance effects
- Strong assumptions in genetic distances