

# Evaluating Accuracy and Reliability of Brain-Behavior Models Using Diffusion MRI

McKenzie Paige Hagen<sup>1, </sup>, John Kruper<sup>1</sup>, Keshav Motwani<sup>2</sup>, Eardi Lila<sup>2</sup>, Jason Yeatman<sup>3</sup>, Ariel Rokem<sup>1, 4</sup>



<sup>1</sup> Department of Psychology, University of Washington <sup>2</sup> Department of Biostatistics, University of Washington <sup>3</sup> Graduate School of Education, Stanford University <sup>4</sup> University of Washington eScience Institute

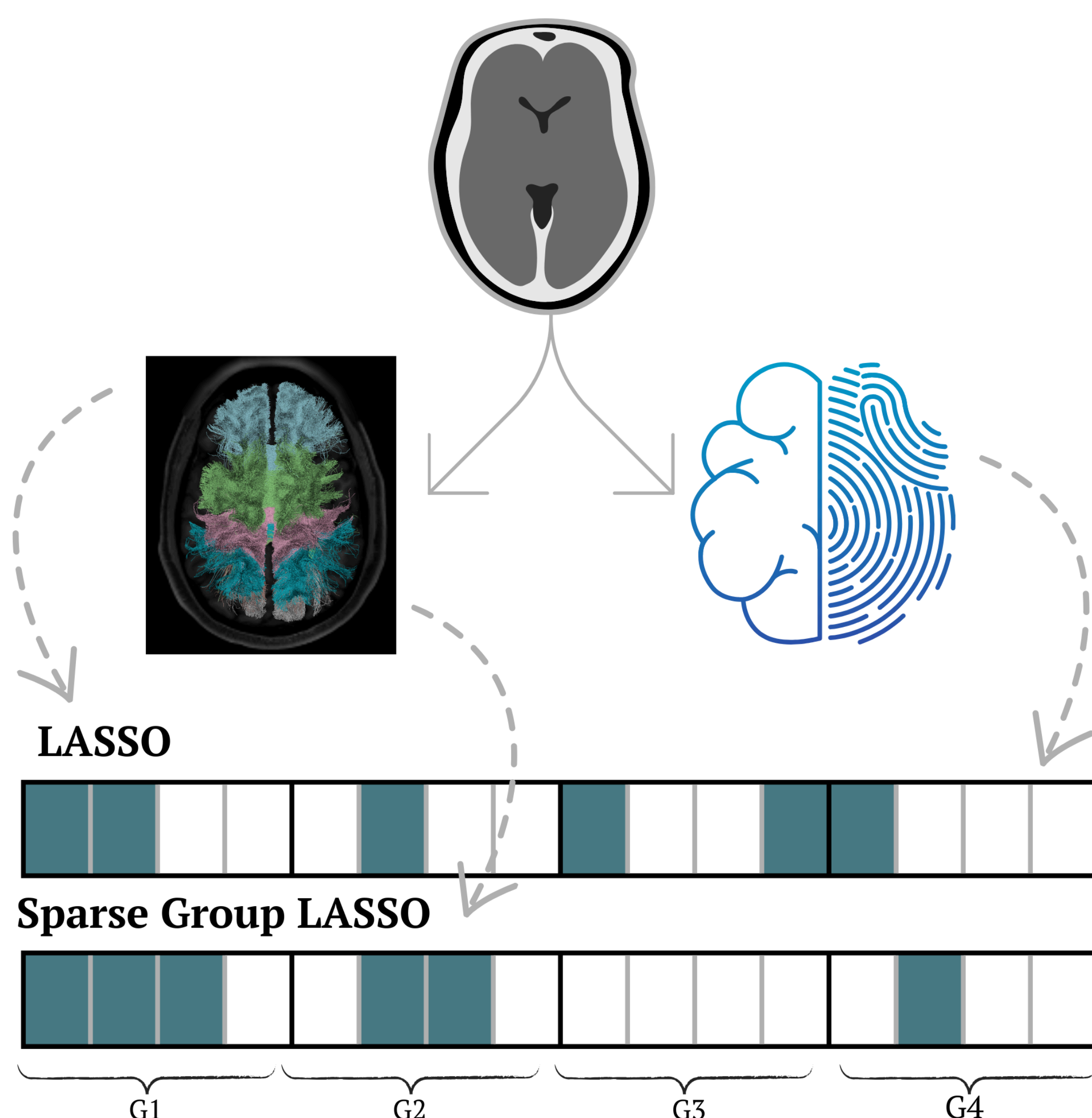
## Background

- Diffusion MRI (dMRI) measures tissue properties of white matter, which contains long-range connections between different brain regions.
- Brain-behavior models can be used link neuroimaging features and phenotypes.
- Researchers have several options to process dMRI data into input features for brain-behavior modeling.
- Characteristics of feature sets determine which regularization methods are suitable for modeling fitting.

**Question:** How do sets of features derived from different dMRI processing methods compare in model accuracy and variability?

## Methods

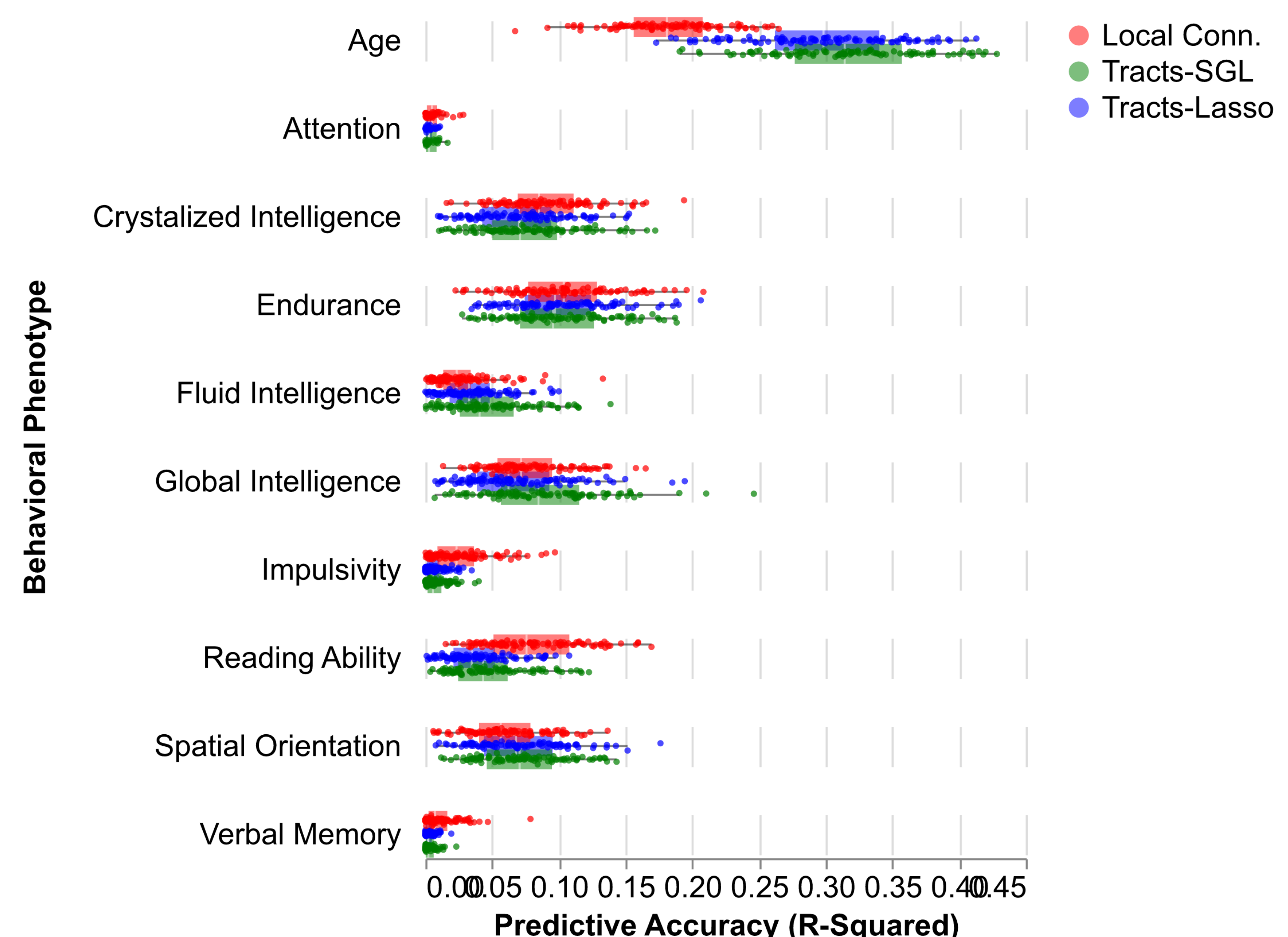
- Diffusion MRI from 1,041 Human Connectome Project participants.
- Processed into “tract profiles”(1) (9,600 features) using [pyAFQ](#) and “local connectome fingerprint” (128,894 features) using DSI-Studio (2) (shared by (3)).
- LASSO regularized models fit to tract profiles and local connectomes separately.
- Sparse Group LASSO (SGL) models fit to tract profiles, grouped by tract.
- Prediction targets were various cognitive phenotypes drawn from prior predictive modeling investigations(3).
- Models implemented in [R](#) and trained using nested group cross-validation and bootstrap resampling by family.
- Additional models were ran without group cross-validation.



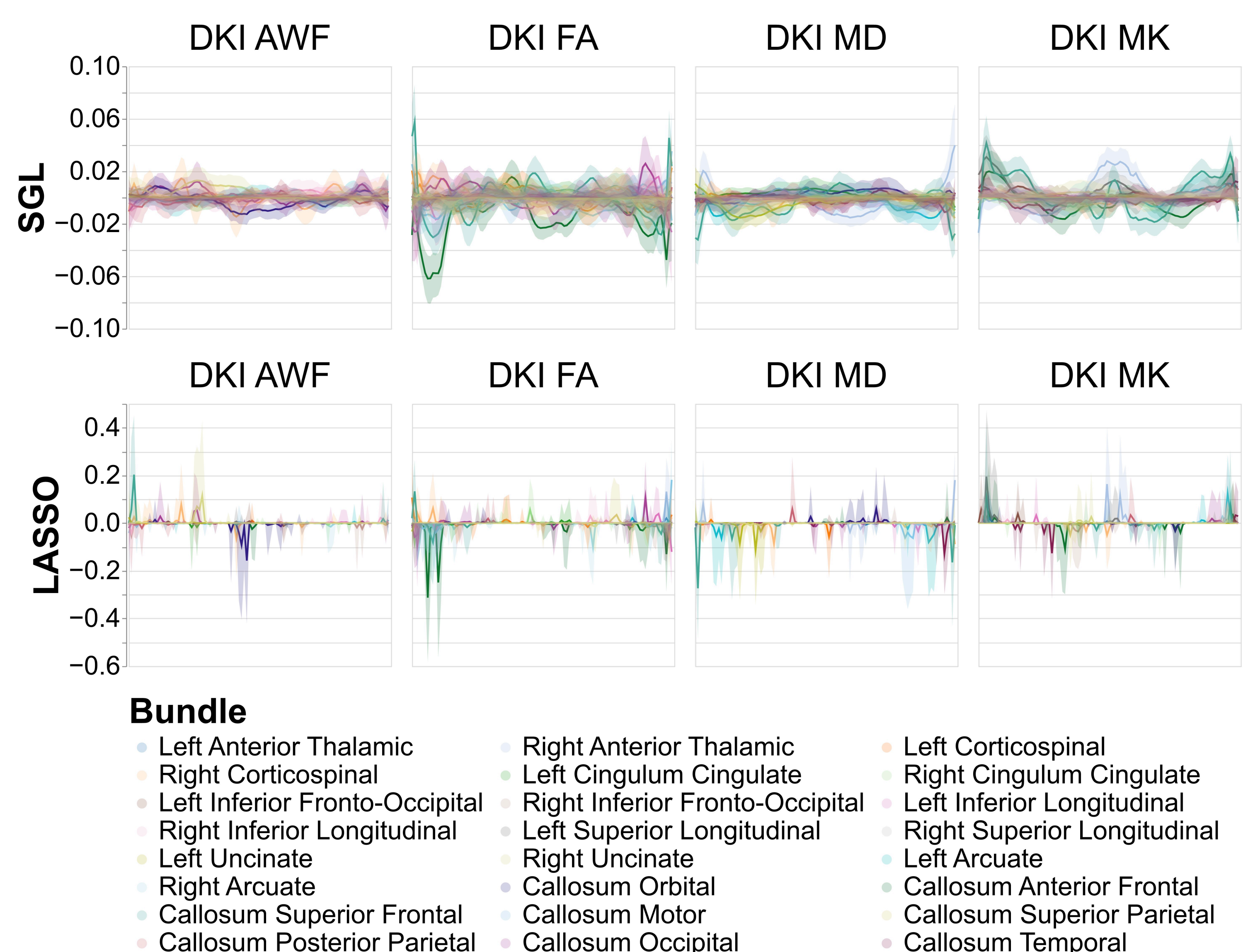
## Conclusions

- Phenotypes vary in their ability to be predicted from features derived from imaging data.
- Tractometry contains equivalent predictive information as compared to local connectome, with much fewer features.
- Utilizing a variety of regularization methods and feature engineering can maximize interpretability of models through more stable model weights.
- Despite concerns about “leakage”, there was no significant difference between model accuracies with and without grouping by family for cross-validation splits for all phenotypes.
- There’s more space here!

## Results



**Model accuracies for behavioral phenotypes.** All models, regardless of regularization method or feature set performed statistically equivalently.



**“Age” prediction model weights across tracts for SGL and LASSO.** Solid lines show the mean model weight across bootstraps for every tract, across every node, and the shaded areas show the 95% confidence intervals of the model weights. Note the reduced y-axis range for SGL.

## Acknowledgements



Data were provided by the Human Connectome Project, WU-Minn Consortium

## References

1. J. Kruper *et al.*, *Apert Neuro* (2021).
2. F.-C. Yeh *et al.*, *PLoS Comput. Biol.* (2016).
3. J. Rasero, A. I. Sentis, F.-C. Yeh, T. Verstynen, *PLoS Comput. Biol.* (2021).