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

McKenzie Paige Hagen[1], John Kruper[1], Keshav Motwani[2], Eardi Lila[2], Jason Yeatman[3], Ariel Rokem[1, 4]

[1] Department of Psychology, University of Washington [2] Department of Biostatistics, University of Washington [3] Graduate School of Education, Stanford University [4] University of Washington eScience Institute

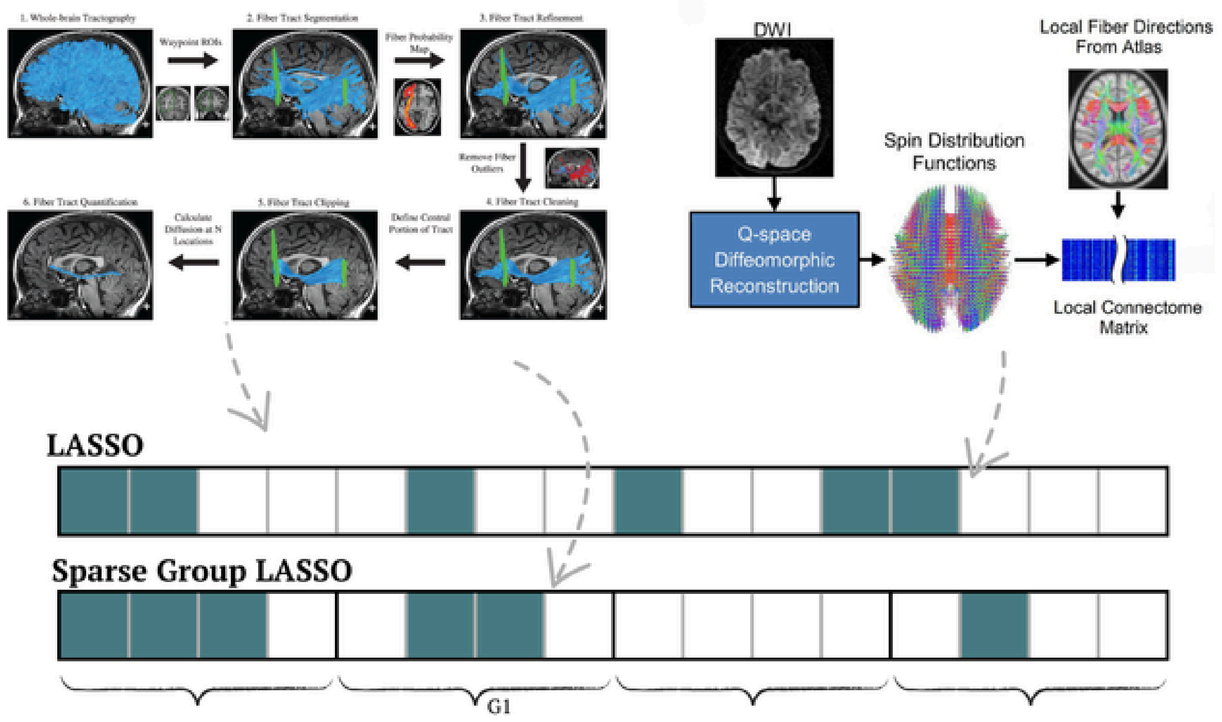
Background

- 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

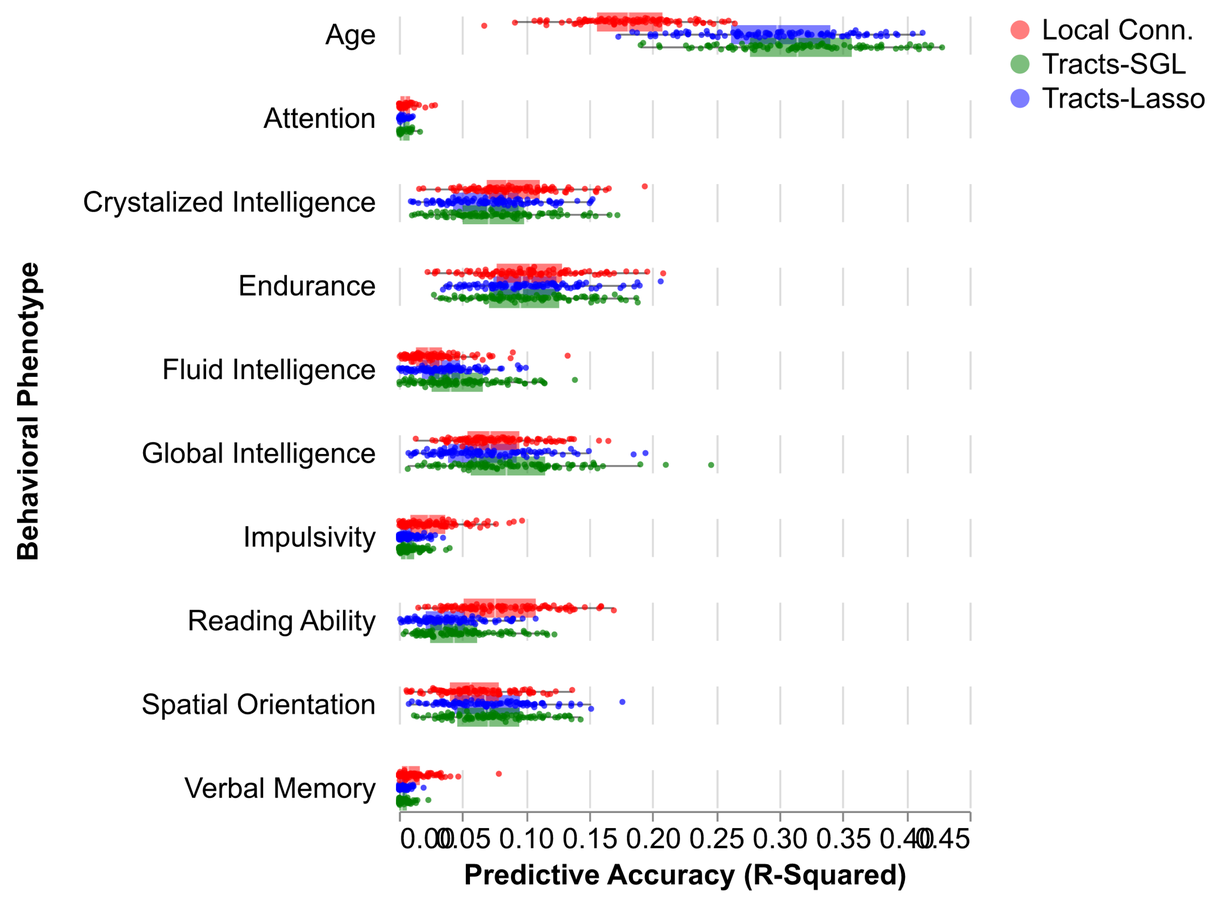
- Processed into "tract profiles" (see (*1*) for processing details and to access data; 9,600 features) using pyAFQ and "local connectome fingerprint" (128,894 features) using DSI-Studio (*2*) (shared by (*3*)).
- LASSO regularizated models fit to tract profiles and local connectomes seperately.
- Sparse 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 boostrap resampling by family.
- Models evaluated with and without assigning related individuals to the same fold to evaluate "leakage".



Conclusions

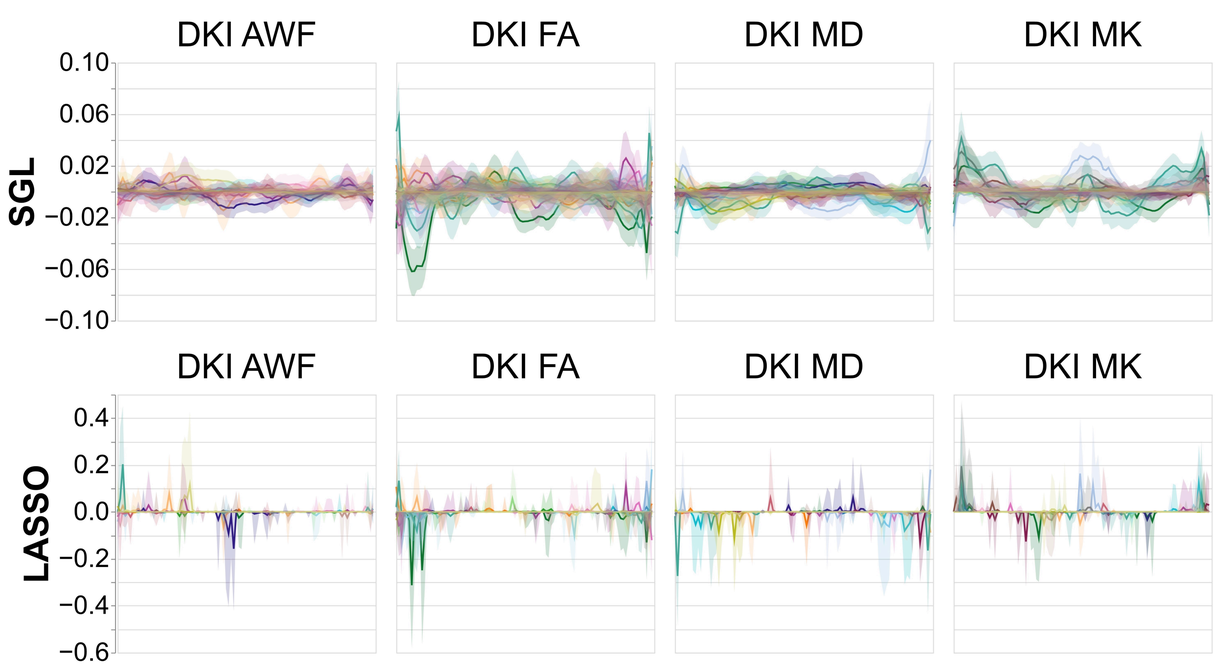
- Tractometry contains equivalent predictive information as compared to local connectome, with much fewer features.
- Utilizing a variety of regularization methods and feature engineering can maximize interpretability of models through more stable model weights.
- There's more space here!
-

Results



Model accuracies for behavioral phenotypes. All models, regardless of regularization method or feature set performed equivalently. These R^2 values are in line with previous literature evaluating brain-behavior predictive models.(*3*)

- Despite concerns about "leakage", there was no difference between model accuracies with and without grouping by family for cross-validation splits for all phenotypes.



"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. The shaded areas show the 95% confidence intervals of the model weights. Note the reduced y-axis range for SGL. This pattern was consistent across phenotypes.

Acknowledgements

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

References

1. J. Kruper *et al.*, *Frontiers in Neuroscience* (2024).
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).