

# 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</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

## Background

- Diffusion MRI measures tissue properties of white matter, which contains long-range connections between different brain regions.
- Raw diffusion data can be engineered into different “feature sets” for use in predictive models.
- Comparison of a variety of feature sets and models for both accuracy of predicutions and variance/bias of beta coefficients can help determine which models are more optimal for different feature sets.

## Methods

- We used diffusion data from 1041 participants from the Human Connectome Project, processed into “tract profiles” using pyAFQ (cite) and “local connectome” features using DSI-Studio (cite).
- LASSO models were run on both tract profiles and local connectome to predict a variety of cognitive outcomes.
- Sparse Group LASSO (SGL) models were run on only tract profiles, to take advantage of the inherent tract groupings.
- Models were implemented in  and trained using nested cross-validation and bootstrap resampling (link to code).

- Explanation of / diagram of SGL here.

- spacing

- more spacing

- For spacing

## Main Results

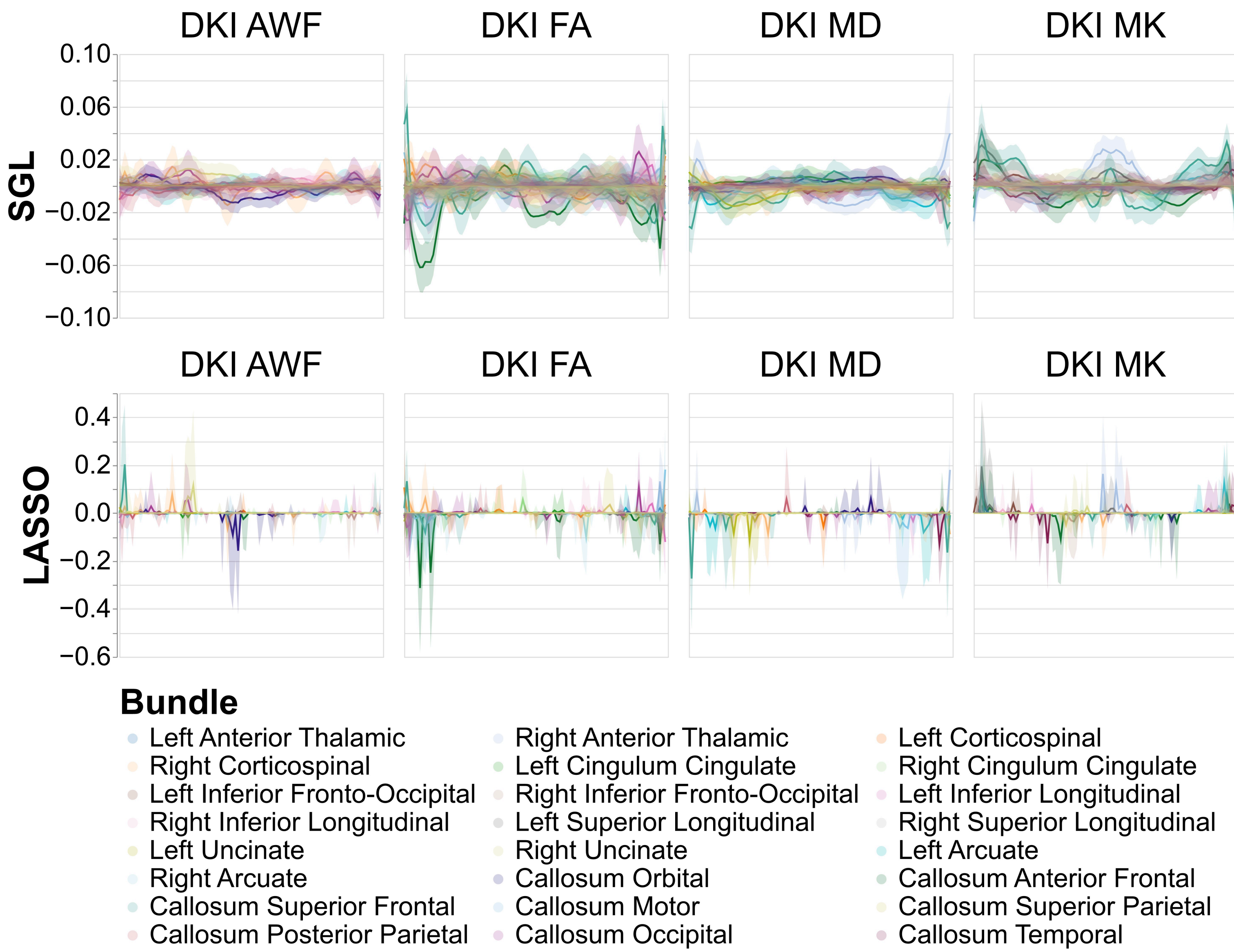
- There were no significant differences in the accuracy of models using tract profiles and local connectomes as their input features (Table 1.)
- There were significant differences in the variability of model weights for LASSO and SGL models (Figure 2.)

	LASSO	LC	SGL
Age	3e-01	1.8e-01	3.1e-01
Crystalized Intelligence	7.2e-02	9e-02	7.6e-02
Fluid Intelligence	3.5e-02	2.7e-02	4.8e-02
Global Intelligence	6.8e-02	7.7e-02	8.9e-02
Impulsivity	7.1e-03	2.7e-02	8.2e-03
Endurance	1e-01	1e-01	1e-01
Verbal Memory	3.3e-03	1.2e-02	3.7e-03
Reading Ability	3.9e-02	8.1e-02	4.8e-02
Attention	3.3e-03	6.4e-03	4.4e-03
Spatial Orientation	7e-02	6.1e-02	7.2e-02

Table 1. This can be replaced by a condensed version of accuracy fig .

## Additional Results

- The inclusion of bootstrapping in our analytic pipeline appeared to decrease the accuracy of the models.\*\* maybe
- Avoiding leakage by not spltting realted subjects across the train/test splits did not have a significant effect.
- Relevant figure here:



Caption for the picture.

## Conclusions (rough wording)

- The selection of model and feature set might not be influential on the accuracy, but may result in less variable, more interpretable models.
- Tract profiles and local connectome have similar accuracies, but the grouping of tract profiles combined with SGL is good\*
- Splitting families across the train/test splits is bad practice, but didn’t have a large effect on the outcome.

## References

Created with (Allaire et al. 2024)

Allaire, JJ, Yihui Xie, Christophe Dervieux, Jonathan McPherson, Javier Luraschi, Kevin Ushey, Aron Atkins, et al. 2024. *Rmarkdown: Dynamic Documents for r*. <https://github.com/rstudio/rmarkdown>.