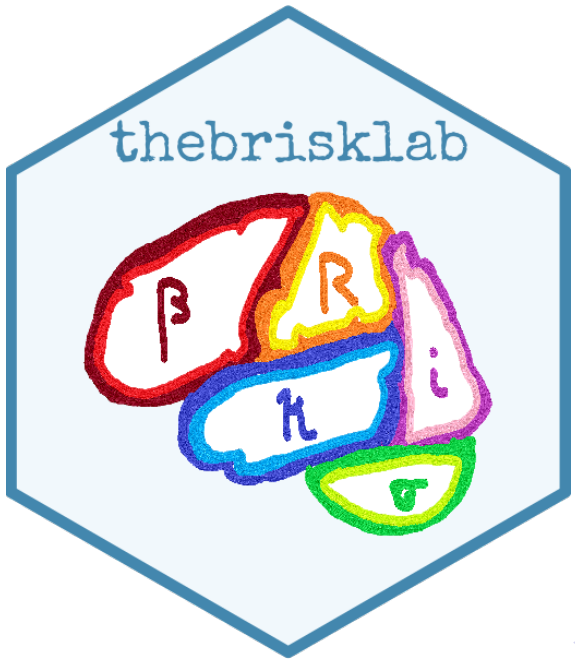
A Missing Data Method for Deconfounding in Neuroimaging Studies

Benjamin Risk

benjamin.risk@emory.edu



Joint work

Preprint:

Nebel, M.B.[1,2], Lidstone, D.E.[1,2], Wang, L.[4], Benkeser, D.[4], Mostofsky, S.H.[1,2,3], and Risk, B.B.[4] (2022). Accounting for motion in fMRI: What part of the spectrum are we characterizing in autism spectrum disorder? *bioRxiv.*

[1] Center for Neurodevelopmental and Imaging Research, Kennedy Krieger Institute
[2] Department of Neurology, Johns Hopkins University School of Medicine
[3] Department of Psychiatry and Behavioral Science, Johns Hopkins University School of Medicine
[4] Department of Biostatistics & Bioinformatics, Rollins School of Public Health, Emory University

Motion quality control causes massive data loss

- Motion in the scanner produces artifacts (Power et al., 2012).
- Lenient criteria: < 5 min data after removing frames with > 3 mm or $3°$ from previous frame (Fassbender et al., 2017).
- Strict criteria: mean framewise displacement $> .2$ mm or < 5 min data after excluding FD $> .25$ mm (Ciric et al., 2017).

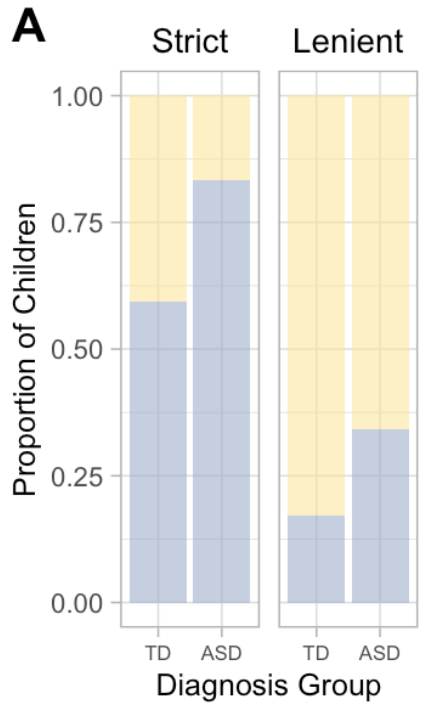


Figure: QC removes 60% and 83% of TD and ASD, respectively, under strict and 16% and 30% under lenient.

The problem: motion exclusion criteria in functional MRI causes sampling bias

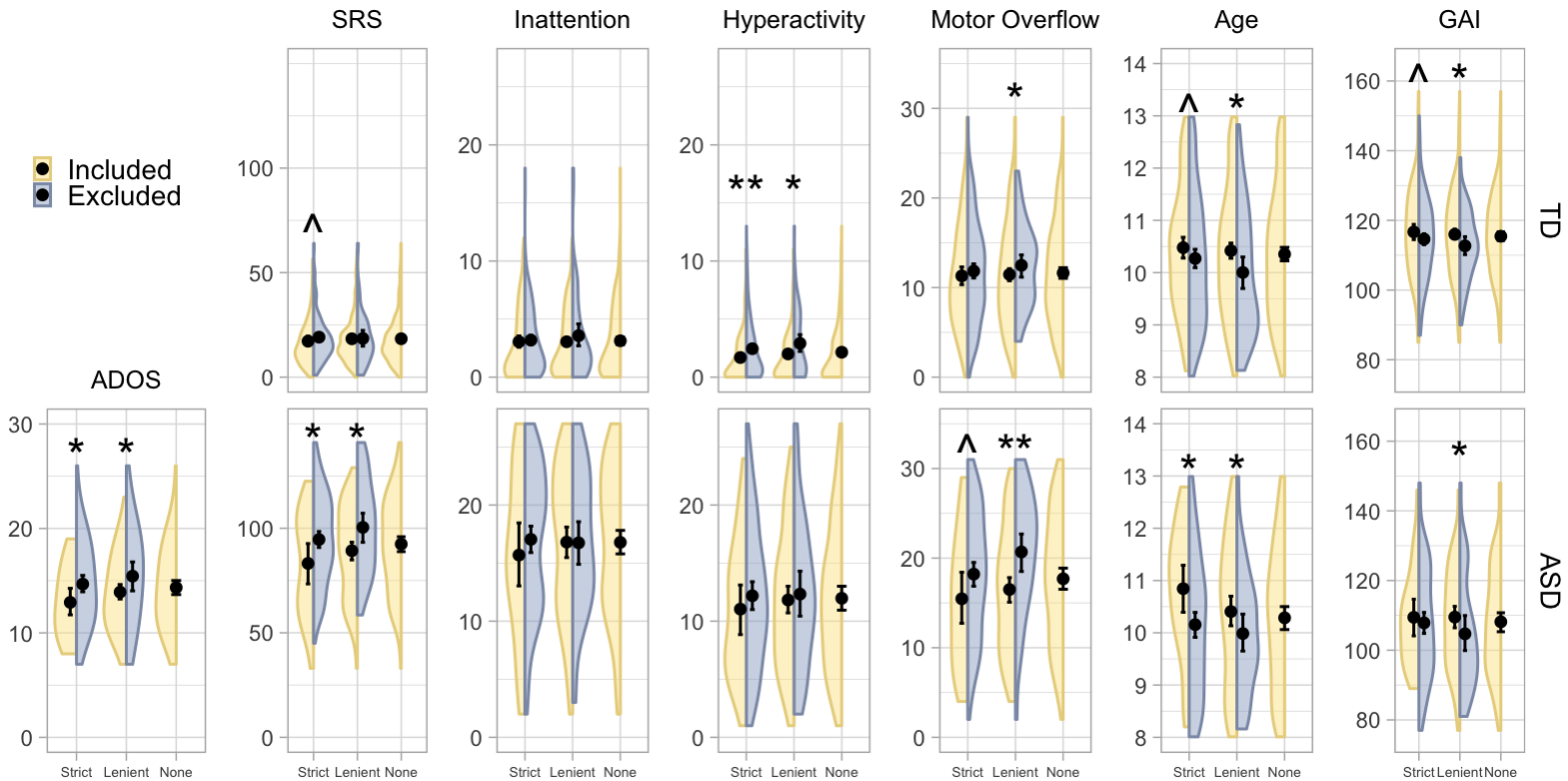


Figure: During quality control, **more severe cases of autism are excluded.** FDR-adjusted p value: ** <0.05; *<0.1; ^ <0.2.

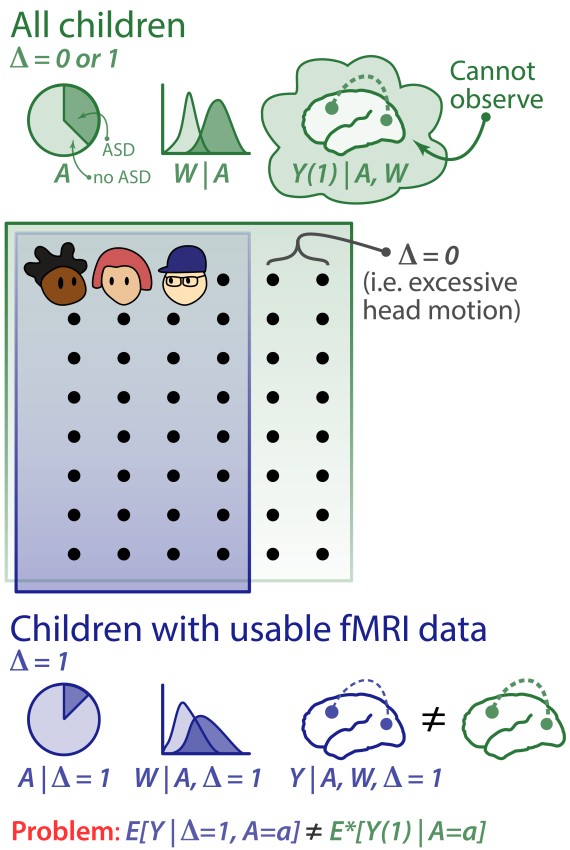
Confounding

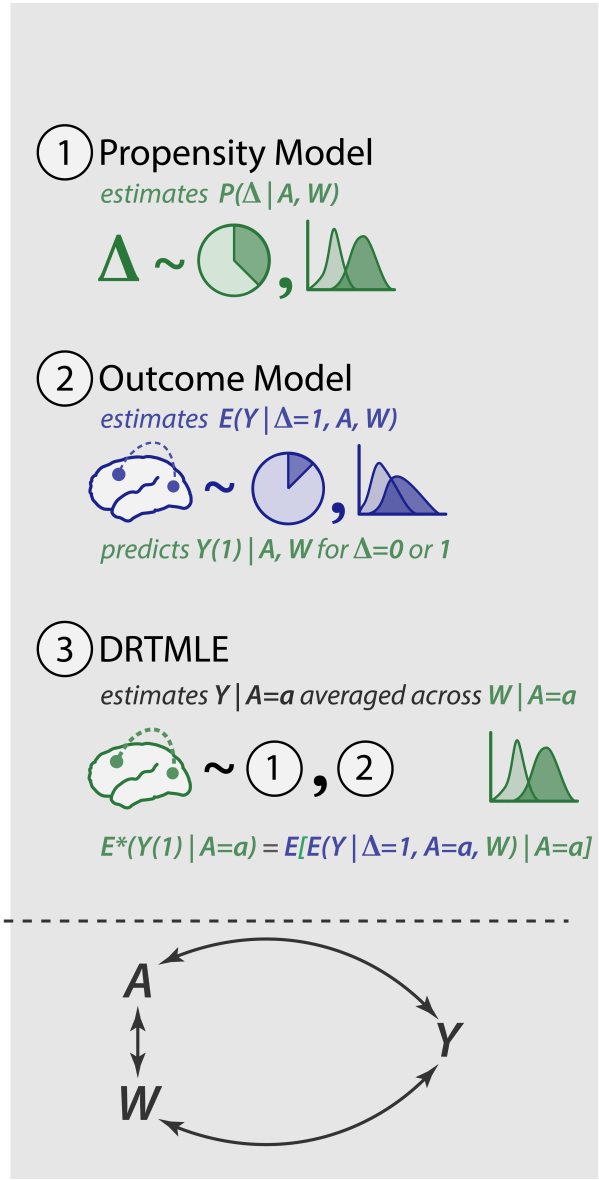
- $Y(1)$ is the counterfactual that a participant's scan is usable. Define associational parameter:

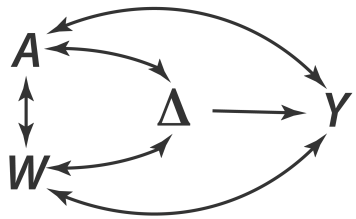
$$\begin{aligned}\psi^* &= E^*[Y(1)|A=1] - E^*[Y(1)|A=0] \\ &= E^*\left\{E^*\left(Y(1)|A=1,W\right)|\, A=1\right\} \\ &\quad - E^*\left\{E^*\left(Y(1)|A=0,W\right)|A=0\right\}\end{aligned}$$

- $\psi_{naive} = E[Y|\Delta=1,A=1] - E[Y|\Delta=1,A=0]$.
- Define confounding: $\psi_{naive} \neq \psi^*$.
- Confounding bias and selection bias: concepts overlap (Lash et al., 2021); see (Hernan and Robins, 2020) p. 80 for detailed discussion.
- Key: lack of exchangeability between usable and unusable data.
- Bias can arise when $\Delta \leftrightarrow W$, $W \leftrightarrow Y$. Then $E^*[Y(1)|A=1] \neq E[Y|\Delta=1,A=1]$

Graphical overview







Target Parameter and Identifiability Assumptions

- Define our target parameter, the deconfounded group difference:

$$\psi = E\{E(Y \mid \Delta = 1, A = 1, W) \mid A = 1\} - E\{E(Y \mid \Delta = 1, A = 0, W) \mid A = 0\}.$$

- Identifiability assumptions: $\psi^* = \psi$ if
  - (A1.1) *Mean exchangeability (conditional randomization)*: for $a = 0, 1$, $E^*\{Y(1) \mid A = a, W\} = E^*\{Y(1) \mid \Delta = 1, A = a, W\}$.
  - (A1.2) *Positivity*: for $a = 0,\ 1$ and all possible w, $P(\Delta = 1 \mid A = a, W = w) > 0$.
  - (A1.3) *Causal Consistency*: for all i such that $\Delta_i = 1$, $Y_i(1) = Y_i$.

IPWE and G-computation

- Inverse probability weighted estimator:
  - Fit propensity model: $P(\Delta_i | A_i, W_i)$.
  - Use ensemble of machine learning methods (van der Laan et al., 2007).
  - Propensity model predicting probability of inclusion to upweight usable data with small probabilities of inclusion:

$$\hat{\psi}_{IPWE} = \mathbb{E}_{n, A_i=1}\left(\frac{\Delta_i}{\hat{p}_i} Y_i\right) - \mathbb{E}_{n, A_i=0}\left(\frac{\Delta_i}{\hat{p}_i} Y_i\right).$$

- G-Computation estimator:
  - Fit outcome model:

$$\{Y_i | (\Delta_i = 1)\} = f(A_i, W_i) + \epsilon_i.$$

  - Predict values for both $\Delta_i = 0$ and $\Delta_i = 1$:

$$\hat{\psi}_{GComp} = \mathbb{E}_{n, A_i=1} \hat{Y}_i - \mathbb{E}_{n, A_i=0} \hat{Y}_i.$$

Doubly robust targeted minimum loss based estimation

- Benkeser et al. (2017) developed a doubly robust targeted minimum loss-based estimator: if *at least* one of the two regressions is consistently estimated, both $\hat{\psi}$ *and its SE* are consistently estimated.

  1. Fit propensity model.
  2. Fit outcome model.
  3. Apply DRTMLE to propensities and predicted outcomes. Involves a special iterative logistic regression.

Data Analysis

- Resting-state fMRI scans from Kennedy Krieger Institute (either 5:20 or 6:45 seconds in length) collected from 2007-2020.
- 137 ASD children without an intellectual disability and 348 typically developing.
- Use the lenient criteria.
- 96 ASD and 292 TD pass lenient criteria.

$\Delta \leftrightarrow W$

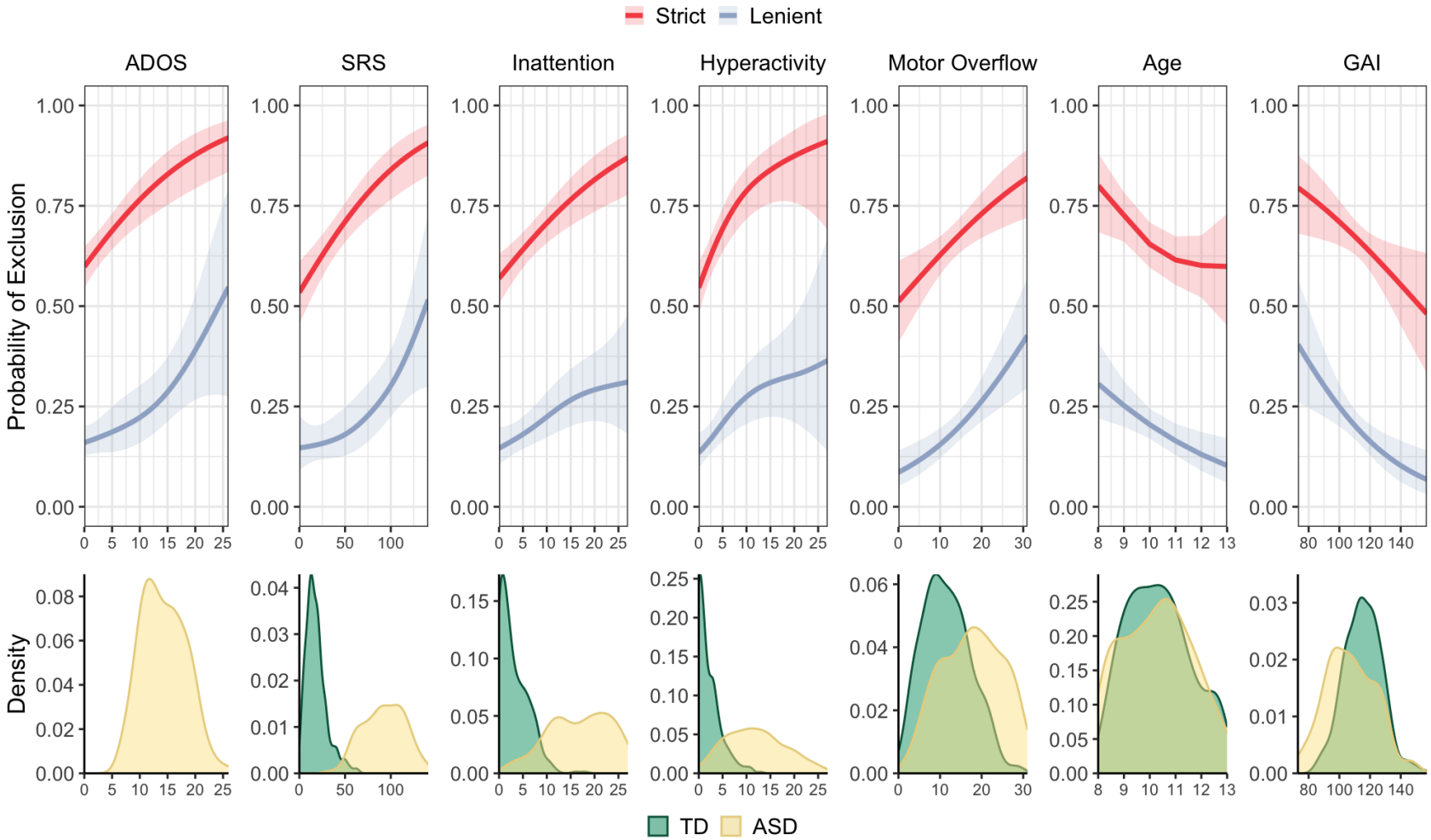


Figure: Behavioral variables are related to the probability that data are excluded.

Data Analysis: Functional Connectivity

- Group ICA with 30 components, extract subject-specific time courses, calculate partial correlations reproducing pipeline from Lombardo et al. (2019).
- Retained 18 signal components, resulting in 153 edges.
- For each edge, fit a linear model: Fisher transformed partial correlation $\sim$ diagnosis and controlling for variables that are not balanced between ASD and TD but should be (motion variables, sex, race, SES).
- Use adjusted residuals: Add diagnosis effect back into residuals.
- Naive approach with t-statistic is then approximately equal to p-value from a linear model controlling for motion, sex, race, SES (similar to Di Martino et al. 2014).

$W \leftrightarrow Y$

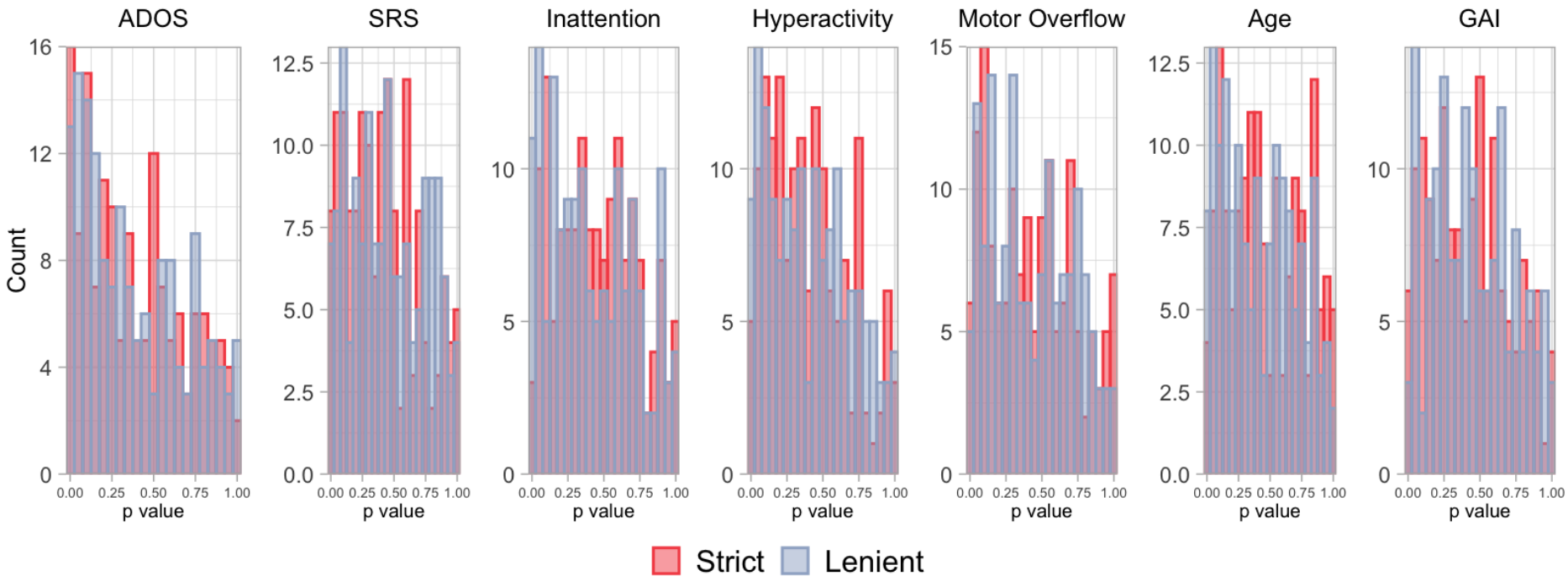


Figure: P values for generalized additive models of the relationship between edgewise functional connectivity (153 edges) in participants with usable rs-fMRI data. Blue: lenient criteria. Red: strict criteria.

Data Analysis

SuperLearner with 10-fold CV for propensity and outcome models: `SL.earth`, `SL.glmnet`, `SL.gam`, `SL.glm`, `SL.ranger`, `SL.ridge`, `SL.step`, `SL.step.interaction`, `SL.svm`, `SL.xgboost`.

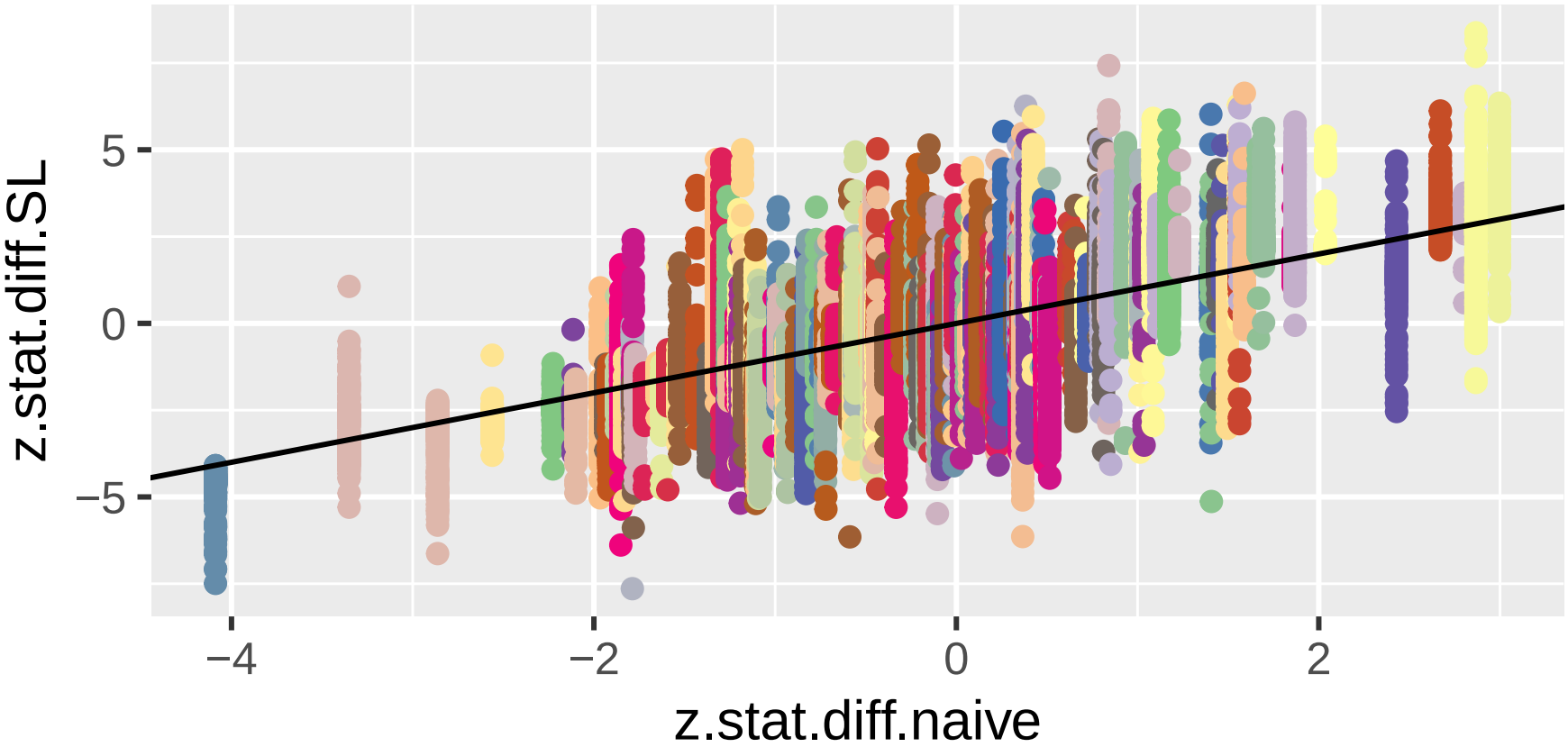
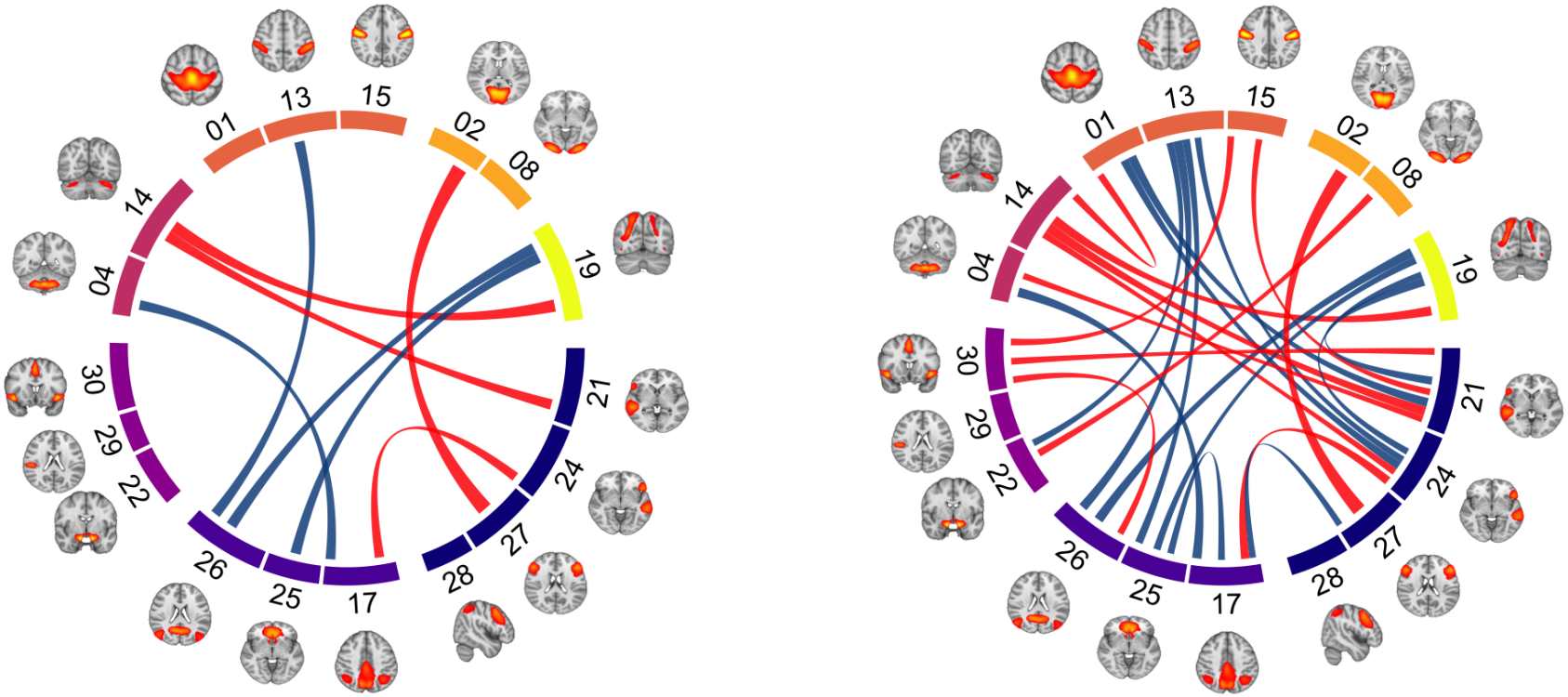


Figure: Used average z-statistic from 400 seeds.

Results



(a) Naïve

(b) DRTMLE

Figure: Z-stats of ASD-TD difference in partial correlations. Thresholded at FDR=0.20. Blue: ASD>TD. Naïve (left): 8 edges, DRTMLE (right): 25 edges.

Results, cont.

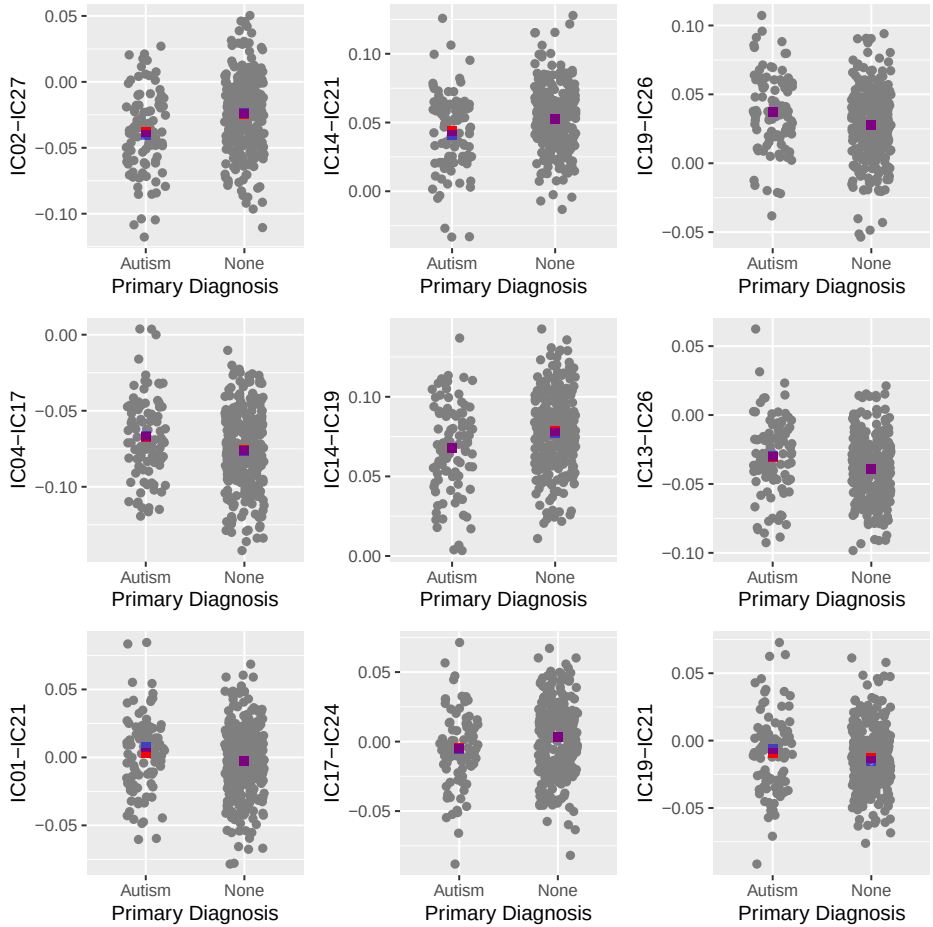


Figure: The changes in the means were small, but this drove the large differences in significance. Red: Naive. Transparent blue: DRTMLE.

Discussion and limitations

- Participant exclusion due to motion quality control creates sampling biases.
- We use DRTMLE to estimate the deconfounded group difference in a large study of autism spectrum disorder.
- More extensive differences between ASD and TD using DRTMLE.
- Limitations: only addressed missing outcome. Dataset also missing covariate values.
- Limitations: possible issues with smaller sample sizes. Develop permutation tests.
- Extend method for covariates we want to be balanced between groups (e.g., motion, age, sex, SES) versus those we want to differ (autism severity).
- Additional info: github.com/thebrisklab
- Thank you!

References I

Benkeser, D., Carone, M., van der Laan, M. J., and Gilbert, P. B. (2017). Doubly robust nonparametric inference on the average treatment effect. *Biometrika*, 104(4):863–880.

Ciric, R., Wolf, D. H., Power, J. D., Roalf, D. R., Baum, G. L., Ruparel, K., Shinohara, R. T., Elliott, M. A., Eickhoff, S. B., Davatzikos, C., Gur, R. C., Gur, R. E., Bassett, D. S., and Satterthwaite, T. D. (2017). Benchmarking of participant-level confound regression strategies for the control of motion artifact in studies of functional connectivity. *NeuroImage*, 154:174–187.

Di Martino, A., Yan, C.-G., Li, Q., Denio, E., Castellanos, F. X., Alaerts, K., Anderson, J. S., Assaf, M., Bookheimer, S., Dapretto, M., Deen, B., Delmonte, S., Dinstein, I., Ertl-Wagner, B., Fair, D., Gallagher, L., Kennedy, D., Keown, C. L., Keysers, C., Lainhart, J. E., Lord, C., Luna, B., Menon, V., Minshew, N. J., Monk, C., Mueller, S., Müller, R.-A., Nebel, M. B., Nigg, J. T., O'Hearn, K., Pelphrey, K. A., Peltier, S. J., Rudie, J. D., Sunaert, S., Thioux, M., Tyszka, J. M., Uddin, L. Q., Verhoeven, J. S., Wenderoth, N., Wiggins, J. L., Mostofsky, S. H., and Milham, M. P. (2014). The autism brain imaging data exchange: Towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular Psychiatry*, 19(6):659–667.

References II

Fassbender, C., Mukherjee, P., and Schweitzer, J. B. (2017). Reprint of: Minimizing noise in pediatric task-based functional MRI; Adolescents with developmental disabilities and typical development. *NeuroImage*, 154:230–239.

Hernan, M. A. and Robins, J. M. (2020). *Causal Inference: What If*. Chapman Hall/CRC, Boca Raton.

Lash, T. L., VanderWeele, T. J., Haneuse, S., and Rothman, K. J. (2021). *Modern Epidemiology, Fourth Edition*. Wolters Kluwer.

Lombardo, M. V., Eyler, L., Moore, A., Datko, M., Barnes, C. C., Cha, D., Courchesne, E., and Pierce, K. (2019). Default mode-visual network hypoconnectivity in an autism subtype with pronounced social visual engagement difficulties. *eLife*, 8:e47427.

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., and Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*, 59(3):2142–2154.

van der Laan, M. J., Polley, E. C., and Hubbard, A. E. (2007). Super learner. *Statistical Applications in Genetics and Molecular Biology*, 6(1).