Nonparametric motion adjustment in studies of functional connectivity alterations in autistic children: Part I

Jialu Ran[1], Sarah Shultz[2], Benjamin Risk[1] and David Benkeser[1]

[1] *Department of Biostatistics and Bioinformatics,*

[2] *Department of Pediatrics, Department of Neuroscience*

Emory University

Co-authors

Figure: Jialu Ran (left), Sarah Shultz (center), David Benkeser (right)

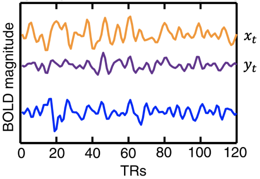
Overview

Introduction: Autism spectrum disorder

Autism spectrum disorder (ASD):

- Approximately 1 in 36 children in the US (Maenner 2023)
- Deficits in social communication and interaction; restricted and repetitive behaviors, interests, and activities
- Functional connectivity is used to study ASD

fMRI (functional magnetic resonance imaging)



brain activity (time series) of three selected brain regions

$$\text{Functional connectivity}_i(\text{seed}, \text{region}_j) = Corr(Y_{i,seed,t}, Y_{i,j,t}) = Y_{ij}$$

- In ASD, disruptions of functional connectivity thought to involve the default mode network. (Yerys et al. 2015)

Default mode decreased connectivity

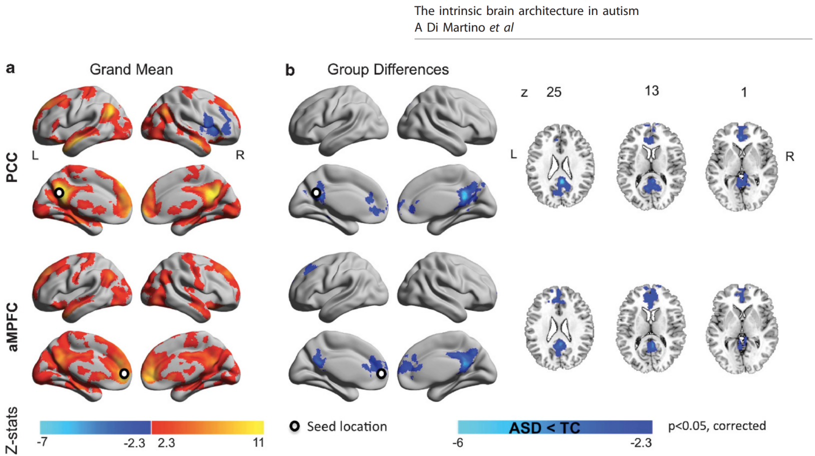


Figure: Di Martino et al. (2014)

Introduction: Movement during scan

Problem — Motion:

- ASD participants are more likely to move during scan.
- Patterns of correlation induced by motion can mimic the connectivity hypothesis of autism (Deen and Pelphrey 2012).
- Difficult to disentangle neural and motion related effects.

Motion artifacts

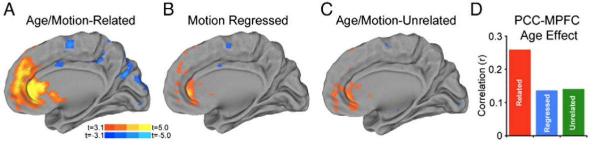
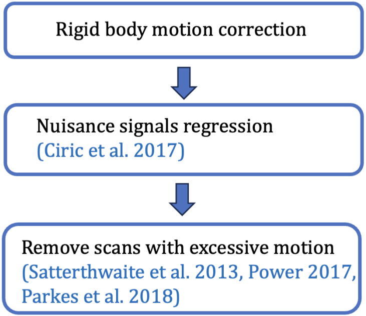


Fig. 6. Effect of motion on estimates of age-related connectivity change from a posterior cingulate seed. In a sample of 421 subjects where age and motion were related, increasing subject age was associated with increased connectivity between the PCC and the MPFC (A). This effect, while still significantly present, was attenuated when motion was included as a confound regressor in the group level analysis (B) or when the *age/motion-unrelated* subsample of 348 subjects was used (C). The correlation of age with pairwise PCC-MPFC connectivity was reduced substantially when motion was taken into account (D).

Figure: Satterthwaite et al. (2012): age-related differences partly due to younger children moving more than older children.

Introduction: Motion Quality Control

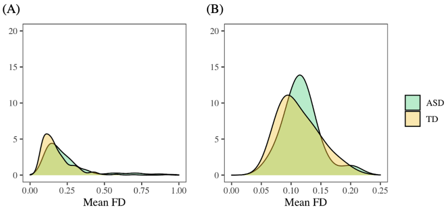


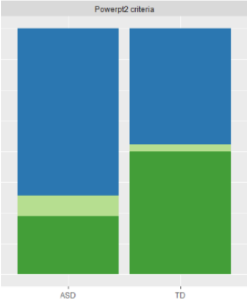
- Existing motion control using regression is inadequate because motion patterns are complex (Power et al. 2014).
- Current best practices **remove scans with excessive motion**.

Problem II: Motion QC creates issues

- Motion control leads to drastic reductions in sample size.
- ABCD study with school-aged children **removed $60 - 75\%$** of children due to excessive motion (Marek et al. 2022, Nielsen et al. 2019).
- This creates selection bias, disproportionately selecting for: higher SES, White participants, older, females, higher neurocognitive skills, fewer neurodevelopmental problems (Cosgrove et al. 2022).
- Children removed by motion QC in autism studies tend to have more severe symptomatology (Nebel et al. 2022).

Introduction: Scan Exclusion





Top: ASD Children move more.

Bottom: Most ASD children are removed in a typical analysis. Green: included participants using >5 minutes FD<0.2 mm criteria (Power et al. 2014). Children that move more tend to have more severe symptomatology (Nebel et al. 2022).

Motivation: A fair comparison

Goal

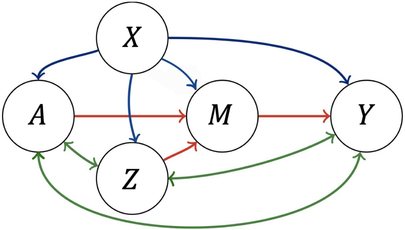
Estimate the neural differences between ASD and non-ASD children while including data from all participants.

Our Proposal – a causal mediation framework

- no participant removal.
- remove both motion bias and QC selection bias.
- view motion as a mediator:
  - direct association of ASD with functional connectivity.
  - indirect effect *mediated by motion*.
  - machine learning to model effects of motion.
  - stochastic intervention to reduce motion.

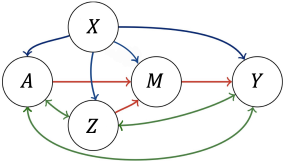
Overview

Causal Framework: Notation



- $Y \in \mathbb{R}$: functional connectivity between two locations in the brain
- $A \in \{0,1\}$: non-ASD (0), ASD (1) [association not causal effect]
- $M \in \mathbb{R}$: a motion variable (mean framewise displacement)
- X: demographic confounders (age, sex and handedness)
- Z: variables intrinsically related to diagnostic group (autism diagnostic score, IQ, medication status)

Causal Framework: Raw association



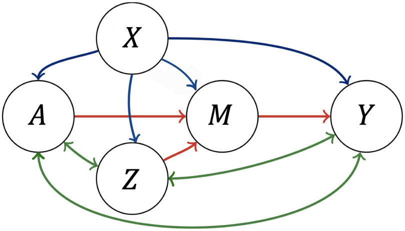
Raw association (balances demographics, includes motion):

$$\theta_{\mathrm{R},1} - \theta_{\mathrm{R},0} = E[E(Y|A=1,X)] - E[E(Y|A=0,X)]$$

where the outer expectation is wrt $P_X(x)$.

Associations of biological interest: $A \leftrightarrow Y$, $A \leftrightarrow Z \leftrightarrow Y$
Associations attributed to motion: $A \rightarrow M \rightarrow Y, A \leftrightarrow Z \rightarrow M \rightarrow Y$
Use **stochastic interventions** to control for motion

Causal Framework: Decomposition



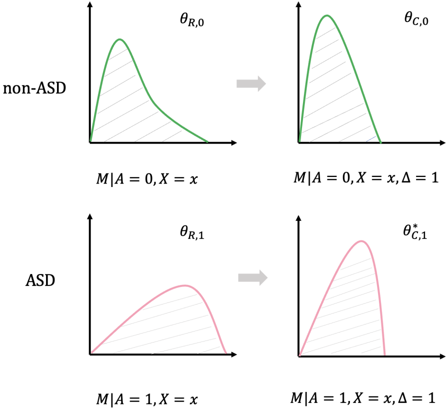
decomposition of the raw association

$$\underbrace{\theta_{R,1} - \theta_{R,0}}_{\substack{\text{total association} \\ \text{ASD vs. non-ASD}}} = \underbrace{\theta_{R,1} - \theta^*_{C,1}}_{\substack{\text{effect of initial} \\ \text{training in ASD group}}} - \underbrace{(\theta_{R,0} - \theta_{C,0})}_{\substack{\text{effect of initial} \\ \text{training in non-ASD group}}}$$

$$+ \underbrace{\theta^*_{C,1} - \theta_{C,1}}_{\substack{\text{effect of additional} \\ \text{training in ASD group}}} + \underbrace{\theta_{C,1} - \theta_{C,0}}_{\substack{\textbf{motion-adjusted} \\ \textbf{neural association}}}$$

Initial training: control for motion

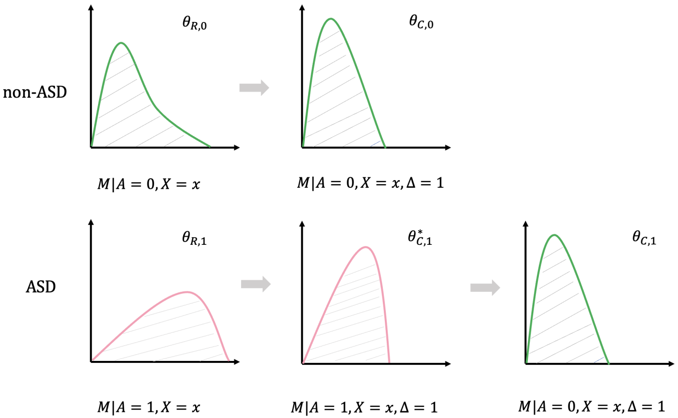
Motion reduced to a tolerable level, Δ:
For all children: replace motion M by $M_a \sim P_{M|\Delta=1,a,X} \longrightarrow Y(M_a)$



$$\theta^*_{C,1} - \theta_{C,0} = E[E_C(Y(M_1) \mid A = 1, X)] - E[E(Y_C(M_0) \mid A = 0, X)]$$

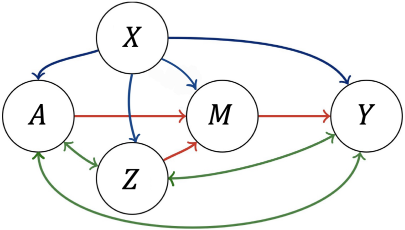
Additional training for ASD children

For ASD children: replace motion M by $M_0 \sim P_{M|\Delta=1,0,X} \longrightarrow Y(M_0)$



$$\theta_{C,1} - \theta_{C,0} = E[E_{C^*}(Y(M_0) \mid A=1, X)] - E[E_C(Y(M_0) \mid A=0, X)]$$

Causal Framework: Decomposition



decomposition of the raw association

$$\underbrace{\theta_{\mathrm{R},1} - \theta_{\mathrm{R},0}}_{\substack{\text{total association} \\ \text{ASD vs. non-ASD}}} = \underbrace{\theta_{\mathrm{R},1} - \theta^*_{\mathrm{C},1}}_{\substack{\text{effect of initial} \\ \text{training in ASD group}}} - \underbrace{(\theta_{\mathrm{R},0} - \theta_{\mathrm{C},0})}_{\substack{\text{effect of initial} \\ \text{training in non-ASD group}}}$$

$$+ \underbrace{\theta^*_{\mathrm{C},1} - \theta_{\mathrm{C},1}}_{\substack{\text{effect of additional} \\ \text{training in ASD group}}} + \underbrace{\theta_{\mathrm{C},1} - \theta_{\mathrm{C},0}}_{\substack{\textbf{motion-adjusted} \\ \textbf{neural association}}}$$

Overview

Identification

Theorem (Identifiability)

Under the following assumptions:

(A1) *Mean exchangeability: for all m such that $P\{p_{M|\Delta=1,0,X}(m \mid X) > 0\} > 0$, $E_C\{Y(m) \mid A = a, X, Z\} = E_C\{Y(m) \mid A = a, M = m, X, Z\}$ a.e. P;*

(A2) *Positivity:*

- (A2.1) *for every x such that $p_X(x) > 0$, we also have $p_{a|X}(x) > 0$ for $a = 0, 1$;*
- (A2.2) *for every (x, z, m) such that $p_X(x) p_{Z|a,X}(z \mid x) p_{M|\Delta=1,0,X}(m \mid x) > 0$, we also have that $p_{M|a,X,Z}(m \mid x, z) > 0$ for $a = 0, 1$.*

(A3) *Causal Consistency: for any child with observed motion value $M = m$, the observed functional connectivity measurement Y is equal to the counterfactual functional connectivity measurement $Y(m)$.*

The counterfactual $\theta_{C,a}$ is identified by θ_a, where

$$\theta_{\mathbf{1}} = \iiint \mu_{Y|\mathbf{1},M,X,Z}(m,x,z)\, p_{Z|\mathbf{1},X}(z \mid x)\, p_{M|\Delta=1,\mathbf{0},X}(m \mid x)\, p_X(x) dz\, dm\, dx$$

$$\theta_{\mathbf{0}} = \iiint \mu_{Y|\mathbf{0},M,X,Z}(m,x,z)\, p_{Z|\mathbf{0},X}(z \mid x)\, p_{M|\Delta=1,\mathbf{0},X}(m \mid x)\, p_X(x) dz\, dm\, dx$$

Estimation: plug-in estimator

Estimand:

$$\theta_a = \iiint \mu_{Y|a,M,X,Z}(m,x,z)\, p_{Z|a,X}(z \mid x)\, p_{M|\Delta=1,0,X}(m \mid x)\, p_X(x) dz\, dm\, dx$$

Step I:

$$\to \iiint \mu_{Y|a,M,X,Z}(m,x,z)\, p_{Z|a,X}(z \mid x)\, p_{M|\Delta=1,0,X}(m \mid x)\, p_X(x) dz\, dm\, dx$$

→ regress $Y_i \sim M_i, A_i, X_i, Z_i$
→ Flexible form of regression is used: Super Learning (Van der Laan et al. 2007)

- pre-specifying a *library* of candidate regression estimators, e.g.: multivariate adaptive regression splines, LASSO, ridge regression, generalized additive models, generalized linear models (with/without interactions and with/without forward stepwise covariate selection), random forest, and Xgboost
- cross-validation is used to build a weighted combination of these estimators

Estimation: plug-in estimator

Step II: $\theta_a = \iint \int \mu_{Y|a,M,X,Z}(m,x,z) p_{M|\Delta=1,0,X}(m \mid x) dm \, p_{Z|a,X}(z \mid x) dz \, p_X(x) \, dm \, dx$

$$\int \mu_{Y|a,M,X,Z}(m,x,z) p_{M|\Delta=1,0,X}(m \mid x) dm =$$

$$\int \mu_{Y|a,M,X,Z}(m,x,z) \frac{p_{M|\Delta=1,0,X}(m \mid x)}{p_{M|\Delta=1,0,X,Z}(m \mid x,z)} p_{M|\Delta=1,0,X,Z}(m \mid x,z) \, dm =$$

$$E\left[\mu_{Y|a,M,X,Z}(M,X,Z) \frac{p_{M|\Delta=1,0,X}(M \mid X)}{p_{M|\Delta=1,0,X,Z}(M \mid X,Z)} \,\middle|\, \Delta=1, A=0, X=x, Z=z \right].$$

Avoid numeric integration and instead utilize super learning-based mean regression for estimation (Díaz et al. 2021).

Use highly adaptive lasso conditional density estimation (Hejazi et al. 2022).

For $a=0,1$ and $A_i=0$ and $\Delta_i=1$, pseudo-regression on:

$$\mu_{n,Y|a,M,X,Z}(M_i,X_i,Z_i) \frac{p_{n,M|\Delta=1,0,X}(M_i \mid X_i)}{p_{n,M|\Delta=1,0,X,Z}(M_i \mid X_i,Z_i)} \sim X_i, Z_i,$$

then evaluate regression function at $i=1,\ldots,n$.

Estimation: plug-in estimation

Step III:

$$\theta_a = \int \int \left\{ \int \mu_{Y|a,M,X,Z}(m,x,z) p_{M|\Delta=1,0,X}(m \mid x) dm \right\} p_{Z|a,X}(z \mid x) dz \, p_X(x) \, dx$$

Let $\eta_{n,\mu|a,X,Z}$ be the regression estimated in Step II evaluated at $\{X_1, Z_1\}, \ldots, \{X_n, Z_n\}$ for $a = 0, 1$.

Then regress $\eta_{n,\mu|a,X,Z}(X_i, Z_i) \sim X_i$

Step IV:

$$\theta_a = \int \left\{ \iint \mu_{Y|a,M,X,Z}(m,x,z) p_{M|\Delta=1,0,X}(m \mid x) dm \, p_{Z|a,X}(z \mid x) dz \right\} p_X(x) \, dx$$

Let $\xi_{n,\eta|a,X}(X_i)$ be the regression estimated in Step III evaluated $i = 1, \ldots, n$. Then,

$$\theta_{n,a} = \frac{1}{n} \sum_i \xi_{n,\eta|a,X}(X_i).$$

Theorem (Efficient Influence Function)

Define

$$\pi_a(x) = P(A = a|X = x),$$
$$\bar{\pi}_0(x) = P(A = 0|X = x)P(\Delta = 1 \mid A = 0, X = x),$$
$$r_a(m, x, z) = \frac{p_{M|\Delta=1,0,X}(m \mid x)}{p_{M|a,X,Z}(m \mid x, z)}.$$

In a nonparametric model, the efficient influence function for θ_a evaluated on an observation O_i is

$$D_{P,a}(O_i) = \frac{\mathbb{1}_a(A_i)}{\pi_a(X_i)} r_a(M_i, X_i, Z_i)\{Y_i - \mu_{Y|a,M,X,Z}(M_i, X_i, Z_i)\}$$
$$+ \frac{\mathbb{1}_a(A_i)}{\pi_a(X_i)}\{\eta_{\mu|a,Z,X}(X_i, Z_i) - \xi_{\eta|a,X}(X_i)\}$$
$$+ \frac{\mathbb{1}_{a,1}(A_i, \Delta_i)}{\bar{\pi}_0(X_i)}\left\{\eta_{\mu|a,M,X}(M_i, X_i) - \xi_{\eta|a,X}(X_i)\right\} + \xi_{\eta|a,X}(X_i) - \theta_a.$$

Estimation: one-step estimator

Plug-in estimator

- obtained by sequential regression
- not robust; motivates our one-step estimator (Bickel et al. 1993)

One-step estimator

$$\theta_{n,a}^{+} = \theta_{n,a} + \frac{1}{n}\sum_{i=1}^{n} D_{a,P_n}(O_i)$$

- $D_{a,P}(O)$ [EIF] estimated using sequential regression
- multiple robustness
- asymptotically linear with an analytical confidence interval

Asymptotic linearity of the one-step estimator

Theorem (Asymptotic linearity)

Under the following assumptions,

(i) *Positivity of estimates:* $\pi_{n,a} > \epsilon_1$ *for some* $\epsilon_1 > 0$, $\bar{\pi}_{n,0} > \epsilon_2$ *for some* $\epsilon_2 > 0$, *and* $\frac{p_{n,M|\Delta=1,0,X}}{p_{n,M|a,X,Z}} < \epsilon_3$ *for some* $\epsilon_3 < \infty$*;*

(ii) $n^{1/2}$*-convergence of second order terms:* $\|\xi_{n,\eta|a,X} - \xi_{\eta|a,X}\| \|\pi_{n,a} - \pi_a\| = o_P(n^{-1/2})$, $\|\mu_{n,Y|a,M,X,Z} - \mu_{Y|a,M,X,Z}\|(\|p_{n,M|\Delta=1,0,X} - p_{M|\Delta=1,0,X}\| + \|p_{n,M|a,X,Z} - p_{M|a,X,Z}\|) = o_P(n^{-1/2})$, *and* $\|p_{n,M|\Delta=1,0,X} - p_{M|\Delta=1,0,X}\|(\|\eta_{n,\mu|a,M,X} - \eta_{\mu|a,M,X}\| + \|\pi_{n,\Delta=1|0,X} - \pi_{\Delta=1|0,X}\|) = o_P(n^{-1/2})$.

(iii) $L^2(P)$*-consistent influence function estimate:*

$$\int [\{D_{a,P_\ell}(o) - D_{a,P_n}(o)\}^2] dP(o) = o_P(1),$$

where P_ℓ *denotes the limit of* P_n *as* $n \to \infty$.

(iv) *Donsker influence function estimate:* D_{a,P_n} *falls in a* P*-Donsker class with probability tending to 1.*

Then,

$$\theta^+_{n,a} - \theta_a = \frac{1}{n}\sum_{i=1}^{n} D_{a,P}(O_i) + o_P(n^{-1/2})$$

Multiple robustness

Lemma (Multiple robustness)

Consider some $P' \in \mathcal{M}$. If any one of the following conditions hold, then $E[D_{P',a}(O)] = 0$, i.e., θ_a is consistently estimated:
(i) $p'_{M|\Delta=1,0,X} = p_{M|\Delta=1,0,X}$ and $p'_{M|a,X,Z} = p_{M|a,X,Z}$ and either $\pi'_a = \pi_a$ or $\xi'_{\eta|a,X} = \xi_{\eta|a,X}$;
(ii) $\mu'_{Y|a,M,X,Z} = \mu_{Y|a,M,X,Z}$, $\pi'_a = \pi_a$, $\pi'_{\Delta=1|0,X} = \pi_{\Delta=1|0,X}$, and $\eta'_{\mu|a,M,X} = \eta_{\mu|a,M,X}$;
(iii) $\mu'_{Y|a,M,X,Z} = \mu_{Y|a,M,X,Z}$ and $p'_{M|\Delta=1,0,X} = p_{M|\Delta=1,0,X}$ and either $\pi'_a = \pi_a$ or $\xi'_{\eta|a,X} = \xi_{\eta|a,X}$.

Multiple robustness

- $E[D_{P',a}(O)] = 0$ implies θ_a is consistently estimated.
- Thus, only *some* of the nuisance parameters need to be consistently estimated to achieve consistency of our estimate of θ_a.
- For example, condition (i) implies that obtaining consistent estimates of the conditional motion densities, $p_{M|\Delta=1,0,X}$ and $p_{M|a,X,Z}$, and the conditional probability of ASD as a function of covariates π_a is sufficient to ensure a consistent estimator of θ_a.

Simultaneous confidence bands

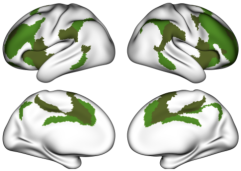
$$\begin{pmatrix} \theta^+_{n,1,1} - \theta^+_{n,0,1} \\ \vdots \\ \theta^+_{n,1,J} - \theta^+_{n,0,J} \end{pmatrix} - \begin{pmatrix} \theta_{1,1} - \theta_{0,1} \\ \vdots \\ \theta_{1,J} - \theta_{0,J} \end{pmatrix} \to N\left\{ \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \mathrm{Cov}\begin{pmatrix} D_{P,1,1}(O) - D_{P,0,1}(O) \\ \vdots \\ D_{P,1,J}(O) - D_{P,0,J}(O) \end{pmatrix} \right\}$$

- Calculate asymptotic covariance matrix for the J locations.
- Estimate simultaneous confidence bands.
- Idea is from (Ruppert et al. 2003), where we generate realizations from the multivariate normal distribution with covariance defined from EIF to generate FWER-control with $z_{\max,1-\alpha}$.

Overview

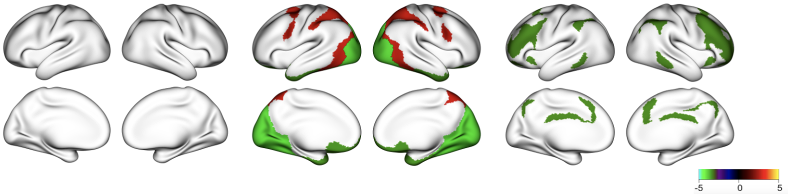
Simulation: Evaluating estimators in the context of ASD

Truth



Significant associations

naïve (with participant removal) naïve (all data) our method (all data)



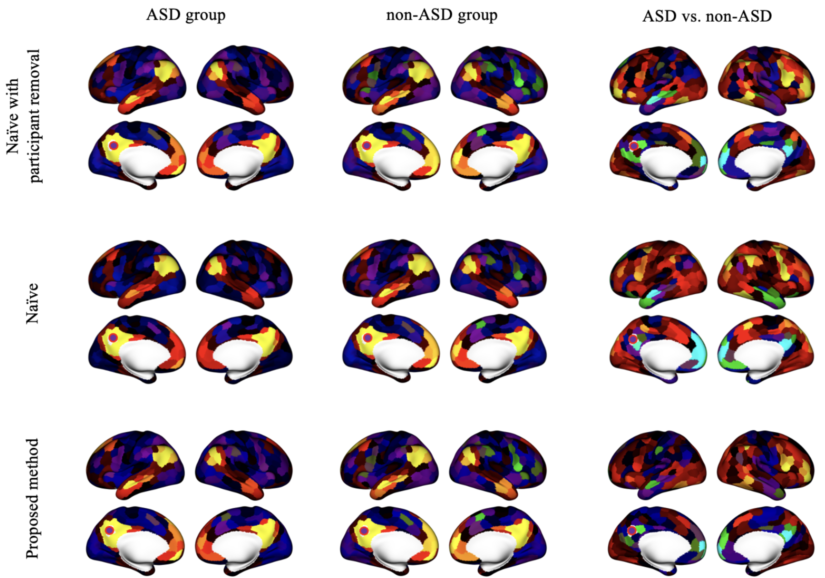
Overview

Data Analysis: ABIDE Data

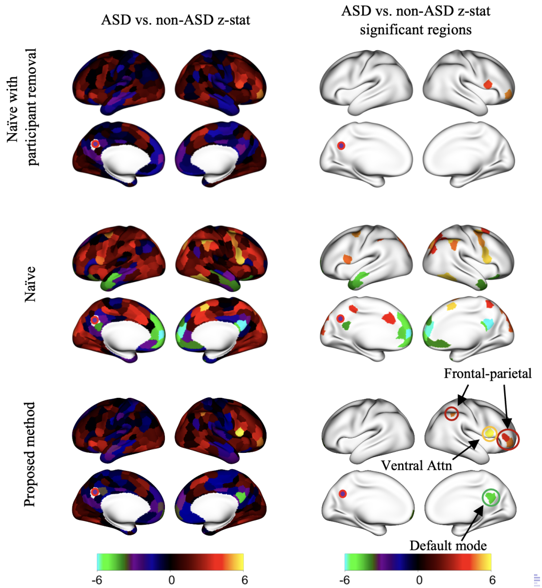
School-age children from Autism Brain Imaging Data Exchange (ABIDEI and ABIDEII) Dataset (Di Martino et al. 2014; 2017)

- in-house preprocessing [fmriprep with 9 parameters nuisance regression]
- site harmonization [neuroCombat]
- variables:
    - A: 245 TD ($A = 0$), 132 ASD ($A = 1$) [377 8-13 yo children]
    - X: age, sex, handedness
    - Z: autism diagnostic observation schedule, IQ, medication status
    - M: mean frame-wise displacement (FD)
    - $\Delta = 1$: > 5 minutes of data free from ≥ 0.2 framewise displacement (Power et al. 2014) [126 TD (51%), 34 ASD (26%)]
    - Y_j: correlation between seed region in DMN and region j, $j = 1, ..., 400$. (Schaeffer 400 atlas).

Data Analysis: Estimation



Data Analysis: Significance



Overview

Discussion

We introduce a causal mediation framework to estimate the neural association of ASD with functional connectivity

- Stochastic intervention to estimate what ASD brain looks like under tolerable motion.
- Incorporate flexible data-adaptive regression techniques.
- In ABIDE data analysis, removed motion artifacts while more efficiently using data.
- Can be viewed as machine-learning based standardization of motion.

Current projects

- Prospective study at Emory.
- All children receive training in mock scanner with MoTrak.
- Autistic children receive additional training.
- https://www.brainconnectivitystudy.org/.

Acknowledgement

Thank you!

All authors were supported by the National Institute of Mental Health of the National Institutes of Health under award number R01 MH129855. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References I

P. J. Bickel, C. A. Klaassen, Y. Ritov, and J. A. Wellner. *Efficient and adaptive estimation for semi-parametric models*, volume 4. Springer, 1993.

K. T. Cosgrove, T. J. McDermott, E. J. White, M. W. Mosconi, W. K. Thompson, M. P. Paulus, C. Cardenas-Iniguez, and R. L. Aupperle. Limits to the generalizability of resting-state functional magnetic resonance imaging studies of youth: An examination of abcd study® baseline data. *Brain imaging and behavior*, 16(4):1919–1925, 2022.

B. Deen and K. Pelphrey. Perspective: brain scans need a rethink. *Nature*, 491(7422):S20–S20, 2012.

A. Di Martino, C.-G. Yan, Q. Li, E. Denio, F. X. Castellanos, K. Alaerts, J. S. Anderson, M. Assaf, S. Y. Bookheimer, M. Dapretto, et al. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular psychiatry*, 19(6):659–667, 2014.

A. Di Martino, D. O'connor, B. Chen, K. Alaerts, J. S. Anderson, M. Assaf, J. H. Balsters, L. Baxter, A. Beggiato, S. Bernaerts, et al. Enhancing studies of the connectome in autism using the autism brain imaging data exchange ii. *Scientific data*, 4(1):1–15, 2017.

I. Díaz, N. S. Hejazi, K. E. Rudolph, and M. J. van Der Laan. Nonparametric efficient causal mediation with intermediate confounders. *Biometrika*, 108(3):627–641, 2021.

N. S. Hejazi, M. J. van der Laan, and D. C. Benkeser. haldensify: Highly adaptive lasso conditional density estimation in R. *Journal of Open Source Software*, 2022. doi: 10.21105/joss.04522. URL https://doi.org/10.21105/joss.04522.

M. J. Maenner. Prevalence and characteristics of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 sites, united states, 2020. *MMWR. Surveillance Summaries*, 72, 2023.

References II

S. Marek, B. Tervo-Clemmens, F. J. Calabro, D. F. Montez, B. P. Kay, A. S. Hatoum, M. R. Donohue, W. Foran, R. L. Miller, T. J. Hendrickson, S. M. Malone, S. Kandala, E. Feczko, O. Miranda-Dominguez, A. M. Graham, E. A. Earl, A. J. Perrone, M. Cordova, O. Doyle, L. A. Moore, G. M. Conan, J. Uriarte, K. Snider, B. J. Lynch, J. C. Wilgenbusch, T. Pengo, A. Tam, J. Chen, D. J. Newbold, A. Zheng, N. A. Seider, A. N. Van, A. Metoki, R. J. Chauvin, T. O. Laumann, D. J. Greene, S. E. Petersen, H. Garavan, W. K. Thompson, T. E. Nichols, B. T. Yeo, D. M. Barch, B. Luna, D. A. Fair, and N. U. Dosenbach. Reproducible brain-wide association studies require thousands of individuals. *Nature 2022 603:7902*, 603(7902):654–660, 3 2022. ISSN 1476-4687. doi: 10.1038/s41586-022-04492-9. URL https://www.nature.com/articles/s41586-022-04492-9.

M. B. Nebel, D. E. Lidstone, L. Wang, D. Benkeser, S. H. Mostofsky, and B. B. Risk. Accounting for motion in resting-state fmri: What part of the spectrum are we characterizing in autism spectrum disorder? *NeuroImage*, 257:119296, 2022.

A. N. Nielsen, D. J. Greene, C. Gratton, N. U. Dosenbach, S. E. Petersen, and B. L. Schlaggar. Evaluating the prediction of brain maturity from functional connectivity after motion artifact denoising. *Cerebral Cortex*, 29(6):2455–2469, 2019.

J. D. Power, A. Mitra, T. O. Laumann, A. Z. Snyder, B. L. Schlaggar, and S. E. Petersen. Methods to detect, characterize, and remove motion artifact in resting state fmri. *Neuroimage*, 84:320–341, 2014.

D. Ruppert, M. P. Wand, and R. J. Carroll. *Semiparametric regression*. Number 12. Cambridge university press, 2003.

T. D. Satterthwaite, D. H. Wolf, J. Loughead, K. Ruparel, M. A. Elliott, H. Hakonarson, R. C. Gur, and R. E. Gur. Impact of in-scanner head motion on multiple measures of functional connectivity: relevance for studies of neurodevelopment in youth. *Neuroimage*, 60(1):623–632, 2012.

M. J. Van der Laan, E. C. Polley, and A. E. Hubbard. Super learner. *Statistical applications in genetics and molecular biology*, 6(1), 2007.

B. E. Yerys, E. M. Gordon, D. N. Abrams, T. D. Satterthwaite, R. Weinblatt, K. F. Jankowski, J. Strang, L. Kenworthy, W. D. Gaillard, and C. J. Vaidya. Default mode network segregation and social deficits in autism spectrum disorder: Evidence from non-medicated children. *NeuroImage: Clinical*, 9:223–232, 2015.

Overview

Simulation: Confirming theoretical properties of estimators

- Simulation Setting
  $X \sim \text{Bin}(1, \frac{1}{2})$
  $A \sim \text{Bin}(1, \text{expit}(X - \frac{1}{4}))$
  $Z \sim \text{Bin}(1, \text{expit}(\frac{5}{4}A - \frac{1}{2}))$
  $M \sim \text{N}(1 + A + X/2 - Z/4, 1)$
  $Y \sim \text{N}(-1 + X/2 - Z/3 - A/4 + M/5, 1)$
  sample size $n \in \{200, 500, 1000, 2000, 4000\}$
- evaluate proposed estimators of θ_0 and θ_1

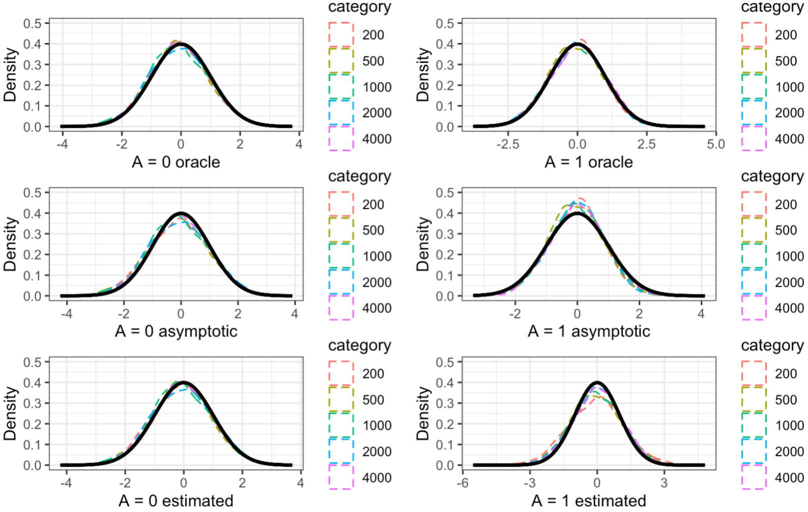
Simulation: Confirming theoretical properties of estimators

Case I: all nuisance parameters are consistently estimated at appropriate rates

n	$\theta^{\text{cf}}_{n,0}$				$\theta^{\text{cf}}_{n,1}$			
	$n^{1/2}$ bias	$n^{1/2}$ sd	sd ratio	cover	$n^{1/2}$ bias	$n^{1/2}$ sd	sd ratio	cover
200	-0.235	2.063	1.075	0.929	-0.246	2.358	1.430	0.851
500	-0.150	1.938	0.986	0.951	-0.310	2.369	1.205	0.900
1000	-0.141	2.003	1.028	0.940	-0.113	2.333	1.110	0.922
2000	-0.026	1.977	1.033	0.940	-0.077	2.328	1.056	0.931
4000	-0.006	1.913	1.014	0.950	0.076	2.074	0.914	0.979

Table: All nuisance parameters are consistently estimated at appropriate rates with the use of cross-fitting

Simulation: Confirming theoretical properties of estimators



Simulation: Multiple Robustness

Case II: five scenarios in which only certain combinations of nuisance parameters were correctly specified

Setting	n	$\text{bias}_{\theta^{\text{cf}}_{n,0}}$	$\text{sd}_{\theta^{\text{cf}}_{n,0}}$	$\text{bias}_{\theta^{\text{cf}}_{n,1}}$	$\text{sd}_{\theta^{\text{cf}}_{n,1}}$
$p_{M\mid\Delta=1,0,X}, p_{M\mid a,X,Z}, \pi_a$ correct	200	0.0085	0.1492	0.0915	0.1608
	500	0.0076	0.0890	0.0568	0.0912
	1000	0.0024	0.0627	0.0423	0.0700
	2000	0.0043	0.0440	0.0335	0.0502
	4000	0.0037	0.0309	0.0281	0.0358
$p_{M\mid\Delta=1,0,X}, p_{M\mid a,X,Z}, \xi_{\eta\mid a,X}$ correct	200	-0.0190	0.1458	-0.0318	0.1827
	500	-0.0116	0.0855	-0.0158	0.1030
	1000	-0.0100	0.0600	-0.0082	0.0762
	2000	-0.0040	0.0410	-0.0025	0.0537
	4000	-0.0006	0.0294	0.0022	0.0379
$\pi_a, \bar{\pi}_0, \eta_{\mu\mid a,M,X}, \mu_{Y\mid a,M,X,Z}$ correct	200	-0.0037	0.1472	-0.0137	0.1679
	500	-0.0023	0.0880	-0.0061	0.1012
	1000	-0.0043	0.0616	-0.0026	0.0742
	2000	-0.0007	0.0432	0.0001	0.0527
	4000	0.0009	0.0305	0.0019	0.0372
$\pi_a, p_{M\mid\Delta=1,0,X}, \mu_{Y\mid a,M,X,Z}$ correct	200	-0.0171	0.1512	-0.0273	0.1810
	500	-0.0096	0.0879	-0.0142	0.1019
	1000	-0.0090	0.0621	-0.0074	0.0747
	2000	-0.0032	0.0433	-0.0018	0.0530
	4000	-0.0002	0.0306	0.0019	0.0374
$p_{M\mid\Delta=1,0,X}, \xi_{\eta\mid a,X}, \mu_{Y\mid a,M,X,Z}$ correct	200	-0.0190	0.1458	-0.0318	0.1827
	500	-0.0116	0.0855	-0.0158	0.1030
	1000	-0.0100	0.0600	-0.0082	0.0762
	2000	-0.0040	0.0410	-0.0025	0.0537
	4000	-0.0006	0.0294	0.0022	0.0379

Simulation: Evaluating estimators in the context of ASD

Simulation setting: a data generating process that mimics the real data

- n = 400
- First, use real data to estimate functional connectivity between default mode network (seed region) and 6 resting-state networks defined using Yeo 7 parcellation
- A, M, X, Z: similar in distribution to those in the observed data
- $\Delta = 1$ if $M \leq 0.2$
- Y's follow a multivariate normal distribution
- true associations between 4 regions are set to 0, the remaining 2 regions are assigned non-zero associations

Simulation: Evaluating estimators in the context of ASD

	True association		Proposed Method	Proposed Method (cross-fitting)	naïve with participant removal	naïve
Region 1	0	Bias	0.0045	**0.0041**	-0.0187	-0.0644
		sd	0.0195	0.0196	**0.0191**	0.0208
		MSE$\times 10^3$	**0.4016**	0.4020	0.7145	4.5784
		Type I error	0.0390	**0.0150**	0.1040	0.8660
Region 2	0	Bias	0.0072	**0.0063**	0.0177	0.0608
		sd	0.0299	0.0249	0.0240	**0.0224**
		MSE$\times 10^3$	0.9466	**0.6569**	0.8902	4.2036
		Type I error	0.0320	**0.0110**	0.0830	0.7180
Region 3	0	Bias	0.0075	**0.0069**	0.0156	0.0553
		sd	0.0187	0.0185	**0.0180**	0.0179
		MSE$\times 10^3$	0.4047	**0.3888**	0.5680	3.3807
		Type I error	0.0460	**0.0170**	0.0720	0.8310
Region 4	0	Bias	-0.0036	**-0.0034**	-0.0178	-0.0662
		sd	**0.0192**	0.0198	0.0196	0.0204
		MSE$\times 10^3$	**0.3824**	0.4031	0.7010	4.7970
		Type I error	0.0200	**0.0050**	0.1080	0.8840
Region 5	-0.0484	Bias	0.0068	**0.0042**	0.0215	0.0694
		sd	**0.0202**	0.0212	0.0205	0.0211
		MSE$\times 10^3$	**0.4532**	0.4673	0.8837	5.2665
		Power	**0.4550**	0.4100	0.1670	0.1190
Region 6	-0.0682	Bias	0.0056	**0.0031**	0.0245	0.0798
		sd	**0.0163**	0.0172	0.0179	0.0214
		MSE$\times 10^3$	**0.2976**	0.3059	0.9195	6.8226
		Power	**0.9380**	0.8860	0.5230	0.0630