Age-related network efficiency and the role of cognitive reserve

Seonjoo Lee
Department of Biostatistics, Columbia University/
Mental Health Data Science Biostatistics Group
New York Psychiatric Institute

Motivating Study : Cognitive Reserve (PI: Yaakov Stern)

- “The concept of **cognitive reserve** provides an explanation for differences between individuals in susceptibility to age-related brain changes or pathology (e.g. Alzheimer's disease), whereby some people can tolerate more of these changes than others and maintain function.”
- Lifetime exposures including educational and occupational attainment, and leisure activities in late life, can increase this reserve.

Stern, Yaakov. "What is cognitive reserve? Theory and research application of the reserve concept." *Journal of the International Neuropsychological Society* 8, no. 3 (2002): 448-460.
Stern, Yaakov. "Cognitive reserve in ageing and Alzheimer's disease." *The Lancet Neurology* 11, no. 11 (2012): 1006-1012.

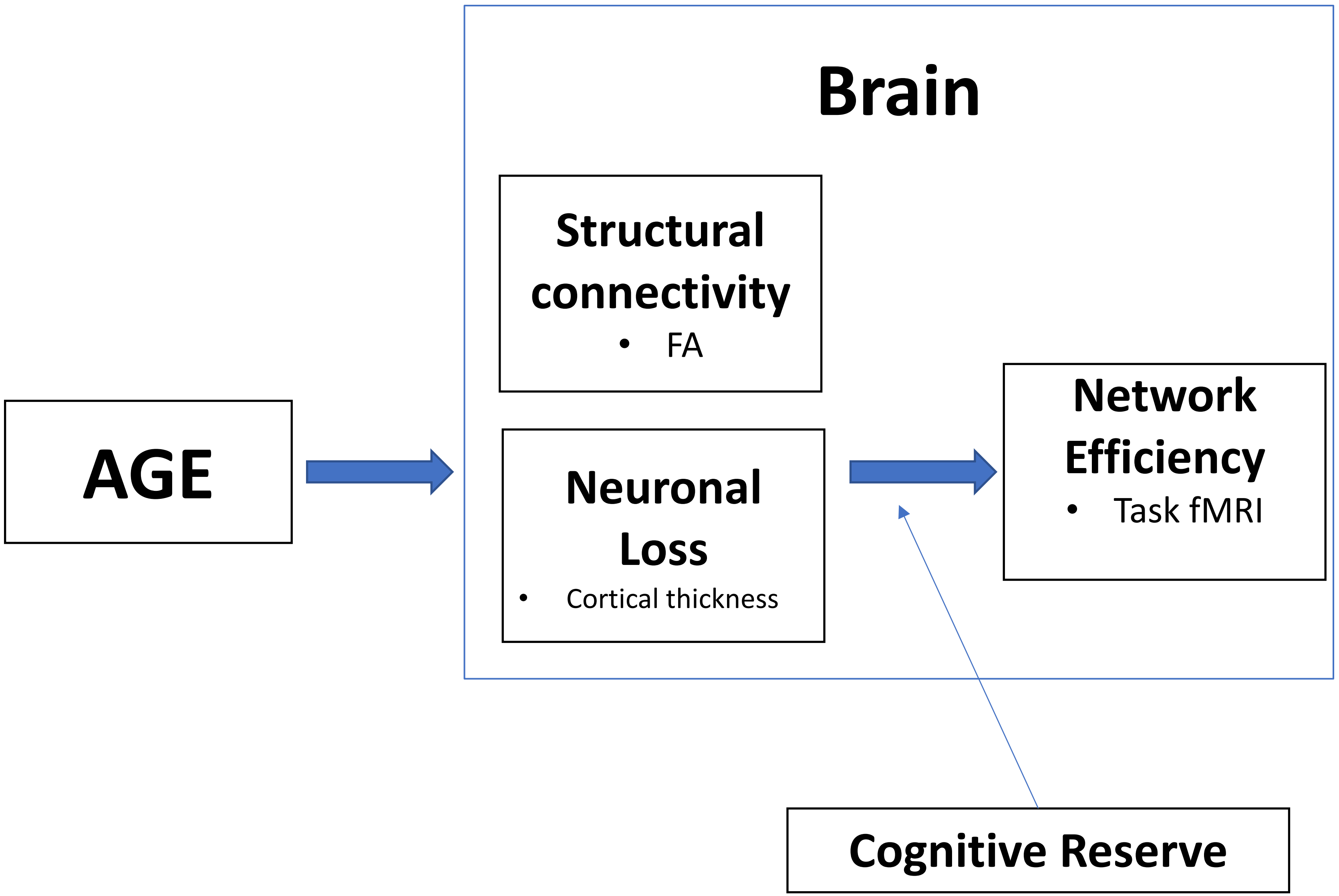
Neural Mechanisms underlying CR

- Although there is ample epidemiologic evidence for cognitive reserve, the **neural substrate of cognitive reserve have not been studied well.**
- The neural implementation of CR might take two forms:
  - **neural reserve:** inter-individual variability in the primary brain networks or cognitive paradigms that underlie the performance of any task (network efficiency or flexibility)
  - **neural compensation**: individuals suffering from brain pathology use brain structures or networks that are not normally used by individuals with intact brains in order to compensate for brain damage

Stern Y. Cognitive reserve. Neuropsychologia. 2009;47:2015–2028.
Stern Y, Habeck C, Moeller J, Scarmeas N, Anderson KE, Hilton HJ, Flynn J, Sackeim H, Van Heertum R. Brain networks associated with cognitive reserve in healthy young and old adults. Cereb.Cortex. 2005;15:394–402.

Neural Reserve (Network Efficiency)

- An individual whose networks are more efficient, have greater capacity, might be more capable of coping with the disruption imposed by aging or brain pathology.
- Efficiency refers to the change in neural activity occurring with a change in task demand. For an equal increase in task demand, someone with greater efficiency requires less of an increase in neural activity than does someone with less efficiency.
- It is possible that CR alters the network efficiency.


AGE
Brain
Structural connectivity
• FA
Neuronal Loss
• Cortical thickness
Network Efficiency
• Task fMRI
Cognitive Reserve

Primary Goals

- To better understand age-related neural reserve decline and the role of cognitive reserve
- To test whether age-related structural changes (neuronal loss/declined connectivity) affect network efficiency, and how they interact with cognitive reserve

Measures

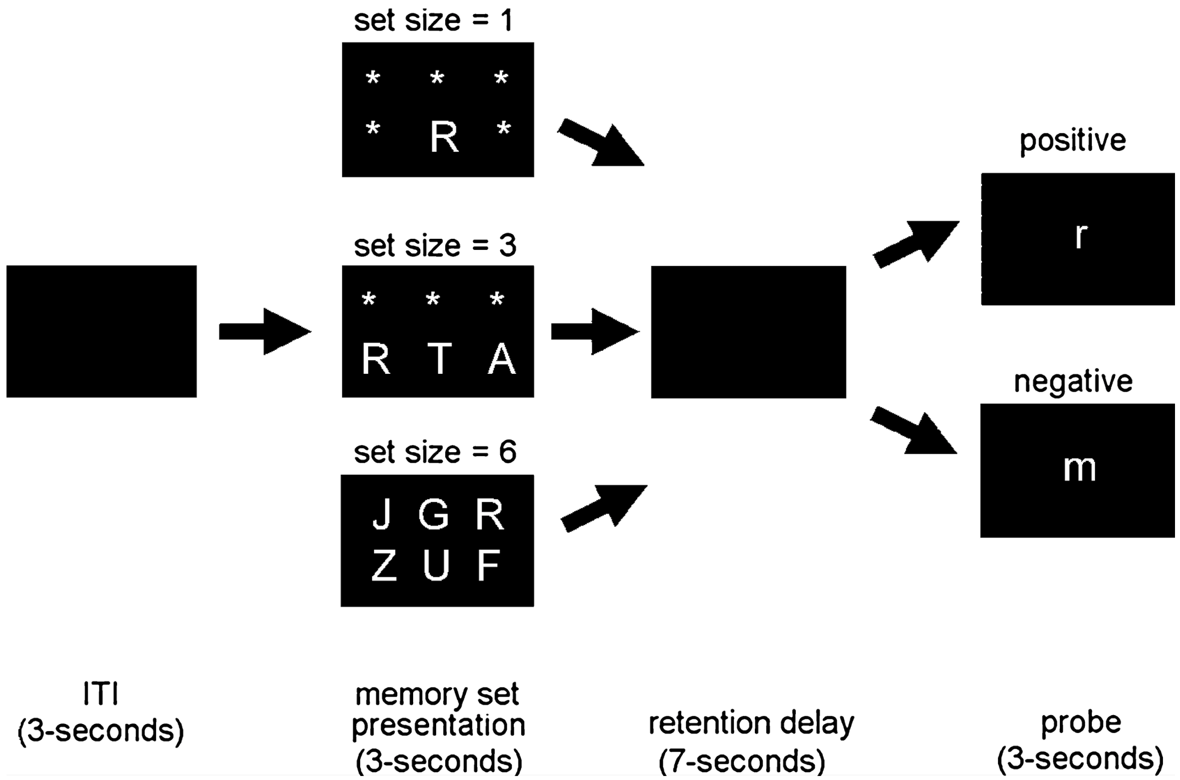
- Structural connectivity (*fractional* anisotropy *map)*
  - **Tract-Based Spatial Statistics (TBSS)** pipeline to compute FA
  - Masked in the major white matter tracks and white matter regions identified in the FSL JHU atlas.
- Neuronal loss – cortical thickness
  - **Freesurfer** pipeline
  - Vertex-level cortical thickness
- Network efficiency
  - Letter Sternberg Task + Ordinal Trend Canonical Variates Analysis
- Cognitive Reserve
  - IQ

Cognitive reserve data analysis

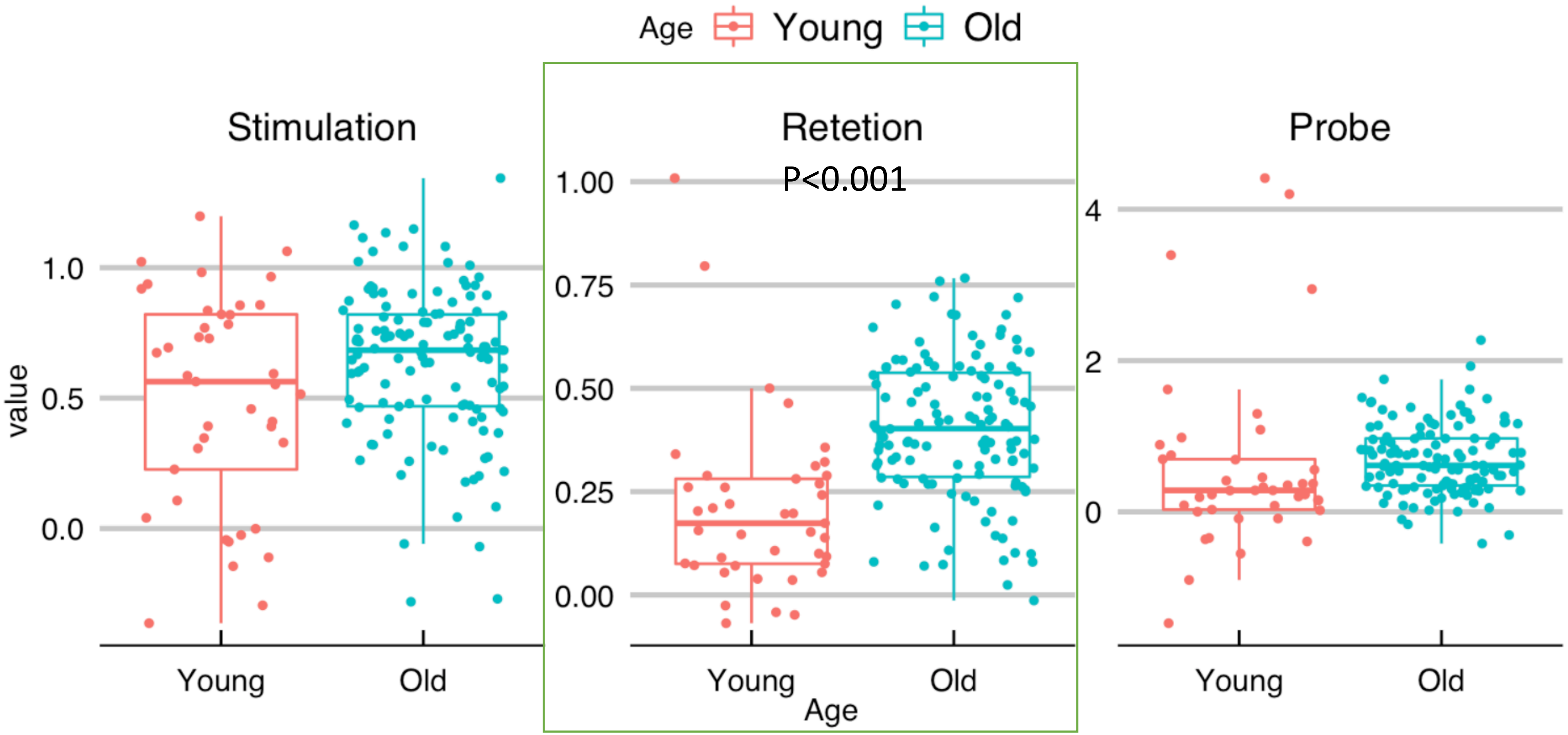
- N=167, 70 (42%) Male
- Mostly White (65.90%)
- 40 young (20-30 yrs) and 127 old (60-70 yrs).
- Multimodal imaging (T1, DTI)

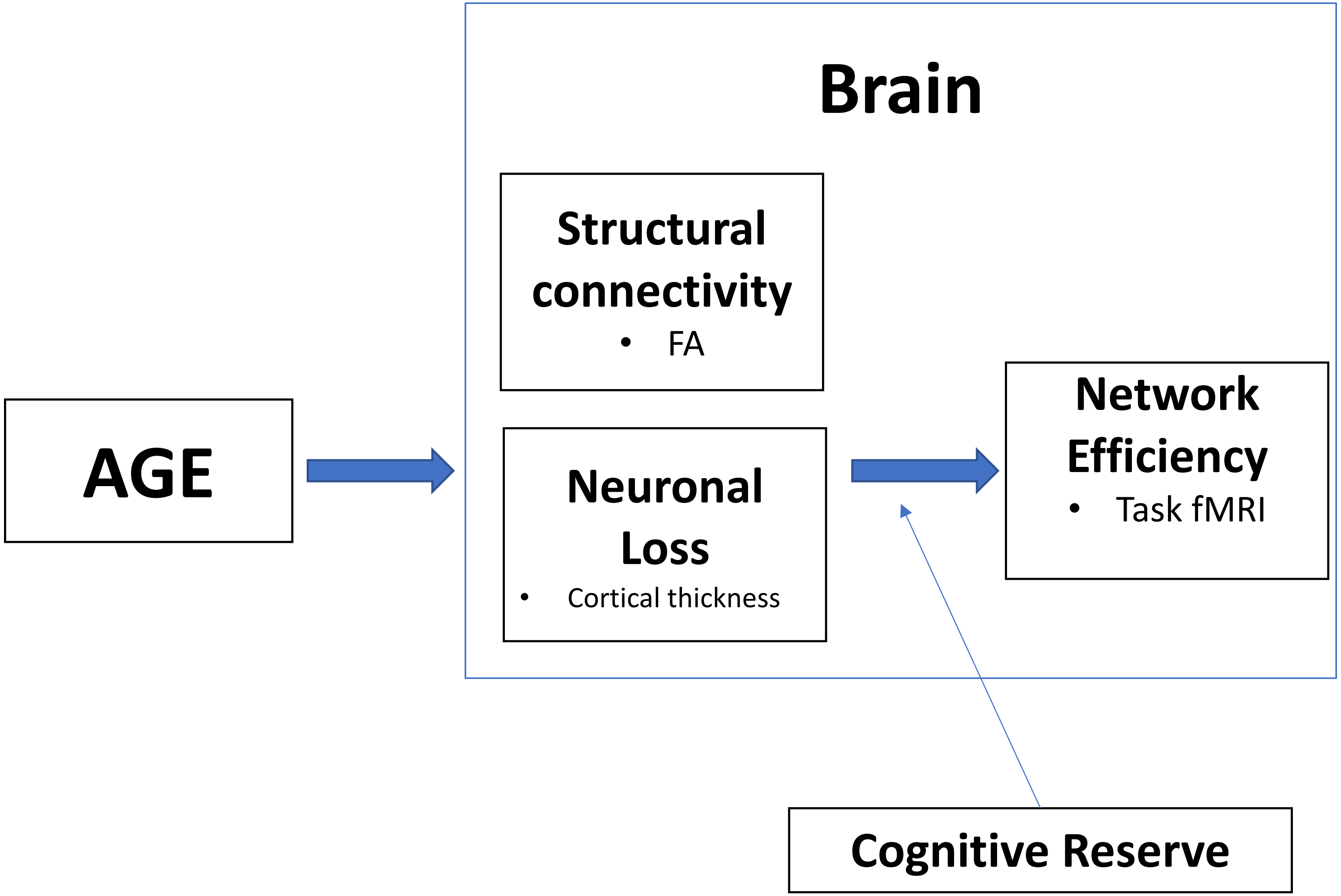
Letter Sternberg Task

- all participants performed a verbal delayed item recognition task in which memory load was manipulated by the number of letters (1, 3, or 6) the subject needed to store in working memory.
- First level analysis: the time series models crossed the load (1, 3 or 6 letters) and task phase (stimulus, retention and recognition) factors to create nine regressors of interest (9 contrast maps).
- Identify patterns of networks which decreases from letter1 to letter 6 using ordinal trend canonical variates analysis (Habeck et al. 2003; Habeck et al. 2005)
- Project data to the identified directions and compute slopes of the score.
- Lower slope indicates higher efficiency.



Total effect (Network Efficiency Measures)




Brain
Structural connectivity
• FA
AGE
Neuronal Loss
• Cortical thickness
Network Efficiency
• Task fMRI
Cognitive Reserve

Challenges

- Multimodal imaging data – age related change in brain is measured via multimodal imaging to study various aspect of brain.
- The dimension of the mediator is very high, and also highly correlated.
- Dimension reduction + feature extraction at the same time.
- We want to find components that are linked each other to characterize underlying brain features.

Analysis Strategy

- Dimension reduction for multimodal structural brain imaging data (FA/Cortical Thickness) ➔ Linked ICA
- Develop high-dimensional moderated mediation analysis to test moderation mediation

Linked ICA

- Flexible to handle different units, signal- and contrast-to-noise ratios, dimensions, spatial smoothness and intensity distribution.

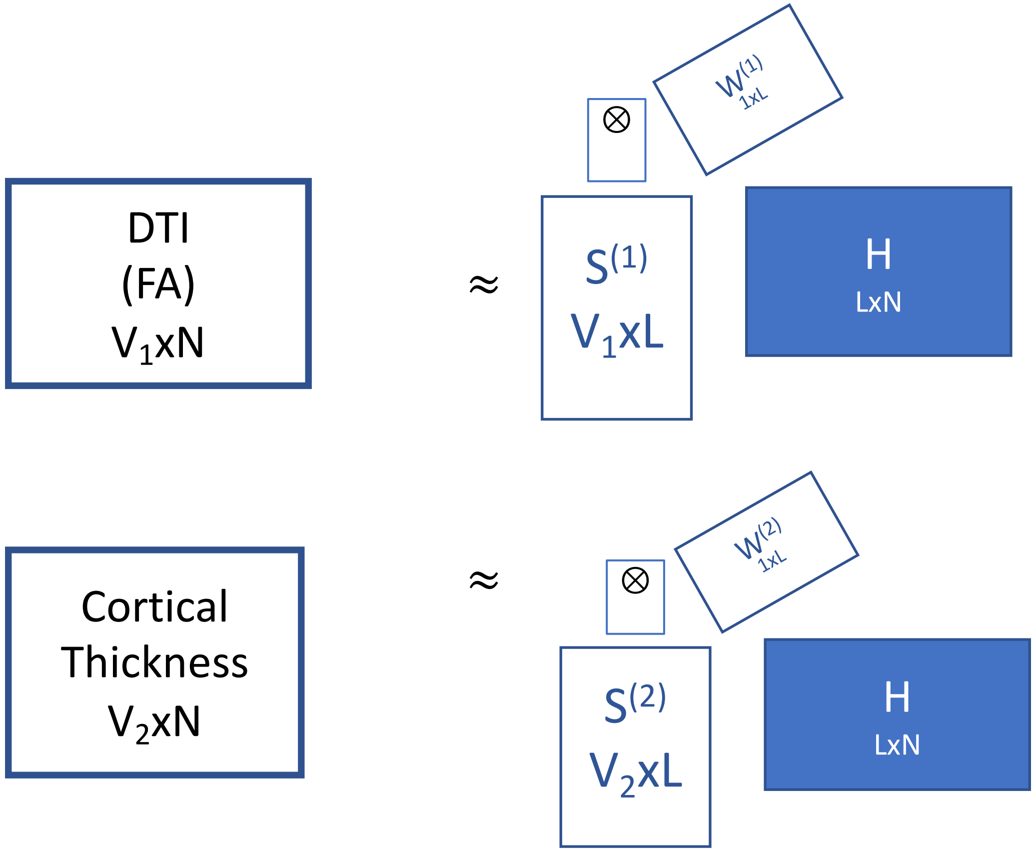
$$X_{v,t,i}^{(k)} = \sum_{l=1}^{L} S_{v,l}^{(k)} W_{t,l}^{(k)} H_{l,i} + E_{v,t,i}^{(k)},$$

  - $X_{v,t,i}^{(k)}$: the data in modality group $k = 1, \dots, K$, modality $t = 1, \dots, T_k$, subject $i = 1, .., N$ and voxel $v = 1, .., V_k$ when the modality group is neuroimaging
  - $S_{v,l}^{(k)}$: the spatial maps for component l in modality group k, voxel v
  - $W_{t,l}^{(k)}$: the modality weightings for component l in modality t of modality group k
  - $H_{l,i}$: the weights for component l in subject i (IC scores)
- In estimation, LICA allows $W_{t,l}^{(k)}$ to be zero, thus we can flexibly identify components from only one or few modalities.

Linked independent component analysis for multimodal data fusion. Neuroimage, Feb 2011, 54(3):2198-2217

Multimodal Imaging Fusion via Linked ICA

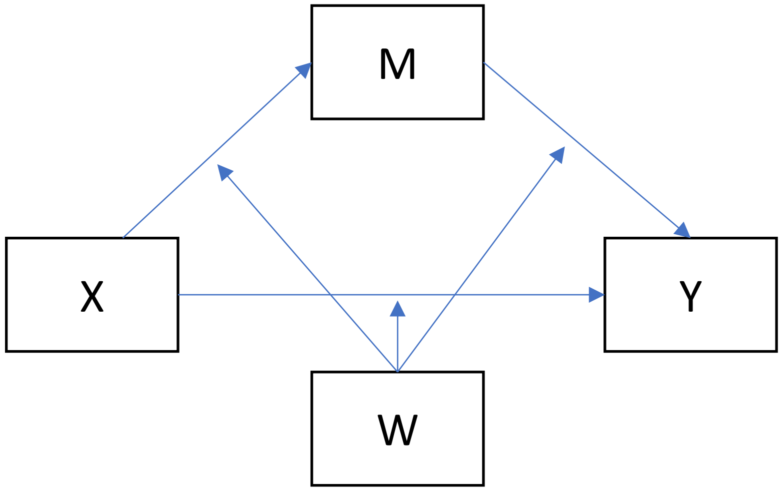
- 40 ICs were extracted.



Linked independent component analysis for multimodal data fusion. Neuroimage, Feb 2011, 54(3):2198-2217

Moderated Mediation Analysis

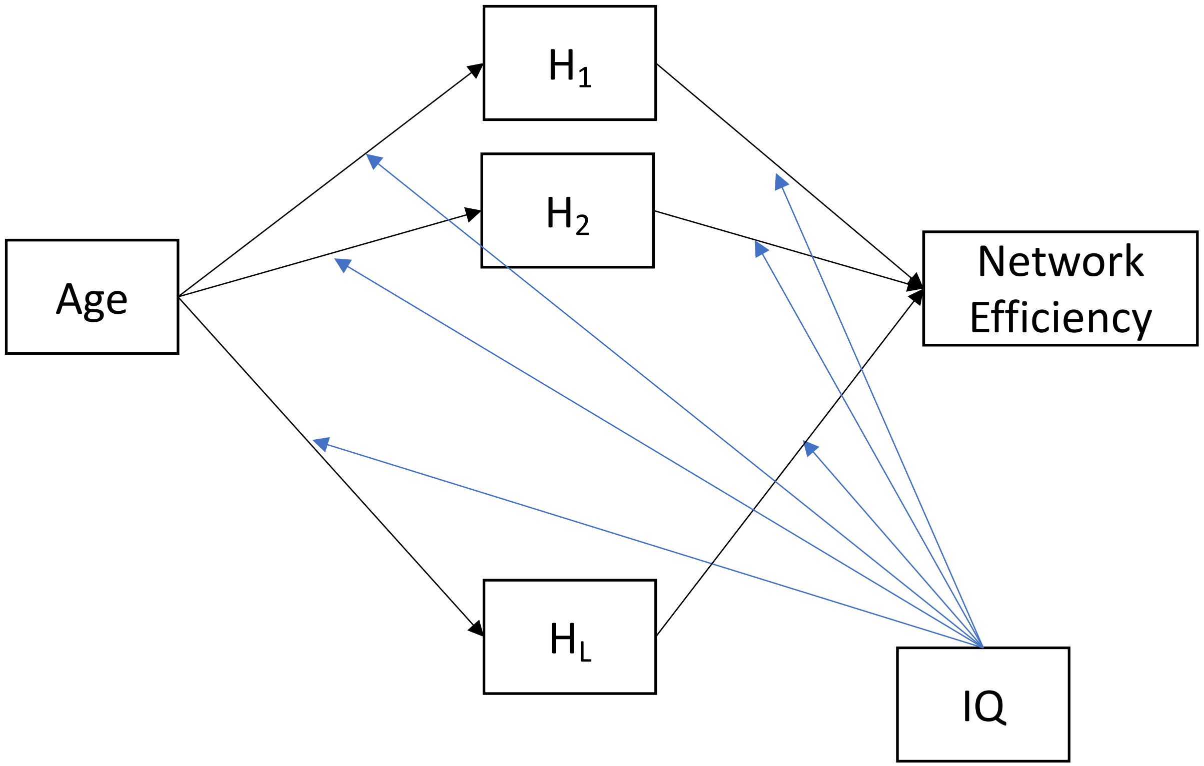
- What is moderated mediation (**conditional indirect effects)?**
- The effect of an independent variable X on an outcome variable Y via a mediator variable M differs depending on levels of a moderator variable W. Specifically, either the effect of X on the M, and/or the effect of M on Y depends on the level of W.



Hayes, Andrew F. "An index and test of linear moderated mediation." *Multivariate Behavioral Research* 50, no. 1 (2015): 1-22.

Moderated Mediation

- High-dimensional multivariate mediator (CR example):



Moderated Mediation

- $H_{l,i} = \alpha_{0l} + \alpha_{1l} Age_i + \alpha_{2l} Age_i * IQ_i + \varepsilon_i$
- $NE_i = \gamma_{0l} + \gamma_{1l} Age_i + \gamma_{2l} Age_i * IQ_i + \sum_{l=1}^{L} \beta_{1l} H_{l,i} + \sum_{l=1}^{L} \beta_{2l} H_{l,i} * IQ_i + E_{v,t,i}^{(k)}$
- Moderated mediation is quantified as

$$(\alpha_{1l} + \alpha_{2l} * IQ)(\beta_{1l} + \beta_{2l} * IQ)$$

- In fact, in the cognitive reserve theory, we don’t believe age to structural brain is moderated by cognitive reserve. Also, we did not find any evidence.

Moderated Mediation

- $H_{l,i} = \alpha_{0l} + \alpha_{1l}Age_i + \alpha_{2l}Age_i * IQ_i + \varepsilon_i$
- $NE_i = \gamma_{0l} + \gamma_{1l}Age_i + \gamma_{2l}Age_i * IQ_i + \sum_{l=1}^{L} \beta_{1l}H_{l,i} + \sum_{l=1}^{L} \beta_{2l}H_{l,i} * IQ_i + E_{v,t,i}^{(k)}$
- Moderated mediation is quantified as

$$(\alpha_{1l} + \alpha_{2l} * IQ)(\beta_{1l} + \beta_{2l} * IQ)$$

- In fact, in the cognitive reserve theory, we don't believe age to structural brain is moderated by cognitive reserve. Also, we did not find any evidence.

Moderated Mediation

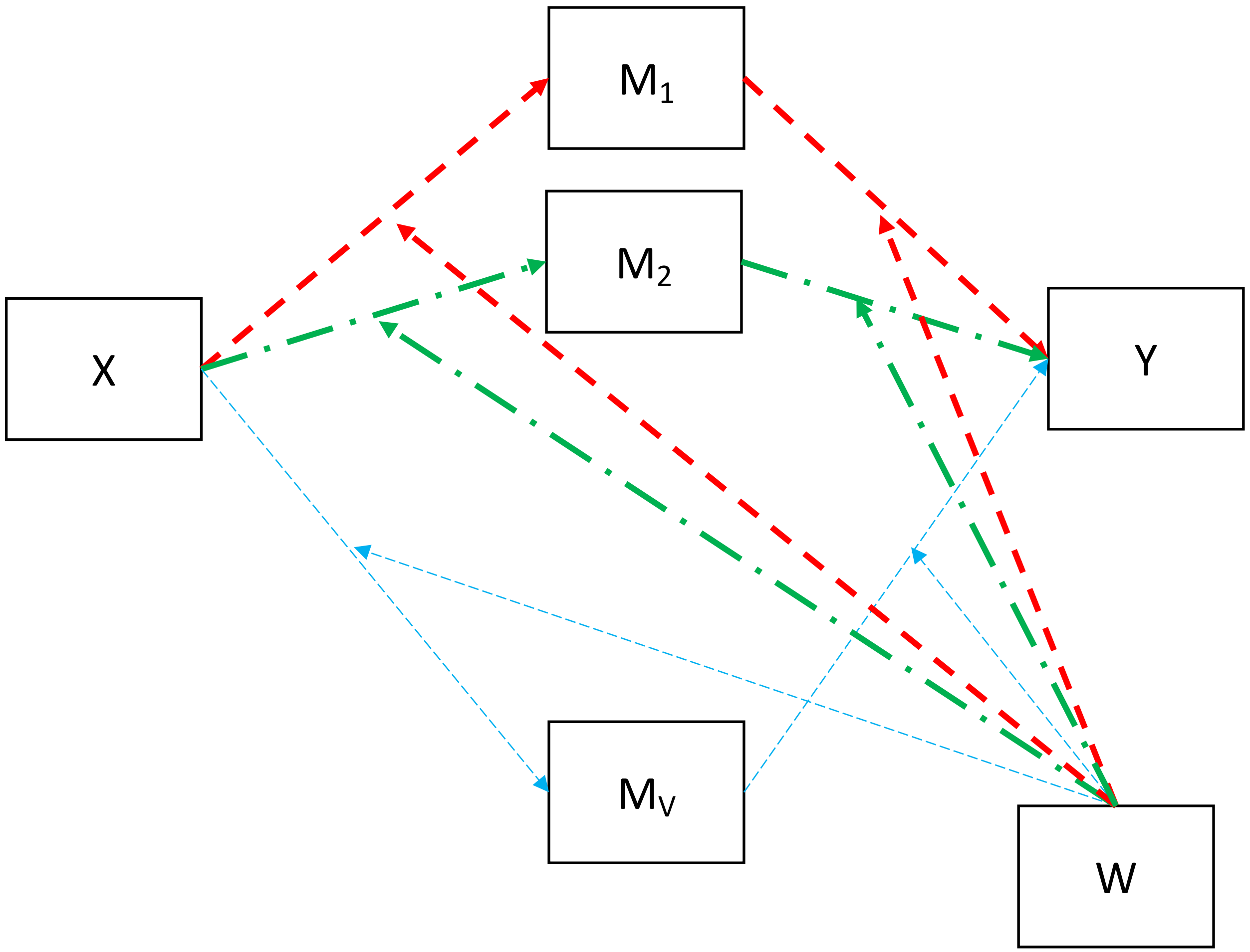
- $H_{l,i} = \alpha_{0l} + \alpha_{1l} Age_i + \varepsilon_i$
- $NE_i = \gamma_{0l} + \gamma_{1l} Age_i + \gamma_{2l} Age_i * IQ_i + \sum_{l=1}^{L} \beta_{1l} H_{l,i} + \sum_{l=1}^{L} \beta_{2l} H_{l,i} * IQ_i + E_{v,t,i}^{(k)}$
- Moderated mediation is quantified as

$$(\alpha_{1l})(\beta_{1l} + \beta_{2l} * IQ)$$

- In fact, in the cognitive reserve theory, we don't believe age to structural brain is moderated by cognitive reserve. Also, we did not find any evidence.

High-dimensional multivariate mediator

- We want to select mediators that have both paths and keep the hierarchical structure between main effects and interactions, i.e. if the interaction terms are non-zero, their main effects have to be in the model. ➔ Group LASSO penalization



Model

- Consider n-dimensional vectors, X (predictor), W(moderator) and n-by-V dimensional matrix M (moderator).
- The moderated mediation model is give as

$$\mathbf{M} = \mathbf{X}\boldsymbol{\alpha}_1^\top + (XW)\boldsymbol{\alpha}_2^\top + \mathbf{W}\boldsymbol{\alpha}_3^\top + \boldsymbol{\epsilon}_M \tag{1}$$

$$\mathbf{Y} = \mathbf{X}c_1 + \mathbf{XW}c_2 + \mathbf{M}\boldsymbol{\beta}_1 + \mathbf{MW}\boldsymbol{\beta}_2 + \epsilon_y. \tag{2}$$

$$M|X = x, W = w \sim MVN(\mathbf{x}\boldsymbol{\alpha}_1^\top + (xw)\boldsymbol{\alpha}_2^\top + \mathbf{w}\boldsymbol{\alpha}_3^\top, \boldsymbol{\Sigma}_M)$$

$$Y|M = m, X = x, W = w \sim N(\mathbf{X}c_1 + \mathbf{XW}c_2 + \mathbf{M}\boldsymbol{\beta}_1 + \mathbf{MW}\boldsymbol{\beta}_2, \sigma_Y^2)$$

Estimation

- -2*loglikelihood:

$$\frac{1}{\sigma_Y^2}\left(\mathbf{Y}-\mathbf{X}c_1+\mathbf{XW}c_2+\mathbf{M}\boldsymbol{\beta}_1+\mathbf{MW}\boldsymbol{\beta}_2\right)^\top\left(\mathbf{Y}-\mathbf{X}c_1+\mathbf{XW}c_2+\mathbf{M}\boldsymbol{\beta}_1+\mathbf{MW}\boldsymbol{\beta}_2\right)$$
$$+\,tr\left(\left(\mathbf{M}-\mathbf{X}\boldsymbol{\alpha}_1^\top-(XW)\boldsymbol{\alpha}_2^\top-\mathbf{W}\boldsymbol{\alpha}_3^\top\right)\boldsymbol{\Sigma}_M^{-1}\left(\mathbf{M}-\mathbf{X}\boldsymbol{\alpha}_1^\top-\mathbf{XW}\boldsymbol{\alpha}_2^\top-\mathbf{W}\boldsymbol{\alpha}_3^\top\right)^\top\right)$$
$$+\,n\log|\boldsymbol{\Sigma}_M|+n\log\sigma_Y^2+constant. \quad (3)$$

- Minimize

-2*loglikelihood $+\lambda\sum_{v=1}^{V}||\gamma_k||_2$,

where γ_k=(α_{1k}, α_{2k}, β_{1k}, β_{2k}).

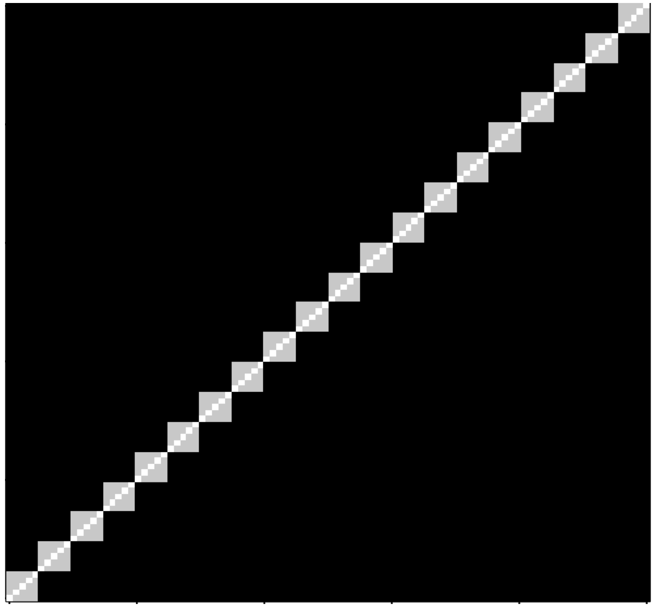
Algorithm

Table 1: **Algorithm: Sparse Moderated Mediation**

Input:	$\mathbf{X}$,$\mathbf{M}$,$\mathbf{Y}$,$\mathbf{W}$, λ
Output:	Parameter estimates of c_1,c_2, c_3, $\boldsymbol{\alpha}_1$,$\boldsymbol{\alpha}_1$,$\boldsymbol{\alpha}_1$,$\boldsymbol{\beta}_1$,$\boldsymbol{\beta}_2$
Step 1.	Initialization:
	Initialize parameters : c_1,c_2, c_3, $\boldsymbol{\alpha}_1$,$\boldsymbol{\alpha}_1$,$\boldsymbol{\alpha}_1$,$\boldsymbol{\beta}_1$,$\boldsymbol{\beta}_2$
	$\hat{\sigma}_y^2 = \frac{1}{n}\left(\mathbf{Y} - \mathbf{U}\boldsymbol{\theta}\right)^\top \left(\mathbf{Y} - \mathbf{U}\boldsymbol{\theta}\right)$,
	where $\mathbf{U} = (\mathbf{X}, (\mathbf{X} * \mathbf{W}), \mathbf{W}, \mathbf{M}, (\mathbf{M} * \mathbf{W}))$
	and $\boldsymbol{\theta} = (c_1, c_2, c_3, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2)^\top$
	$\widehat{\boldsymbol{\Sigma}}_M = \frac{1}{n}(\mathbf{M} - \mathbf{R}\boldsymbol{\theta}_2)^\top(\mathbf{M} - \mathbf{R}\boldsymbol{\theta}_2)$,
	where $\mathbf{R} = (\mathbf{X}, (\mathbf{X} * \mathbf{W}), \mathbf{W})$ and $\boldsymbol{\theta}_2 = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3)$
Step 2.	Regularization: For each λ and τ, iterate until convergence.
	Minimize $\lVert\mathbf{A}^{1/2}\boldsymbol{\gamma} - \mathbf{A}^{-1/2}\mathbf{B}\rVert^2 + \lambda \sum_{v=1}^{V} \lVert\boldsymbol{\gamma}_v\rVert_2$,
	where $\mathbf{A} = \frac{1}{\sigma_y^2}\mathbf{U}^\top\mathbf{U} \oplus \mathbf{X}^\top\mathbf{X}\boldsymbol{\Sigma}_M^{-1} \oplus (\mathbf{X} * \mathbf{W})^\top(\mathbf{X} * \mathbf{W})\boldsymbol{\Sigma}_M^{-1} \oplus \mathbf{W}^\top\mathbf{W}\boldsymbol{\Sigma}_M^{-1}$,
	$\mathbf{B} = \frac{1}{\sigma^2}\mathbf{U}^\top\mathbf{Y} \oplus \boldsymbol{\Sigma}^{-1}\mathbf{M}^\top\mathbf{X} \oplus \boldsymbol{\Sigma}^{-1}\mathbf{M}^\top(\mathbf{X} * \mathbf{W}) \oplus \boldsymbol{\Sigma}^{-1}\mathbf{M}^\top(\mathbf{W})\mathbf{M}^\top\boldsymbol{\Sigma}^{-1}\mathbf{X})$,
	and $\gamma_v = (\alpha_{1v}, \alpha_{2v}, \beta_{1v}, \beta_{2v})^\top$.
	Update $\hat{\sigma}_y^2$ and $\widehat{\boldsymbol{\Sigma}}_M$.

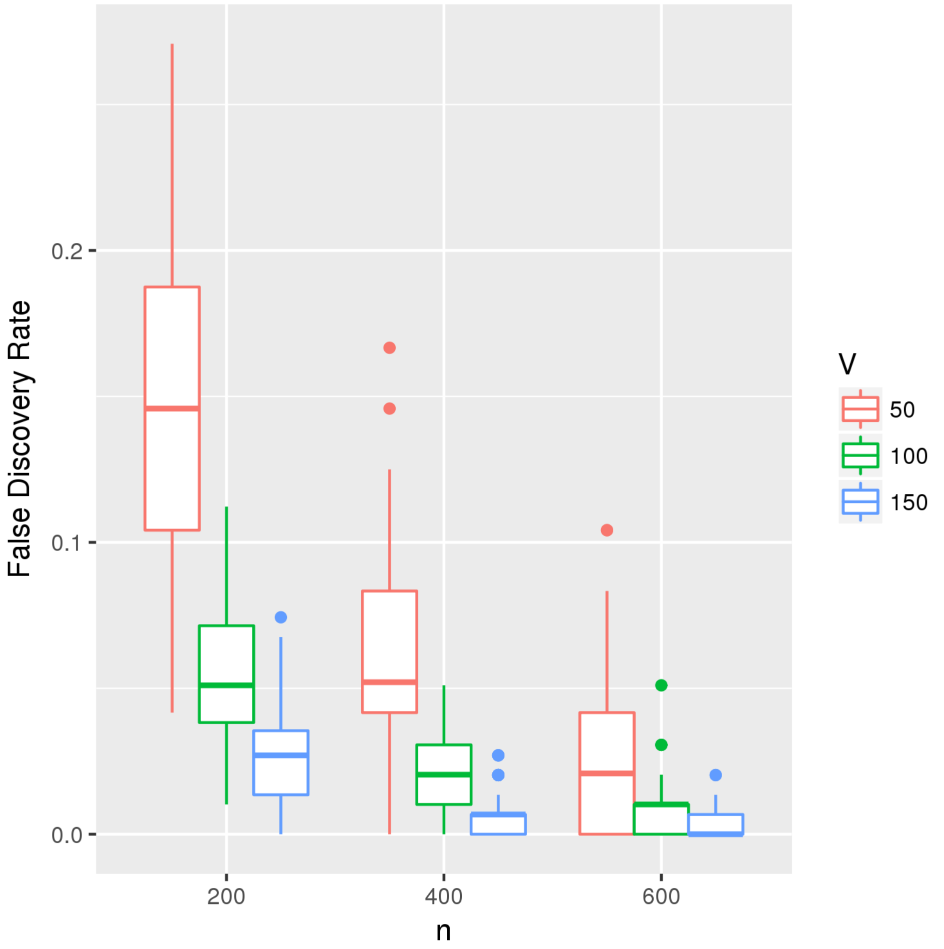
Simulation

- We considered two settings: independent vs. correlated mediators.
- $M_1 = 0.3*X + 0.3*(X*W) + \varepsilon_1$
- $M_2 = 0.3*X + 0.3*(X*W) + \varepsilon_2$
- $Y = 0.3* M_1 + 0.3* M_2 + 0.3* (M_1*W) + 0.3* (M_2*W) + \varepsilon_Y$
- Setting1: ε_1, ε_2~MVN(0, I_V), ε_Y~N(0,1)
- Setting2: ε_1, ε_2~MVN(0, Σ)
- N=200,400,600,V=50,100,150
- 100 simulation runs.

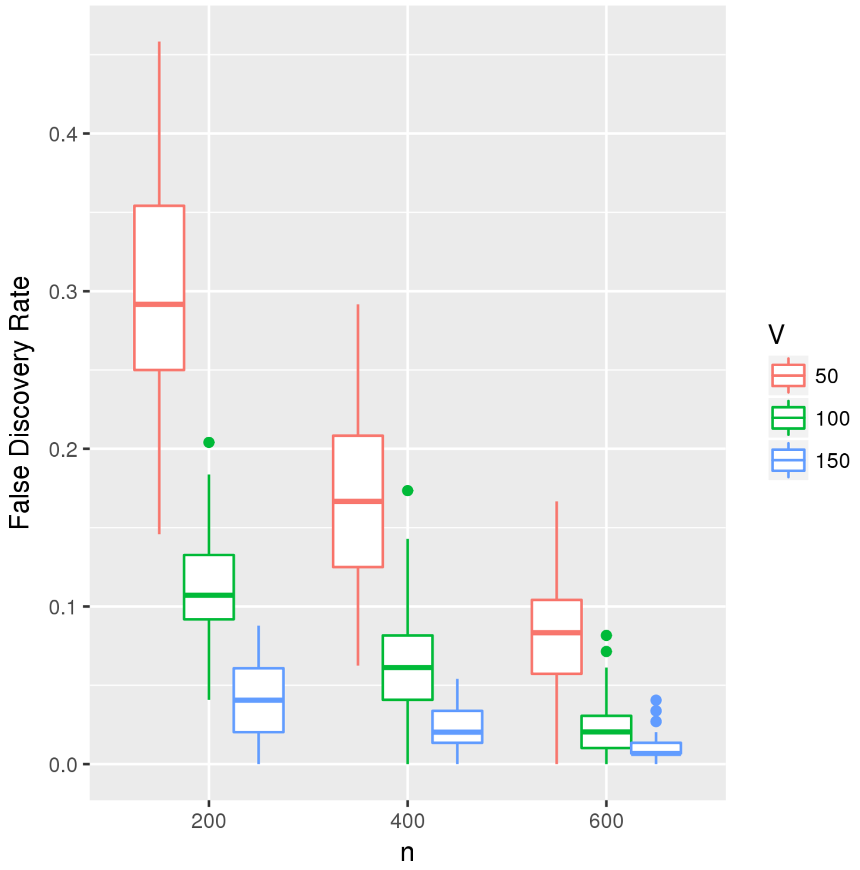


False Positive Rates for Moderation

Independent mediators

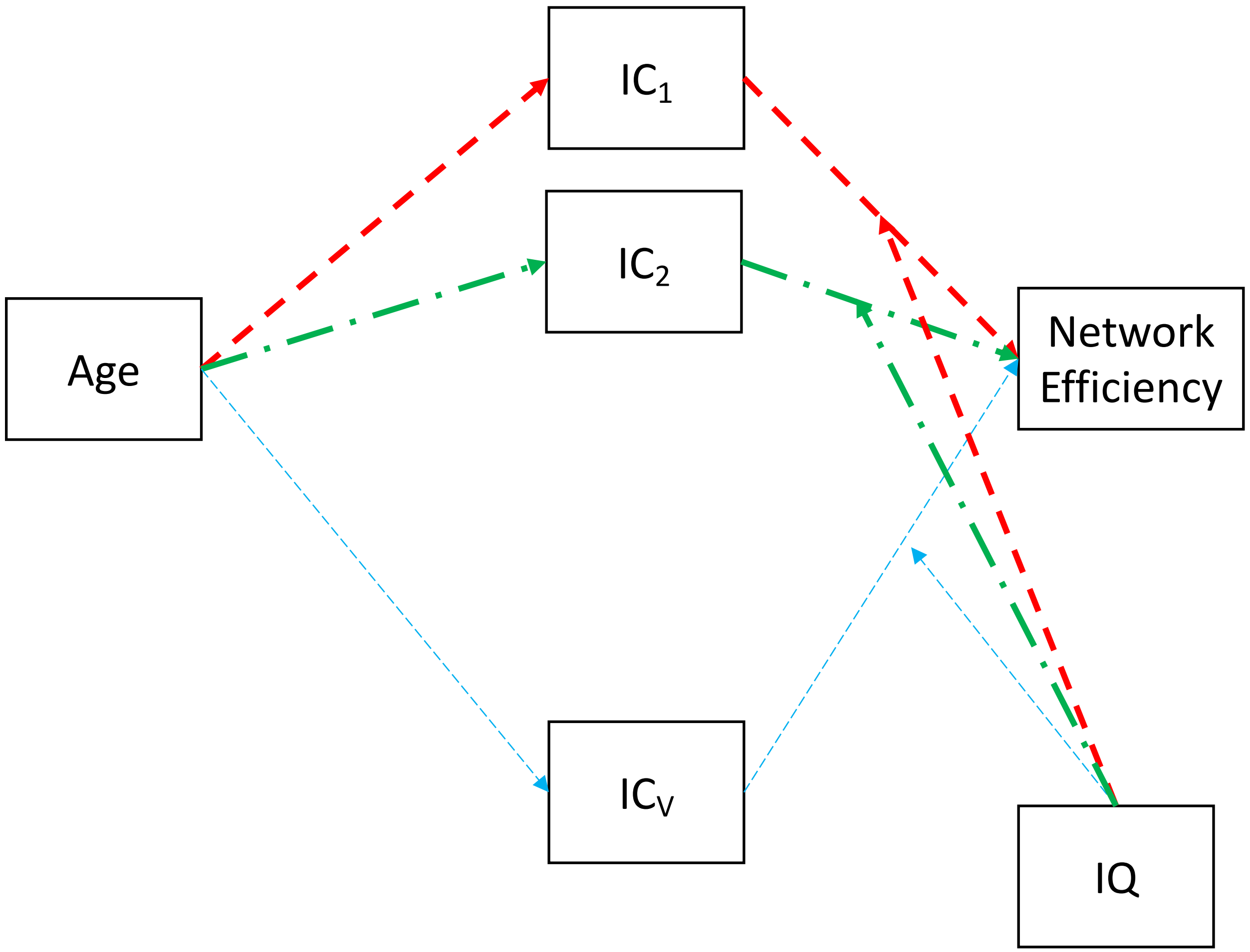


Correlated mediators

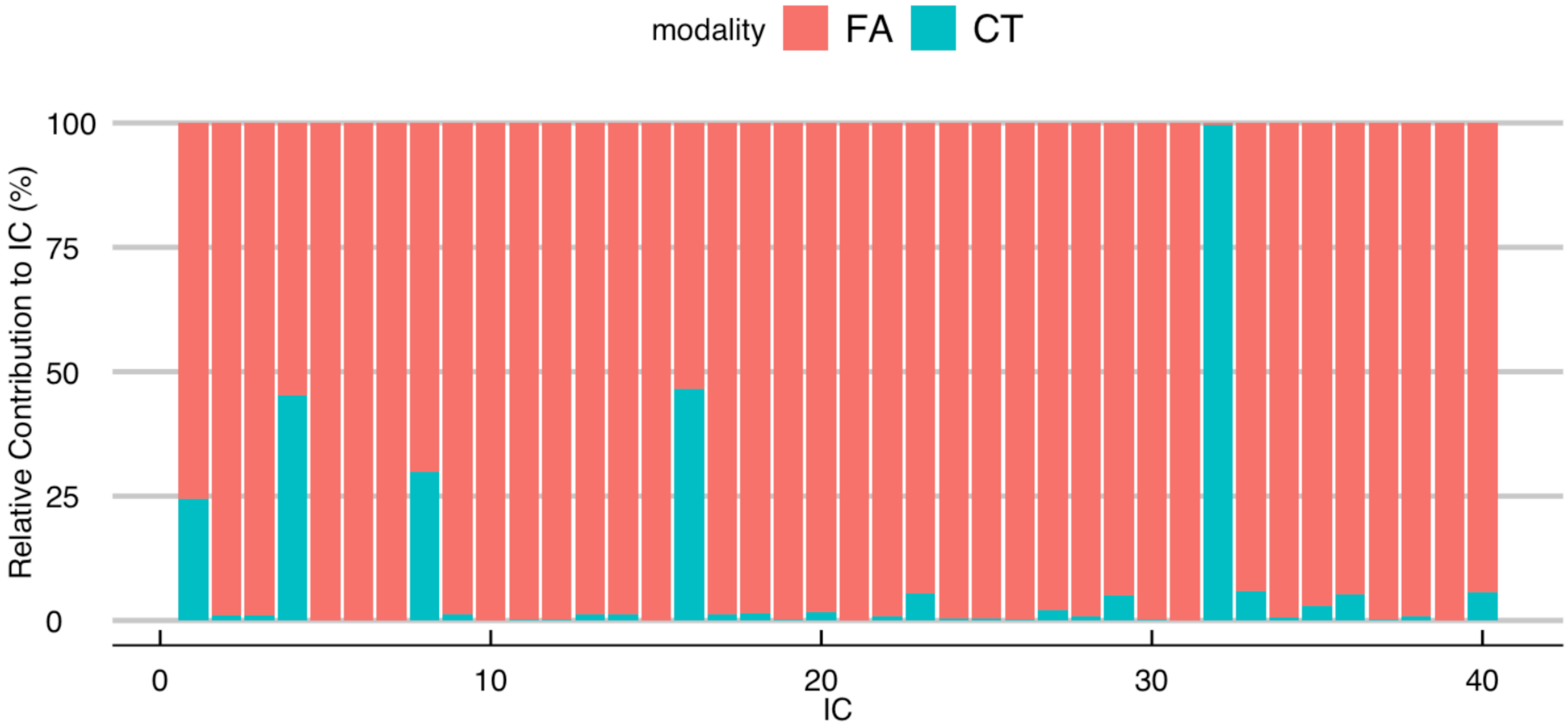


Results

- Tuning parameters were selected via 5-fold cross-validation.
- IC1,2,4,9 were selected as mediators potentially moderated by IQ.
- However, none of IQ moderation was not significantly different from zero (bootstrapping p-values >0.3).
- Network efficiency decreases in older participants(0.195, 95% CI: 0.132 to 0.258).
- Mediation of IC4: -0.032 (95% IC: -0.115 to -0.003)

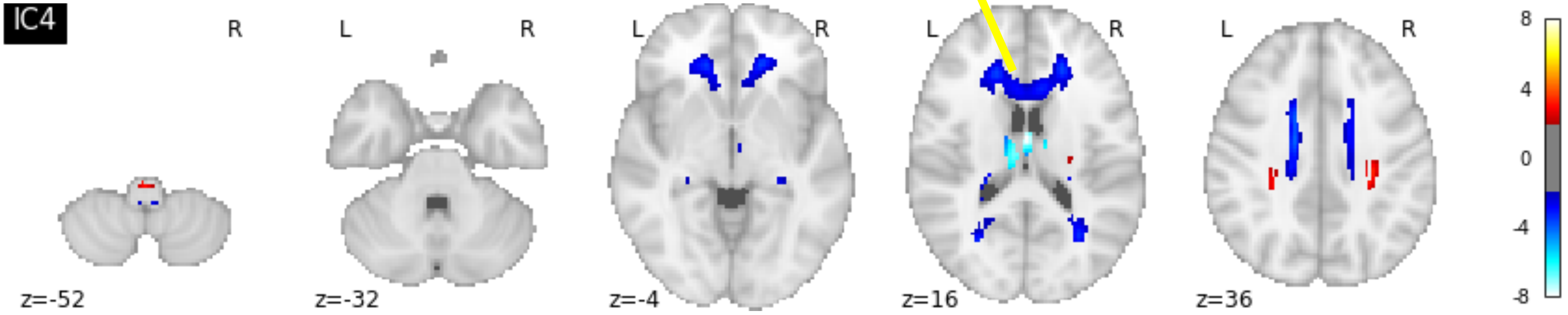


Relative Contribution of each modality

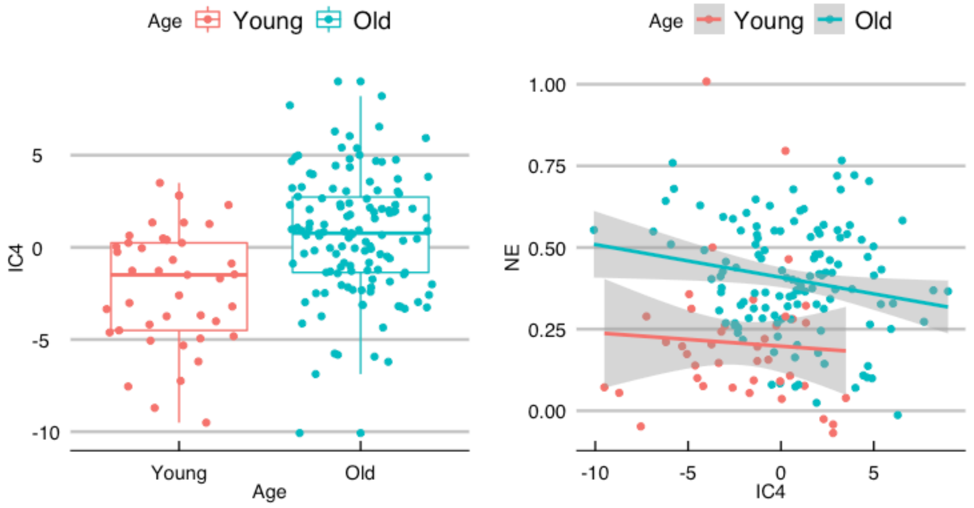


IC4

Corpus Collosum/Forceps minor

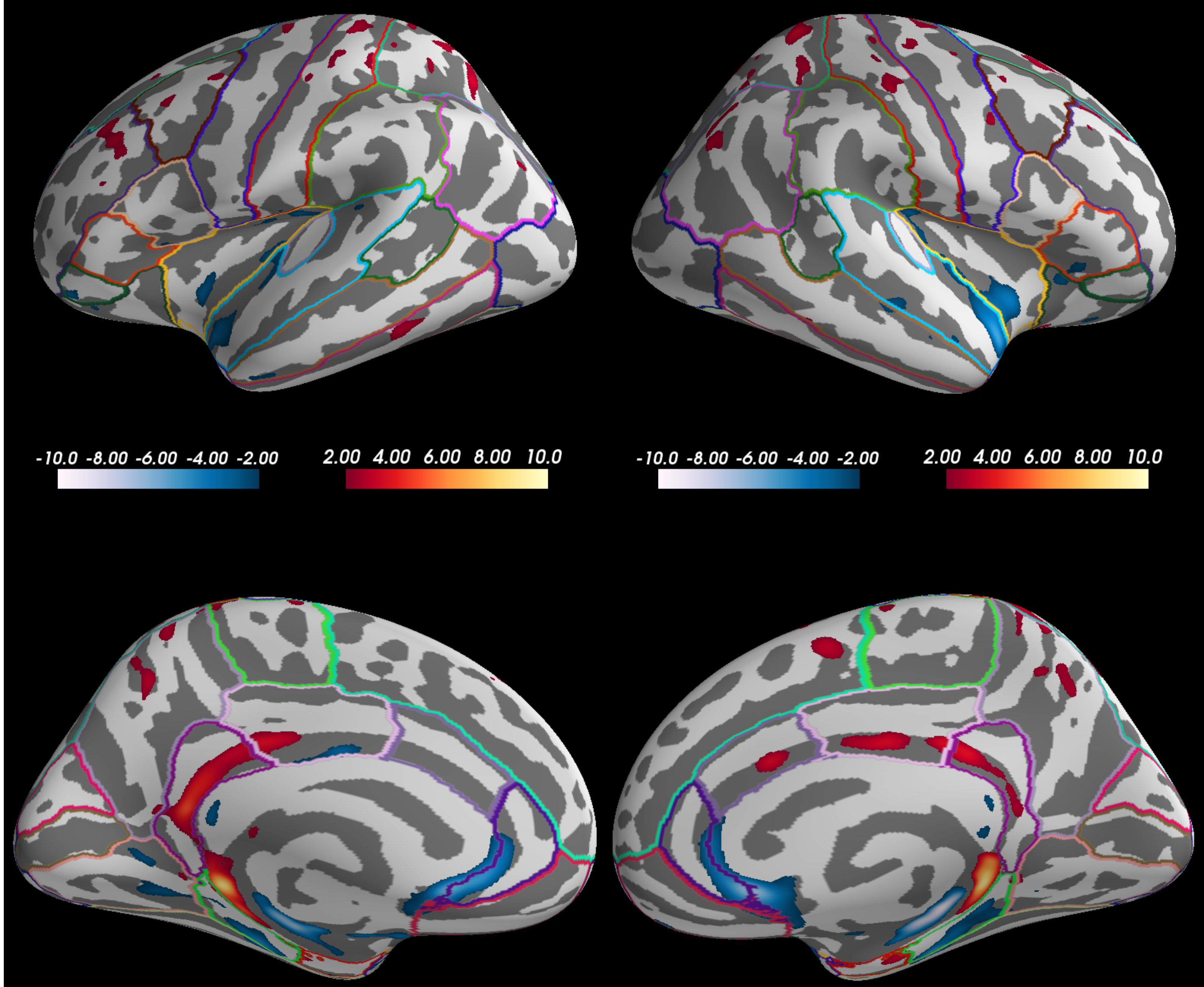


Negative Loadings in Corpus collosum, indicating decreasing structural connectivity



IC4

- Cortical thinning in bilateral rostral ACC and parahippocampal gyrus
- Thicker cortex in bilateral isthmus cingulate cortex
- Mixed pattern in subcortical Area



Future direction

- We tested neural reserve hypothesis employing dimension reduction and sparse moderated mediation analysis
- Cognitive reserve did not alter network efficiency due to aging.
- However, we perhaps are not well powered to test such hypothesis.
- Network efficiency can be quantified as the full profile of activation patterns instead of summarizing into one measure
- The same model will be evaluated in the independent datasets.
- More integrated modeling/inference framework are under development
- An R-package will be available in github.

Acknowledgement

This work is supported by NIH grant K01 AG051348.

Division of cognitive neuroscience, Columbia University

Yaakov Stern

Christian Habeck

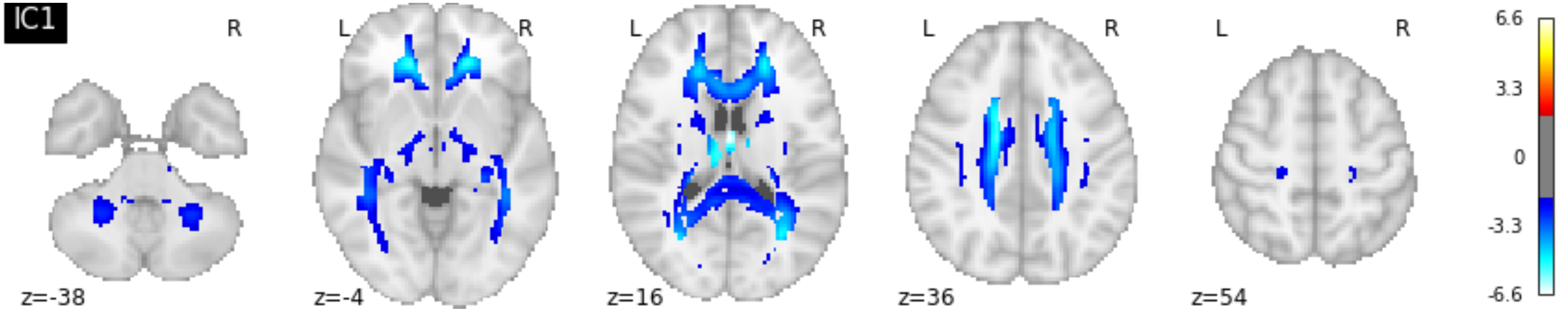
Elaine Gayze

LeeLab (https://leelab.netlify.com/)

THANK YOU

seonjoo.lee@nyspi.columbia.edu

IC1 – overall connectivity/cortical thickness



IC1 – overall connectivity/cortical thickness