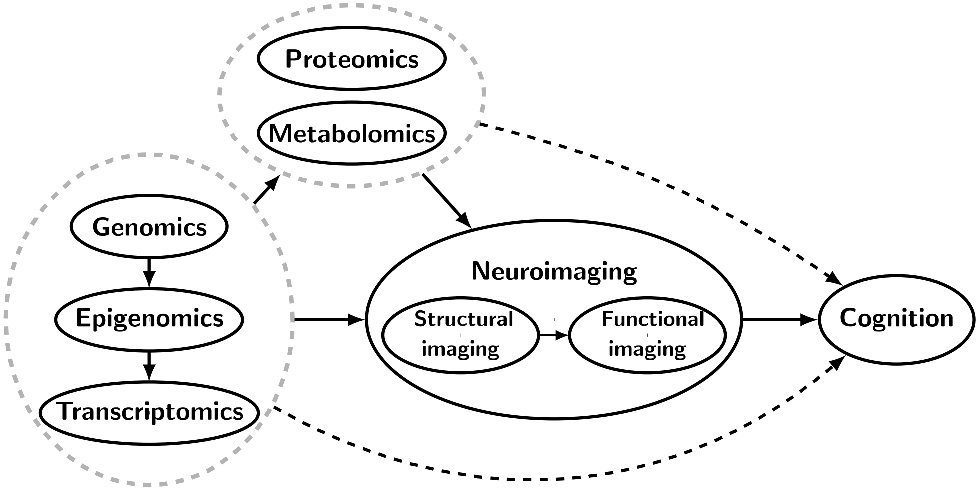
Multiview Data Integration via Mediation/Path Analysis

Yi Zhao

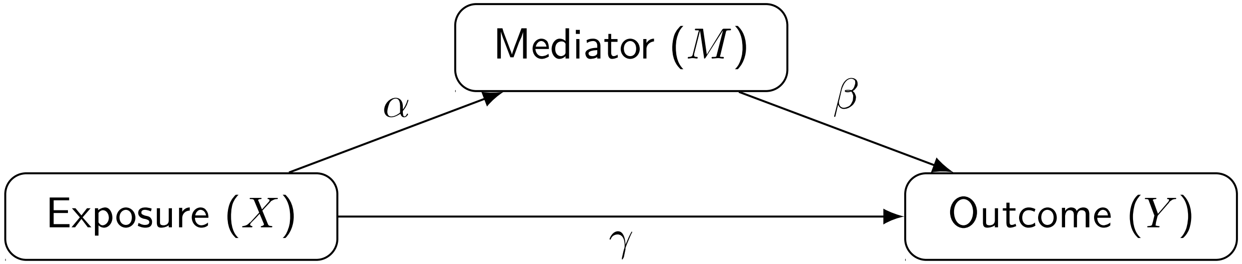
Associate Professor
Siu L. Hui Scholar in Biostatistics
Department of Biostatistics and Health Data Science
Indiana University School of Medicine

October 15, 2024



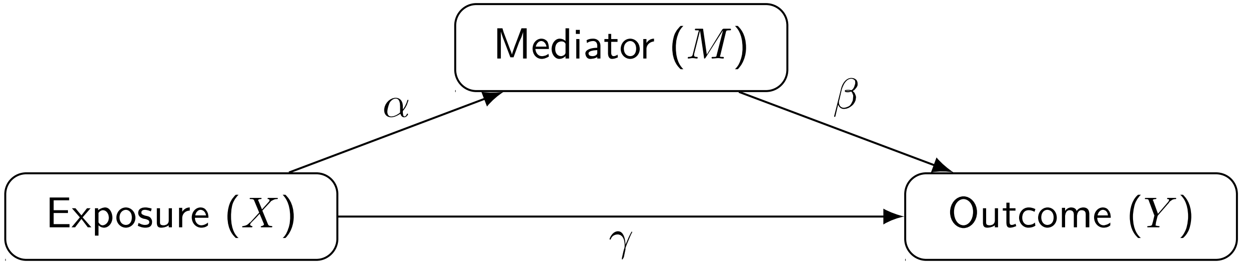
- **Hypothesis**: CSF protein → brain volume → cognition.
- **Approach**: integrate CSF proteomics and brain volumetric data through mediation analysis.
  - Exposure: CSF proteomics data ($r = 320$ peptides).
  - Mediator: brain volume of $p = 145$ ROIs.
  - Outcome: ADNI_MEM, a composite score of memory.
- **Challenges**
  - High-dimensional exposures and high-dimensional mediators.
  - Exposures/mediators are dependent.

Mediation analysis



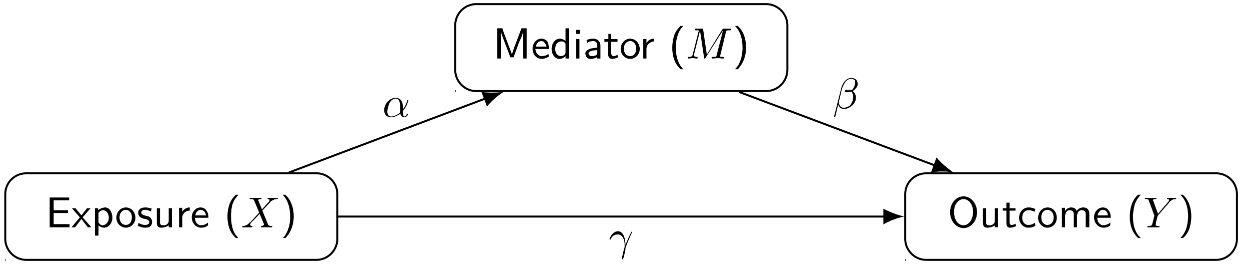
- Quantifies the intermediate effect of the mediator.

Mediation analysis



- Quantifies the intermediate effect of the mediator.
- Helps clarify the underlying causal mechanism.

Mediation analysis



- Quantifies the intermediate effect of the mediator.
- Helps clarify the underlying causal mechanism.
- Popular parametric approach: structural equation modeling (SEM)

$$M = X\alpha + \epsilon_1$$

$$Y = X\gamma + M\beta + \epsilon_2$$

  - $\alpha\beta$: indirect (mediation) effect
  - γ: direct effect

Notation

For subject i, $i = 1, \ldots, n$,

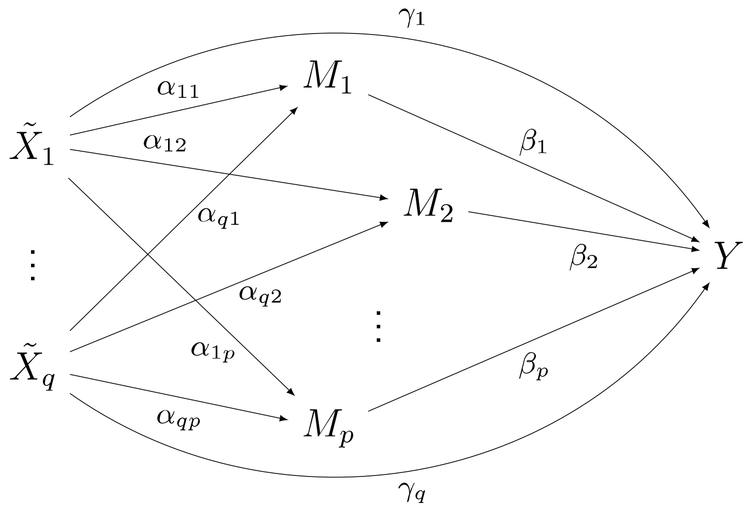
- $\mathbf{X}_i = (X_{i1}, \ldots, X_{ir})^\top \in \mathbb{R}^r$: r-dimensional exposure variables
- $\mathbf{M}_i = (M_{i1}, \ldots, M_{ip})^\top \in \mathbb{R}^p$: p-dimensional mediator vector
- $Y_i \in \mathbb{R}$: univariate outcome

- $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_n)^\top \in \mathbb{R}^{n \times r}$: exposure of n subjects
- $\mathbf{M} = (\mathbf{M}_1, \ldots, \mathbf{M}_n)^\top \in \mathbb{R}^{n \times p}$: mediator of n subjects
- $\mathbf{Y} = (Y_1, \ldots, Y_n)^\top \in \mathbb{R}^n$: outcome of n subjects

Orthogonalization of $\mathbf{X}$

- X's can be dependent
  - Ignoring dependence may raise bias in estimation.
- Assume $\mathbf{X}_i$ follow a multivariate normal distribution with covariance matrix Φ
  - Apply PCA on $\mathbf{X}_i$.
  - $\check{\mathbf{X}}_i = (\check{X}_{i1}, \ldots, \check{X}_{ir})^\top \in \mathbb{R}^r$: transformed data.
  - First q components account for over $s\%$ of the variation in $\mathbf{X}$.
  - Let $\tilde{\mathbf{X}}_i = (\check{X}_{i1}, \ldots, \check{X}_{iq})^\top \in \mathbb{R}^q$, first q elements in $\check{\mathbf{X}}_i$.
- $\check{\mathbf{X}}_j = (\check{X}_{1j}, \ldots, \check{X}_{nj})^\top \in \mathbb{R}^n$: independent exposure component.

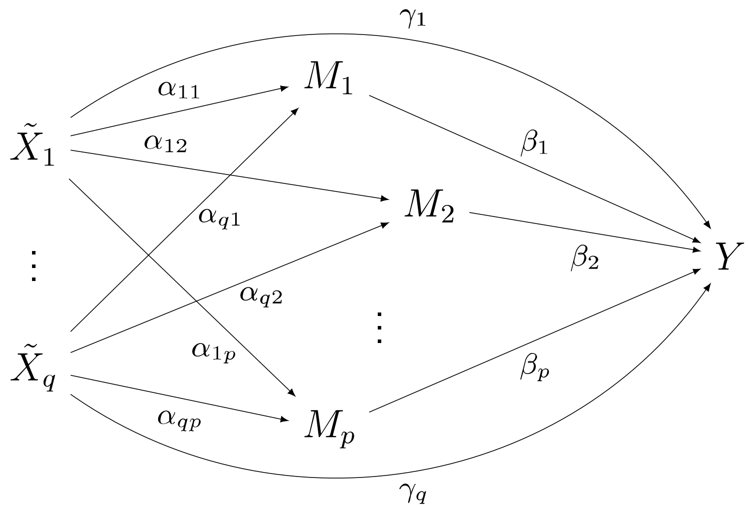
A marginal mediation model



$$\mathbf{M} = \tilde{\mathbf{X}}\boldsymbol{\alpha} + \boldsymbol{\epsilon}, \qquad \mathrm{vec}(\boldsymbol{\epsilon}) \sim \mathcal{N}(\mathbf{0}, \Sigma \otimes \mathbf{I}_n)$$

$$\mathbf{Y} = \tilde{\mathbf{X}}\boldsymbol{\gamma} + \mathbf{M}\boldsymbol{\beta} + \boldsymbol{\eta}, \qquad \boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I}_n)$$

- Exposures are independent/orthogonal.
- Allow the mediators to be dependent.



$$\begin{aligned} \mathbf{M} &= \tilde{\mathbf{X}}\boldsymbol{\alpha} + \boldsymbol{\epsilon}, & \mathrm{vec}(\boldsymbol{\epsilon}) &\sim \mathcal{N}(\mathbf{0}, \Sigma \otimes \mathbf{I}_n) \\ \mathbf{Y} &= \tilde{\mathbf{X}}\boldsymbol{\gamma} + \mathbf{M}\boldsymbol{\beta} + \boldsymbol{\eta}, & \boldsymbol{\eta} &\sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_n) \end{aligned}$$

- $\mathrm{IE}(\tilde{X}_j, M_k) = \alpha_{jk}\beta_k$: indirect effect of $\tilde{X}_j$ on Y through M_k.
- $\mathrm{IE}(\tilde{X}_j) = \sum_{k=1}^{p} \alpha_{jk}\beta_k$: total indirect effect of $\tilde{X}_j$ on Y.
- $\mathrm{DE}(\tilde{X}_j) = \gamma_j$: direct effect of $\tilde{X}_j$ on Y.
- $\mathrm{TE}(\tilde{X}_j) = \mathrm{IE}(\tilde{X}_j) + \mathrm{DE}(\tilde{X}_j) = \sum_{k=1}^{p} \alpha_{jk}\beta_k + \gamma_j$: total effect of $\tilde{X}_j$.
- $\mathrm{TE}_q = \sum_{j=1}^{q} \mathrm{TE}(\tilde{X}_j) = \sum_{j=1}^{q}(\sum_{k=1}^{p} \alpha_{jk}\beta_k + \gamma_j)$: total effect of q components.

Estimation

$$\underset{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}}{\text{minimize}} \ \frac{1}{2}\ell(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}) + \lambda_1\mathcal{R}_1(\boldsymbol{\alpha},\boldsymbol{\beta}) + \lambda_2\mathcal{R}_2(\boldsymbol{\alpha},\boldsymbol{\beta}) + \lambda_3\mathcal{R}_3(\boldsymbol{\gamma})$$

$$\ell = \operatorname{tr}\left\{(\mathbf{M}-\tilde{\mathbf{X}}\boldsymbol{\alpha})^\top(\mathbf{M}-\tilde{\mathbf{X}}\boldsymbol{\alpha})\right\} + (\mathbf{Y}-\tilde{\mathbf{X}}\boldsymbol{\gamma}-\mathbf{M}\boldsymbol{\beta})^\top(\mathbf{Y}-\tilde{\mathbf{M}}\boldsymbol{\gamma}-\mathbf{M}\boldsymbol{\beta})$$

$$\mathcal{R}_1 = \sum_{j=1}^{q}\sum_{k=1}^{p}\left\{|\alpha_{jk}\beta_k| + \phi(\alpha_{jk}^2+\beta_k^2)\right\} + \delta\left(\sum_{j=1}^{q}\sum_{k=1}^{p}|\alpha_{jk}| + \sum_{k=1}^{p}|\beta_k|\right)$$

$$\mathcal{R}_2 = \sum_{j=1}^{q}\sqrt{p}\sqrt{\sum_{k=1}^{p}(\alpha_{jk}\beta_k)^2}$$

$$\mathcal{R}_3 = \sum_{j=1}^{q}|\gamma_j|$$

- $\mathcal{R}_1$: Pathway lasso penalty (Zhao and Luo, 2022), mediator selection.
- $\mathcal{R}_2$: Group lasso penalty (Yuan and Lin, 2007), exposure selection.
- $\mathcal{R}_1+\mathcal{R}_2$: sparse-group lasso penalty (Simon et al. 2013).
- $\mathcal{R}_3$: lasso penalty, direct effect regularization.

Algorithm

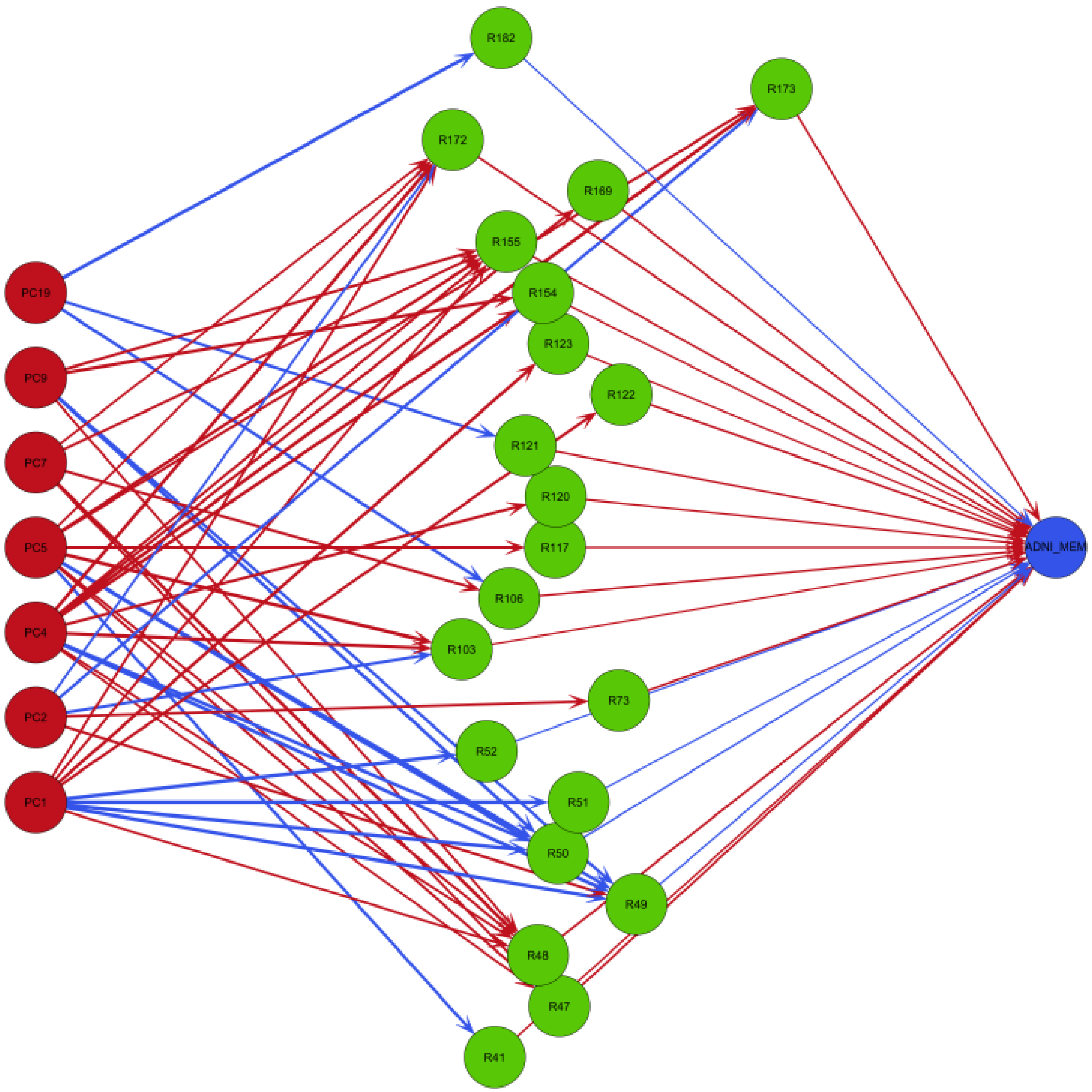
$$\begin{aligned}
\underset{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\mu}}{\text{minimize}} \quad & \frac{1}{2}\ell(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}) + \lambda_1\mathcal{R}_1(\boldsymbol{\mu},\boldsymbol{\alpha},\boldsymbol{\beta}) + \lambda_2\mathcal{R}_2(\boldsymbol{\mu}) + \lambda_3\mathcal{R}_3(\boldsymbol{\gamma}), \\
\text{such that} \quad & \mu_{jk} = \alpha_{jk}\beta_k, \quad \text{for } j = 1,\ldots,q \text{ and } k = 1,\ldots,p.
\end{aligned}$$

- Optimization: augmented Lagrangian
  - $\boldsymbol{\mu}$: sparse group lasso solution
  - $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$: lasso solution
- Tuning parameter selection: Bayesian information criterion (BIC)

ADNI application

- $n = 135$ subjects with mild cognitive impairment (MCI)
  - Pathological mechanism underlying this prodromal stage of AD.
- CSF proteomics (X)
  - Biomarkers Consortium CSF Proteomics MRM.
  - $r = 320$ peptides annotated from 142 proteins.
- Brain volumes (M)
  - $p = 145$ ROIs spanning the entire brain.
  - ROIs from a multi-atlas consensus-based label fusion scheme (Doshi et al., 2016).
  - Standardized by the total intracranial volume.
- Cognition (Y): ADNI_MEM
- Confounding adjustment:
  - Age, gender, Apolipoprotein E ε4 (APOE-ε4), years of education.

- PCA on X: first $q = 20$ PCs, about 85% data variation.
- Regularized mediation on $q = 20$ PCs and $p = 145$ mediators.



- PCA on X : first $q = 20$ PCs, about 85% data variation.
- Regularized mediation on $q = 20$ PCs and $p = 145$ mediators.

Effect of each PC

PC	IE	DE	TE
PC1	0.013	0.138	0.151
PC2	-0.003	-	-0.003
PC4	0.018	-	0.018
PC5	0.012	0.066	0.078
PC6	-	-0.035	-0.035
PC7	0.008	0.168	0.176
PC9	0.007	0.065	0.072
PC11	-	-0.018	-0.018
PC14	-	0.102	0.102
PC15	-	-0.007	-0.007
PC16	-	0.156	0.156
PC19	-0.001	-	-0.001
Total	0.054	0.634	0.688

Mediators with nonzero indirect effect

Region		IE ($\times 10^{-3}$)						
		PC1	PC2	PC4	PC5	PC7	PC9	PC19
R41	Left cerebellum white matter	-	-	-	-1.30	-	-	-
R47	Right hippocampus	-	-	1.12	1.60	1.52	-	-
R48	Left hippocampus	1.34	-	1.59	1.76	3.40	1.55	-
R49	Temporal horn of right lateral ventricle	2.06	-1.01	2.03	1.28	-	1.08	-
R50	Temporal horn of left lateral ventricle	2.55	-	1.78	2.05	-	1.74	-
R51	Right lateral ventricle	1.06	-	-	-	-	-	-
R52	Left lateral ventricle	1.15	-	-	-	-	-	-
R73	Cerebellar vermal lobules VIII-X	-	1.88	-	-	-	-	-
R103	Left anterior insula	-	-1.13	1.27	1.76	-	-	-
R106	Right angular gyrus	-	-	-	-	1.03	-	-1.41
R117	Left entorhinal areas	-	-	-	1.15	-	-	-
R120	Right frontal pole	-	-	1.12	-	-	-	-
R121	Left frontal pole	-	-	-	-	-	-	-1.12
R122	Right fusiform gyrus	1.32	-	-	-	-	-	-
R123	Left fusiform gyrus	1.12	-	-	-	-	-	-
R154	Right middle temporal gyrus	-	-	1.68	-	-	1.66	-
R155	Left middle temporal gyrus	1.01	-	1.09	1.63	1.05	1.35	-
R169	Left precuneus	-	-	1.10	-	-	-	-
R172	Right posterior insula	1.44	-1.30	2.67	1.03	1.00	-	-
R173	Left posterior insula	-	-1.46	2.18	1.09	-	-	-
R182	Right precentral gyrus	-	-	-	-	-	-	1.91

- Top-loaded proteins in PCs 1, 4, and 5.

Protein	Loading	Gene	Direction	Correlation	
				tau	amyloid
PC1					
ProSAAS	0.075	PCSK1N	⇕		
Neuronal growth regulator 1	0.075	NEGR1	↓		
Cell adhesion molecule 3	0.075	CADM3	↓		
Neuroblastoma suppressor of tumorigenicity 1	0.073	NBL1	↑		
Spondin-1	0.073	SPON1	↑	↑	↓
Prostagiandin-H2 D-isomerase	0.073	PTGDS	↓		↓
Monocyte differentiation antigen CD14	0.071	CD14	↑		
VPS10 domain-containing receptor SorCS1	0.069	SORCS1		↑	↓
PC4					
Neuronal pentraxin-2	0.152	NPTX2	↓		↑
Insulin-like growth factor-binding protein 2	-0.146	IGFBP2	⇕	↑	
Beta-2-microglobulin	-0.125	B2M	⇕	↓	
Neurexin-2	0.116	NRXN2	⇕		
Apolipoprotein D	-0.095	APOD	⇕	↑	
Neuronal pentraxin-1	0.093	NPTX1	↓		
Kallikrein-6	-0.083	KLK6	↑	↑	↑
Cystatin-C	-0.066	CST3	⇕	⇕	↑
PC5					
Complement C4-A	-0.180	C4A	↑		
Ectonucleotide pyrophosphatase/phosphodiesterase family member 2	-0.144	ENPP2	↑	↓	
Superoxide dismutase [Cu-Zn]	0.129	SOD1	↓	↑	↓
Complement factor B	0.110	CFB	↑		
Glial fibrillary acidic protein	-0.106	GFAP	↑		
Chromogranin-A	0.105	CHGA	⇕	↑	↑
Mimecan	-0.094	OGN	⇕		
Neurosecretory protein VGF	0.083	VGF	↓		↓
Alpha-1B-glycoprotein	0.075	A1BG	⇕		

↑: consistently upregulated in MCI/AD or positively correlated; ↓: consistently downregulated in MCI/AD or negatively correlated; ⇕: inconsistent reports.

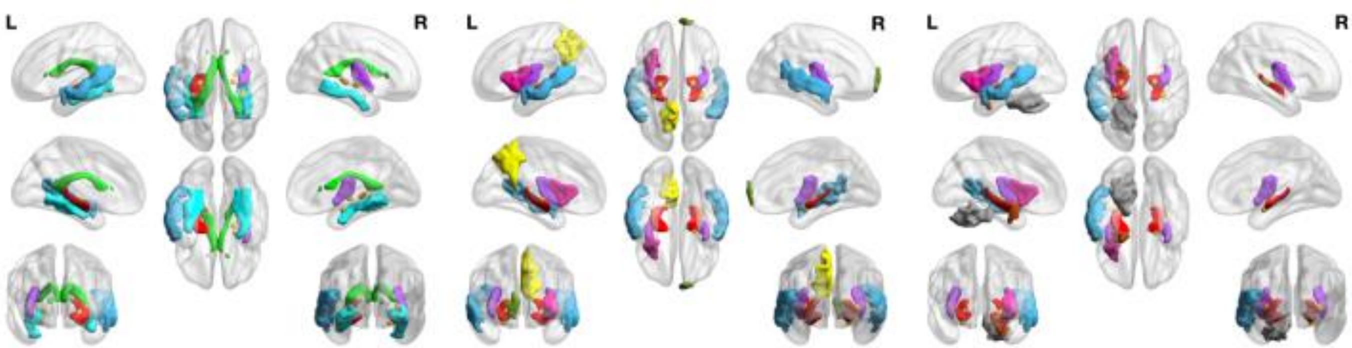
Top-loaded proteins

- Related to Aβ pathology
  - SPON1, SORCS1, PTGDS, CST3, NPTX2, VGF, CHGA.
- Related to tau pathology
  - IGFBP2.
- Related to Aβ and tau pathology
  - KLK6, SOD1.
- AD markers
  - APOD, NRXN2, GFAP.
- Related to brain structure/atrophy
  - NPTX2, CHGA, APOD, NEGR1, B2M, CST3.

Brain mediators

	Region	IE ($\times 10^{-3}$)						
		PC1	PC2	PC4	PC5	PC7	PC9	PC19
R41	Left cerebellum white matter	-	-	-	-1.30	-	-	-
R47	Right hippocampus	-	-	1.12	1.60	1.52	-	-
R48	Left hippocampus	1.34	-	1.59	1.76	3.40	1.55	-
R49	Temporal horn of right lateral ventricle	2.06	-1.01	2.03	1.28	-	1.08	-
R50	Temporal horn of left lateral ventricle	2.55	-	1.78	2.05	-	1.74	-
R51	Right lateral ventricle	1.06	-	-	-	-	-	-
R52	Left lateral ventricle	1.15	-	-	-	-	-	-
R73	Cerebellar vermal lobules VIII-X	-	1.88	-	-	-	-	-
R103	Left anterior insula	-	-1.13	1.27	1.76	-	-	-
R106	Right angular gyrus	-	-	-	-	1.03	-	-1.41
R117	Left entorhinal areas	-	-	-	1.15	-	-	-
R120	Right frontal pole	-	-	1.12	-	-	-	-
R121	Left frontal pole	-	-	-	-	-	-	-1.12
R122	Right fusiform gyrus	1.32	-	-	-	-	-	-
R123	Left fusiform gyrus	1.12	-	-	-	-	-	-
R154	Right middle temporal gyrus	-	-	1.68	-	-	1.66	-
R155	Left middle temporal gyrus	1.01	-	1.09	1.63	1.05	1.35	-
R169	Left precuneus	-	-	1.10	-	-	-	-
R172	Right posterior insula	1.44	-1.30	2.67	1.03	1.00	-	-
R173	Left posterior insula	-	-1.46	2.18	1.09	-	-	-
R182	Right precentral gyrus	-	-	-	-	-	-	1.91

	Region	α							β
		PC1	PC2	PC4	PC5	PC7	PC9	PC19	
R41	Left cerebellum white matter	-	-	-	-0.167	-	-	-	0.007
R47	Right hippocampus	-	-	0.106	0.134	0.130	-	-	0.012
R48	Left hippocampus	0.109	-	0.125	0.131	0.216	0.116	-	0.012
R49	Temporal horn of right lateral ventricle	-0.250	0.148	-0.262	-0.182	-	-0.159	-	-0.006
R50	Temporal horn of left lateral ventricle	-0.287	-	-0.232	-0.245	-	-0.208	-	-0.007
R51	Right lateral ventricle	-0.356	-	-	-	-	-	-	-0.003
R52	Left lateral ventricle	-0.362	-	-	-	-	-	-	-0.003
R73	Cerebellar vermal lobules VIII-X	-	0.143	-	-	-	-	-	0.010
R103	Left anterior insula	-	-0.172	0.202	0.246	-	-	-	0.006
R106	Right angular gyrus	-	-	-	-	0.150	-	-0.181	0.008
R117	Left entorhinal areas	-	-	-	0.168	-	-	-	0.008
R120	Right frontal pole	-	-	0.159	-	-	-	-	0.007
R121	Left frontal pole	-	-	-	-	-	-	-0.156	0.008
R122	Right fusiform gyrus	0.156	-	-	-	-	-	-	0.010
R123	Left fusiform gyrus	0.187	-	-	-	-	-	-	0.006
R154	Right middle temporal gyrus	-	-	0.207	-	-	0.193	-	0.007
R155	Left middle temporal gyrus	0.133	-	0.141	0.181	0.137	0.155	-	0.008
R169	Left precuneus	-	-	0.151	-	-	-	-	0.008
R172	Right posterior insula	0.132	-0.117	0.217	0.108	0.107	-	-	0.010
R173	Left posterior insula	-	-0.166	0.242	0.147	-	-	-	0.008
R182	Right precentral gyrus	-	-	-	-	-	-	-0.237	-0.005



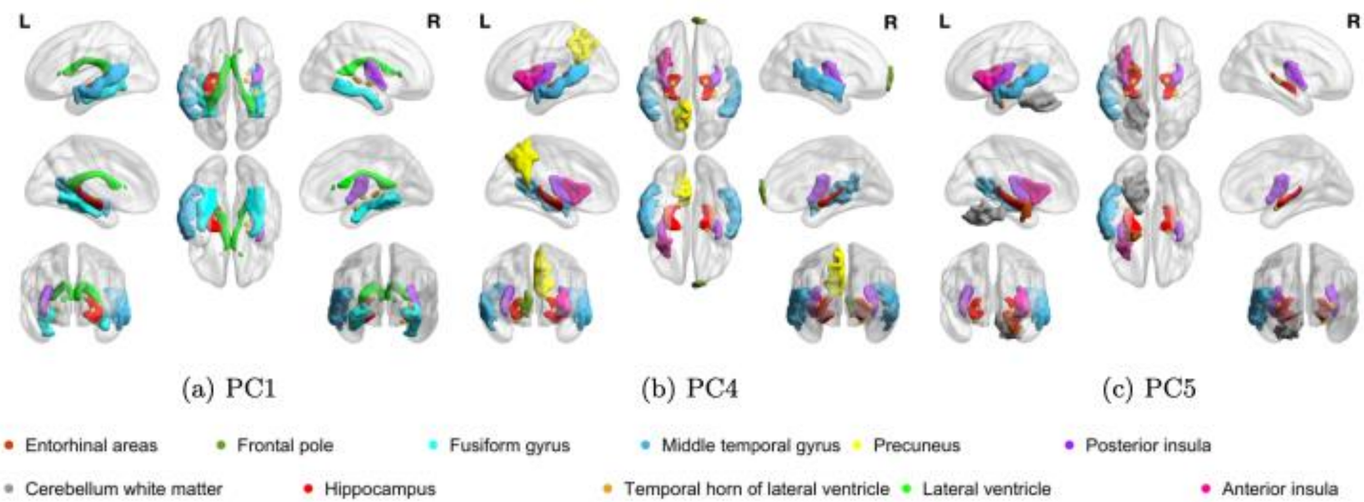
(a) PC1 (b) PC4 (c) PC5

Entorhinal areas · Frontal pole · Fusiform gyrus · Middle temporal gyrus · Precuneus · Posterior insula
Cerebellum white matter · Hippocampus · Temporal horn of lateral ventricle · Lateral ventricle · Anterior insula

Brain mediators

	Region	IE ($\times 10^{-3}$)						
		PC1	PC2	PC4	PC5	PC7	PC9	PC19
R41	Left cerebellum white matter	-	-	-	-1.30	-	-	-
R47	Right hippocampus	-	-	1.12	1.60	1.52	-	-
R48	Left hippocampus	1.34	-	1.59	1.76	3.40	1.55	-
R49	Temporal horn of right lateral ventricle	2.06	-1.01	2.03	1.28	-	1.08	-
R50	Temporal horn of left lateral ventricle	2.55	-	1.78	2.05	-	1.74	-
R51	Right lateral ventricle	1.06	-	-	-	-	-	-
R52	Left lateral ventricle	1.15	-	-	-	-	-	-
R73	Cerebellar vermal lobules VIII-X	-	1.88	-	-	-	-	-
R103	Left anterior insula	-	-1.13	1.27	1.76	-	-	-
R106	Right angular gyrus	-	-	-	-	1.03	-	-1.41
R117	Left entorhinal areas	-	-	-	1.15	-	-	-
R120	Right frontal pole	-	-	1.12	-	-	-	-
R121	Left frontal pole	-	-	-	-	-	-	-1.12
R122	Right fusiform gyrus	1.32	-	-	-	-	-	-
R123	Left fusiform gyrus	1.12	-	-	-	-	-	-
R154	Right middle temporal gyrus	-	-	1.68	-	-	1.66	-
R155	Left middle temporal gyrus	1.01	-	1.09	1.63	1.05	1.35	-
R169	Left precuneus	-	-	1.10	-	-	-	-
R172	Right posterior insula	1.44	-1.30	2.67	1.03	1.00	-	-
R173	Left posterior insula	-	-1.46	2.18	1.09	-	-	-
R182	Right precentral gyrus	-	-	-	-	-	-	1.91

	Region	α							β
		PC1	PC2	PC4	PC5	PC7	PC9	PC19	
R41	Left cerebellum white matter	-	-	-	-0.167	-	-	-	0.007
R47	Right hippocampus	-	-	0.106	0.134	0.130	-	-	0.012
R48	Left hippocampus	0.109	-	0.125	0.131	0.216	0.116	-	0.012
R49	Temporal horn of right lateral ventricle	-0.250	0.148	-0.262	-0.182	-	-0.159	-	-0.006
R50	Temporal horn of left lateral ventricle	-0.287	-	-0.232	-0.245	-	-0.208	-	-0.007
R51	Right lateral ventricle	-0.356	-	-	-	-	-	-	-0.003
R52	Left lateral ventricle	-0.362	-	-	-	-	-	-	-0.003
R73	Cerebellar vermal lobules VIII-X	-	0.143	-	-	-	-	-	0.010
R103	Left anterior insula	-	-0.172	0.202	0.246	-	-	-	0.006
R106	Right angular gyrus	-	-	-	-	0.150	-	-0.181	0.008
R117	Left entorhinal areas	-	-	-	0.168	-	-	-	0.008
R120	Right frontal pole	-	-	0.159	-	-	-	-	0.007
R121	Left frontal pole	-	-	-	-	-	-	-0.156	0.008
R122	Right fusiform gyrus	0.156	-	-	-	-	-	-	0.010
R123	Left fusiform gyrus	0.187	-	-	-	-	-	-	0.006
R154	Right middle temporal gyrus	-	-	0.207	-	-	0.193	-	0.007
R155	Left middle temporal gyrus	0.133	-	0.141	0.181	0.137	0.155	-	0.008
R169	Left precuneus	-	-	0.151	-	-	-	-	0.008
R172	Right posterior insula	0.132	-0.117	0.217	0.108	0.107	-	-	0.010
R173	Left posterior insula	-	-0.166	0.242	0.147	-	-	-	0.008
R182	Right precentral gyrus	-	-	-	-	-	-	-0.237	-0.005



(a) PC1 (b) PC4 (c) PC5

Brain mediators

- The hippocampus and entorhinal cortex
  - Part of the medial temporal lobe, atrophy occurs early, impairments of this system are responsible for the deficit in episodic memory (Nadel and Hardt, 2011).
  - Hippocampal atrophy: best established and validated biomarker across the entire disease spectrum (Jack Jr et al., 2011).
  - Entorhinal cortex: compared to healthy controls, greater atrophy observed in AD patients followed by MCI subjects (Pini et al., 2016).
- The lateral temporal, parietal, and frontal cortex
  - Fusiform (facial recognition), middle temporal gyrus (MTG, language area), precuneus (episodic memories), precentral: atrophy in AD (Parker et al., 2018).
  - Fusiform and MTG: tau deposition (Schultz et al., 2018).
  - Insular atrophy: related to memory deficits in AD (Lin et al., 2017).
- The lateral ventricles
  - Sharp contrast between CSF and surrounding tissues in T1.
  - Enlargement in the lateral ventricles: an important AD marker.

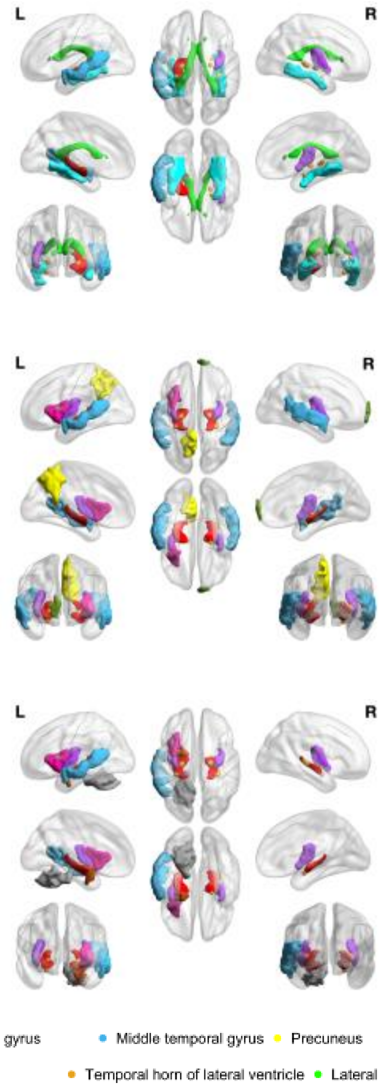
CSF Protein

NBL1, SPON1,
SORCS1, PCSK1N,
PTGDS, NEGR1,
CD14, CADM3

B2M, NPTX2,
IGFBP2, NPTX1,
KLK6, APOD,
NRXN2, CST3

SOD1, VGF,
ENPP2, C4A,
CFB, GFAP, OGN,
CHGA, A1BG

Brain volume

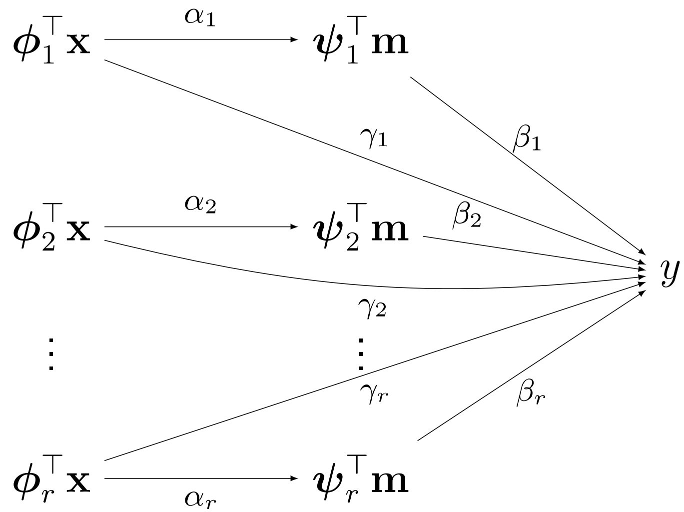


ADNI_MEM

Summary

- Multiview data integration using mediation analysis
  - High-dimensional exposures and high-dimensional mediators.
  - **Orthogonalization** + **marginal model**: allow dependence in exposures and mediators.
  - **Group Lasso** + **Pathway Lasso**: exposure and mediator selection.
  - ADNI data: identify meaningful brain pathways.
- Zhao and Li (2022) *Human Brain Mapping* (doi.org/10.1002/hbm.25800).
- R code available on GitHub (https://github.com/zhaoyi1026/HDExposureMediator).

Another approach



$$
\begin{aligned}
\mathbf{M}\boldsymbol{\psi} &= \mathbf{X}\boldsymbol{\phi} \cdot \alpha + \boldsymbol{\varepsilon}, \\
\mathbf{Y} &= \mathbf{X}\boldsymbol{\phi} \cdot \gamma + \mathbf{M}\boldsymbol{\psi} \cdot \beta + \boldsymbol{\eta},
\end{aligned}
$$

- A simplified model assuming one-to-one correspondence.
- $\boldsymbol{\phi}^\top \mathbf{x}$ exposure component, $\boldsymbol{\psi}^\top \mathbf{m}$ mediator component.
- Identify $(\boldsymbol{\phi}, \boldsymbol{\psi}, \alpha, \beta, \gamma)$ via likelihood-based approach.

ADNI data

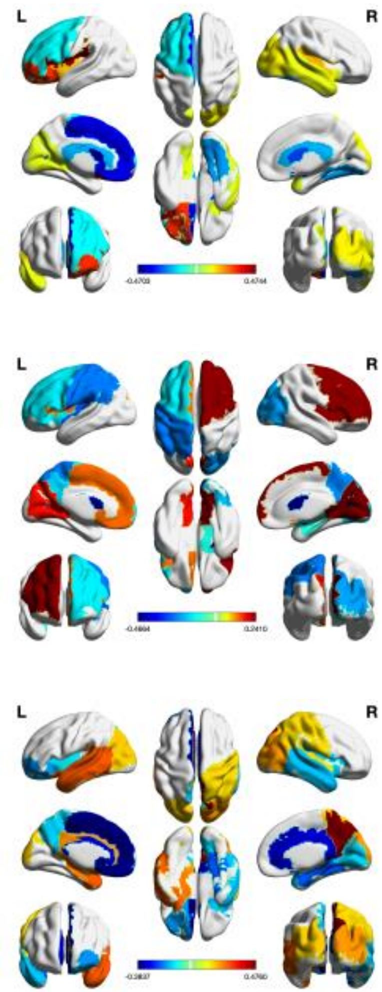
CSF Protein

NPTX2, NRCAM, NRXN1, VGF, PCSK1N, B2M, CH3L1, PTGDS, KLK6, A1BG, CSTN3

KLK6, CLUS, BASP1, A1BG, APOD, APOE, IGFBP2, GFAP

PCSK1N, A1BG, PTGDS, CD14, CFG, NPTX1, B2M, OGN, CSTN3, VGF, C4A, ENPP2, CHGA, CST3, ENO2

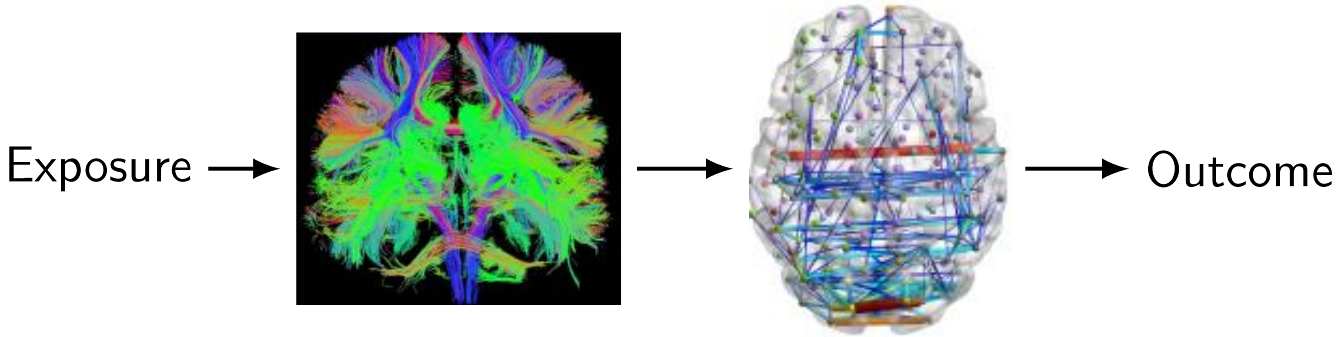
Brain volume

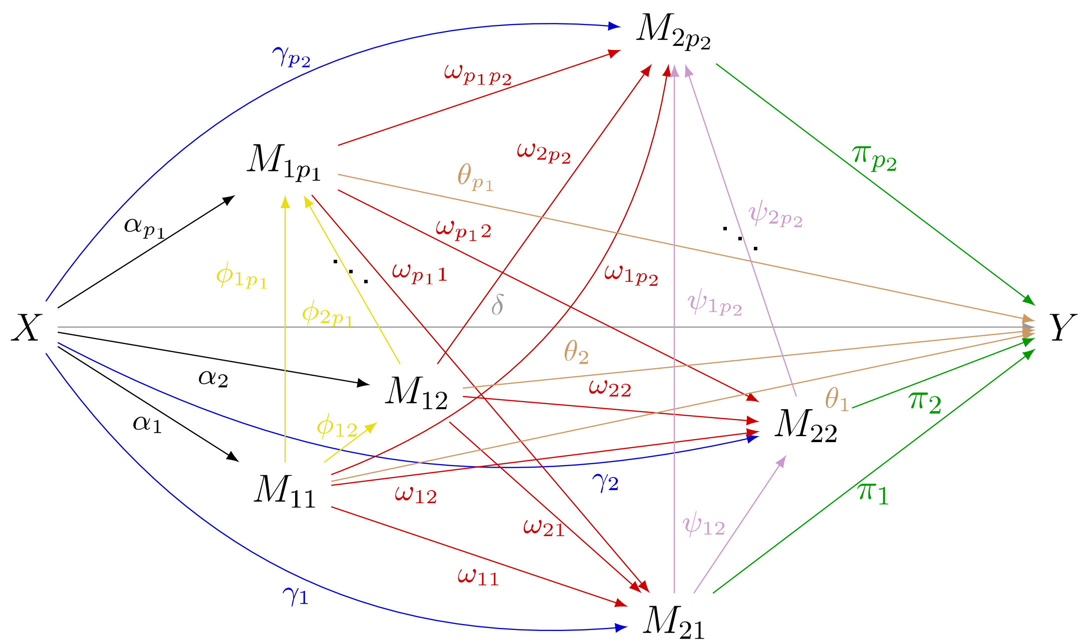


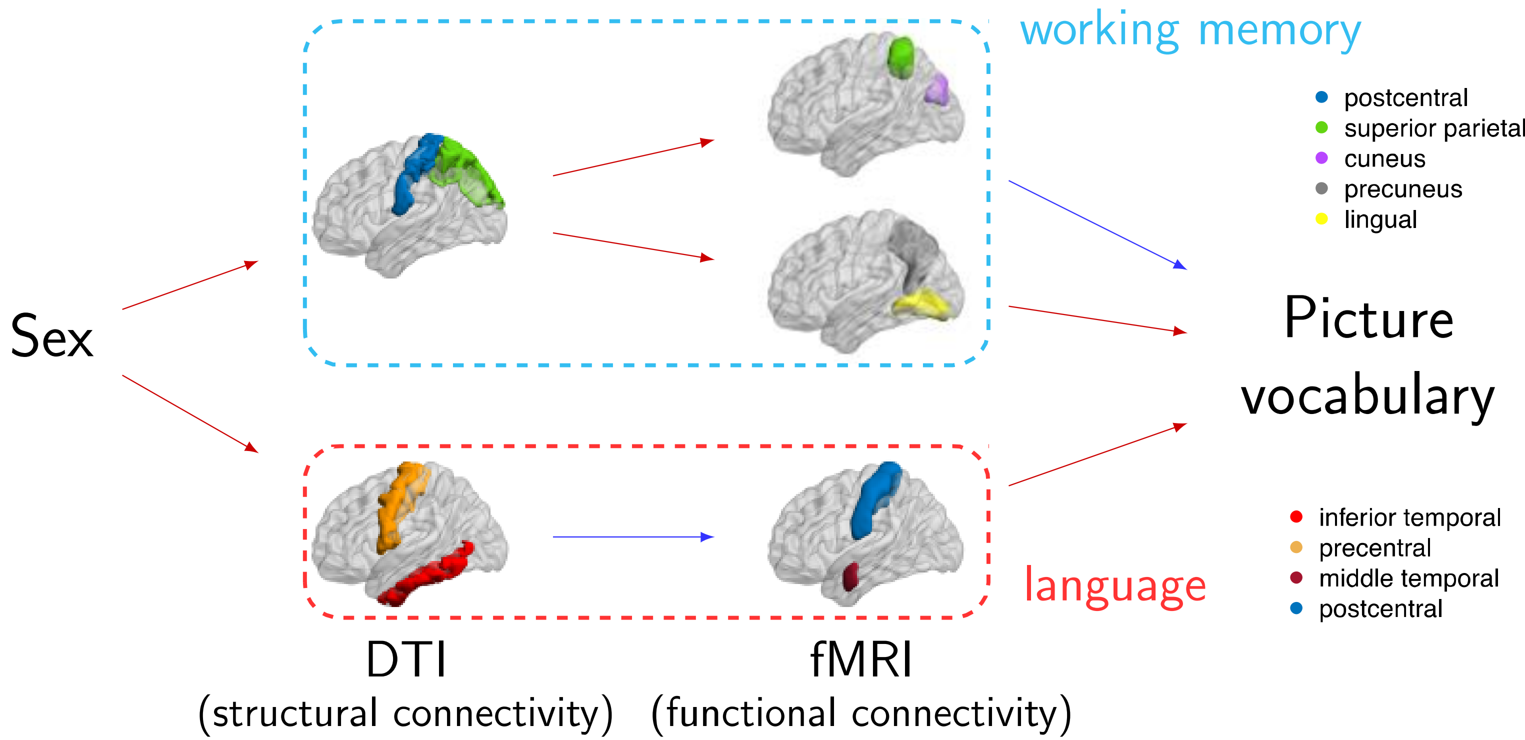
ADNI_MEM

- Zhao (2024) *Statistics in Medicine* (doi.org/10.1002/sim.10215).

Mediation with two blocks of mediators







- Zhao, Li, and Caffo (2021) *Biometrics* (`doi.org/10.1111/biom.13351`).

Acknowledgements

- The Alzheimer's Disease Neuroimaging Initiative (ADNI).
- The Human Connectome Project (HCP).
- Funding: NIMH R01MH126970.

Thank you!