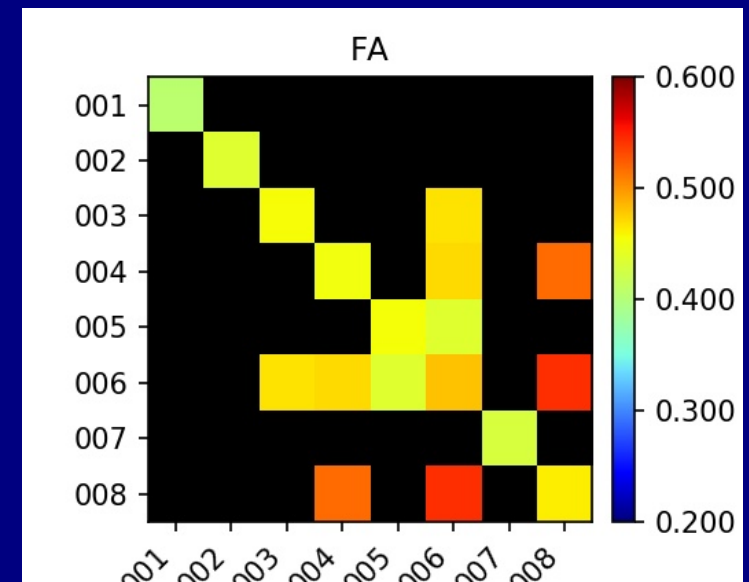


Introduction to:

investigating networks with multivariate modeling



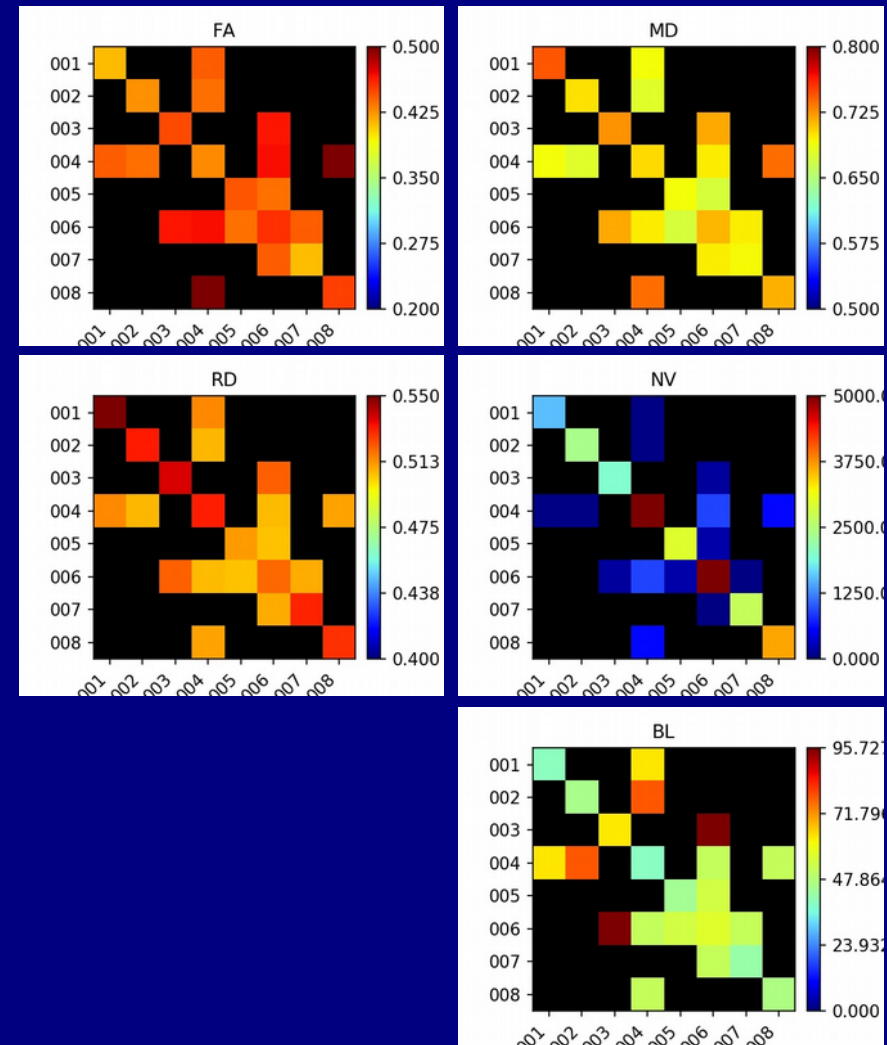
Outline

- + We describe combining DTI or FMRI network results (matrices) with non-MRI data (e.g., age, test scores, characteristics, etc.) for group analysis.
- + General motivation for multivariate modeling (MVM)
- + Case study example

WMC Quantities

For pairs of targets in a network, have an average WMC property (or can map to T1, PD...) →

Have produced sets of localized structural/anatomical quantities for comparison with functional values or behavioral scores, genetics, etc.



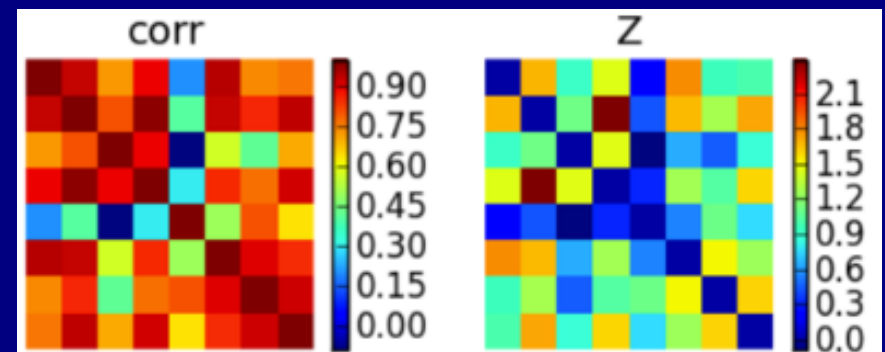
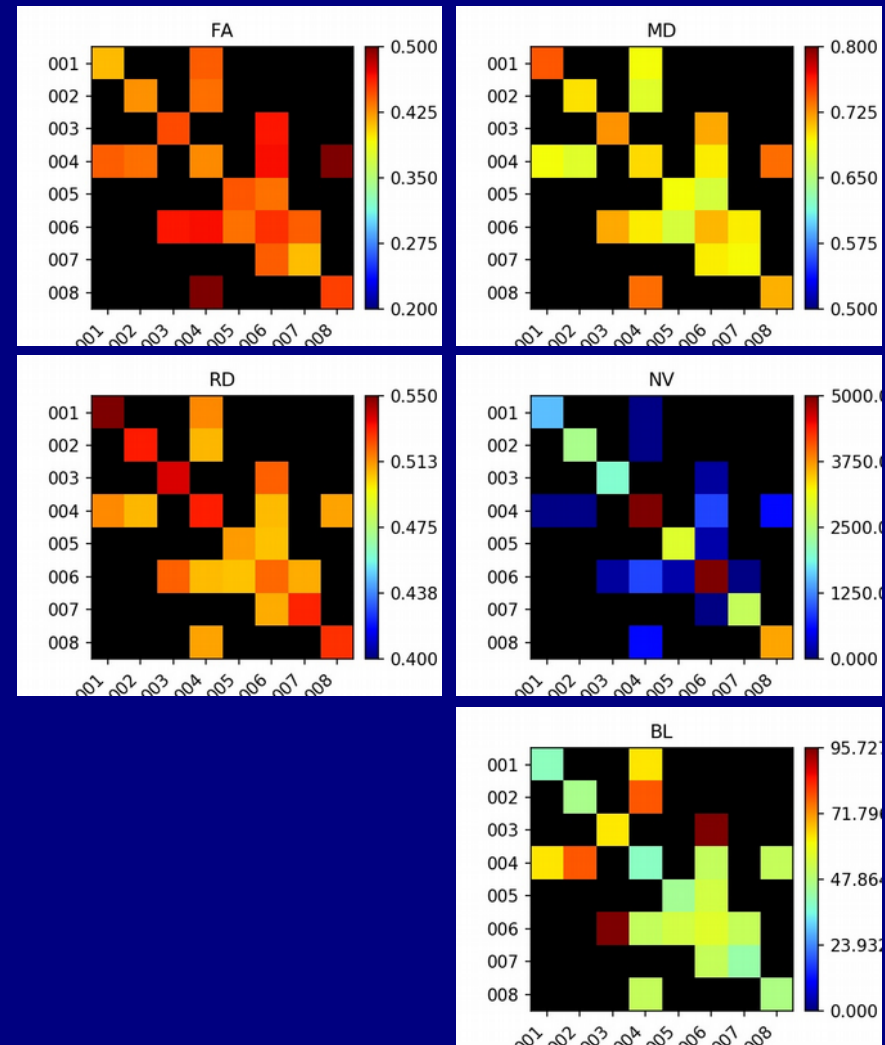
WMC Quantities

For pairs of targets in a network, have an average WMC property (or can map to T1, PD...) →

Have produced sets of localized structural/anatomical quantities for comparison with functional values or behavioral scores, genetics, etc.

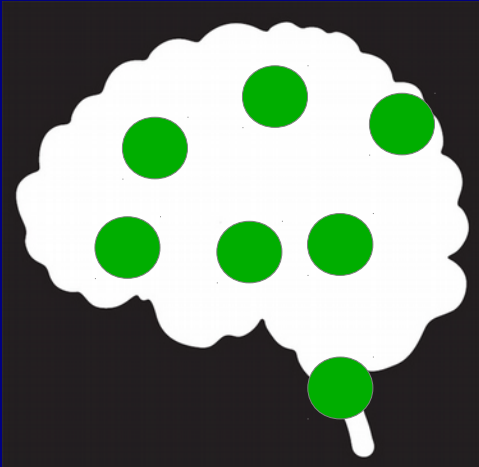
Also works for GM quantities (FC)

3dNetCorr: correlation matrices of average time series in ROIs (e.g., uninflated GM ROIs from 3dROIMaker)



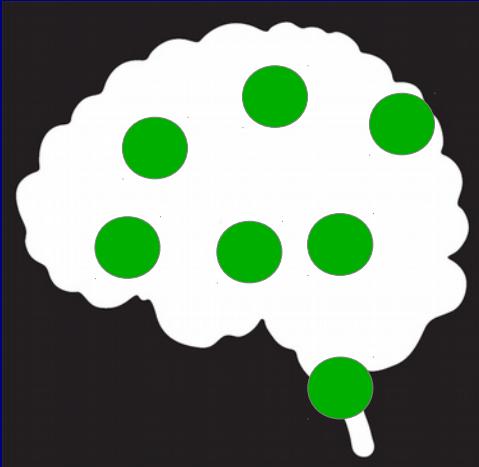
Group Analysis Steps

1) Place network targets

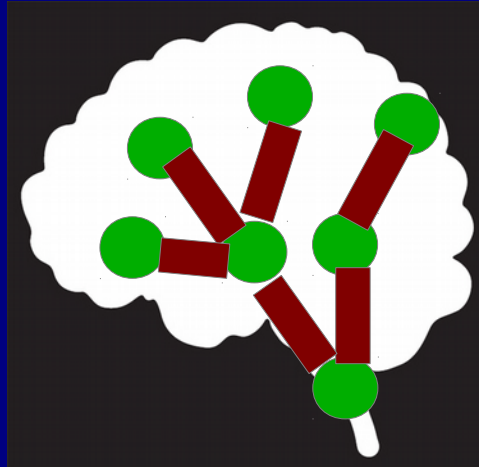


Group Analysis Steps

1) Place network targets

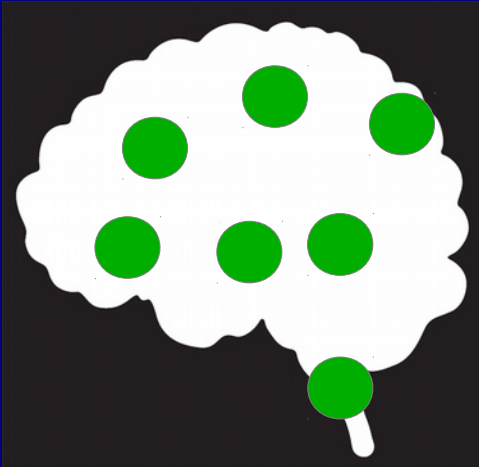


2) Probabilistic tracking

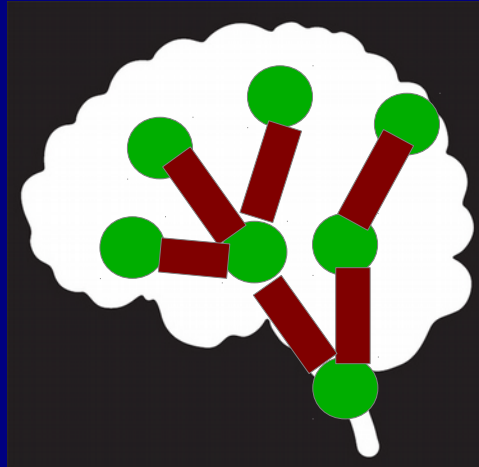


Group Analysis Steps

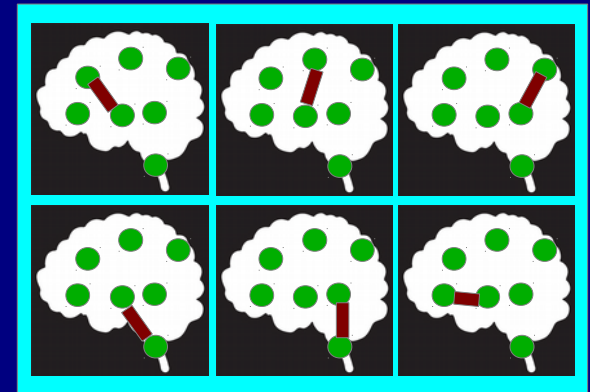
1) Place network targets



2) Probabilistic tracking

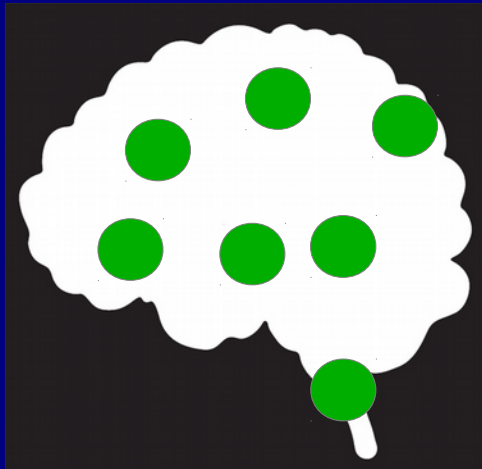


3) set of WM ROIs → set of simultaneous measures

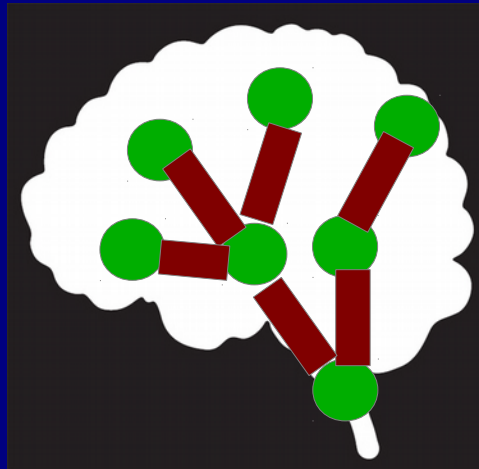


Group Analysis Steps

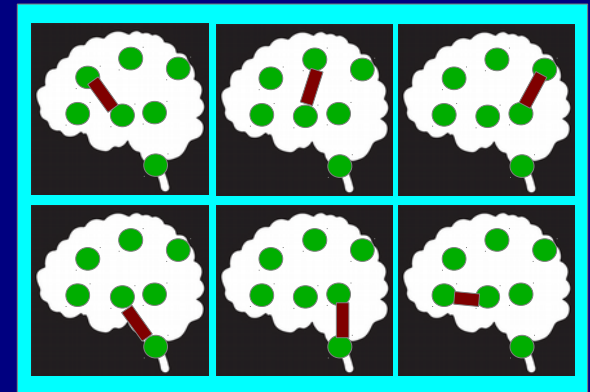
1) Place network targets



2) Probabilistic tracking



3) set of WM ROIs → set of simultaneous measures



4) Network-level test:
multivariate model (MVM)



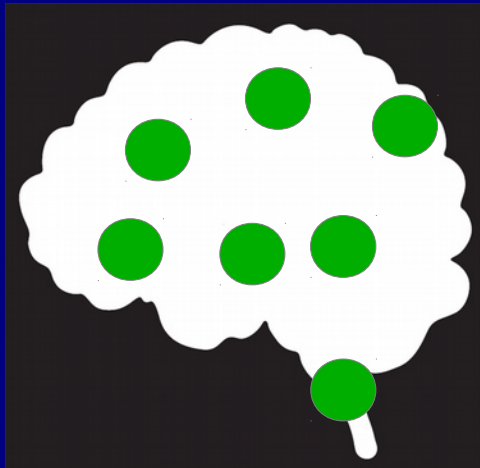
$\{FA_1, FA_2, FA_3, \dots\}$

$\sim \text{var1} + \text{var2} + \text{var3} \dots$

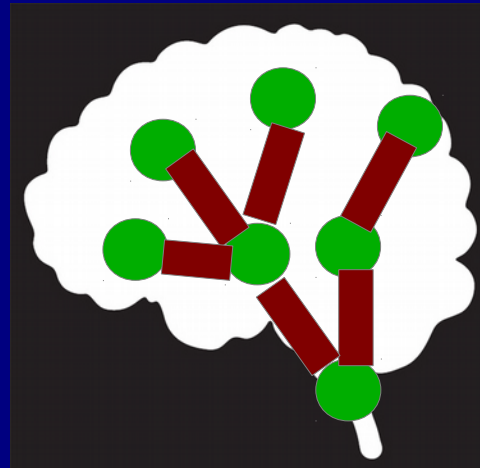
➡ AFNI's 3dMVM, written by G. Chen

Group Analysis Steps

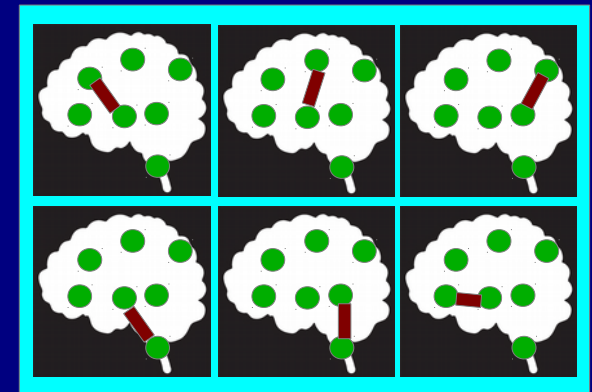
1) Place network targets



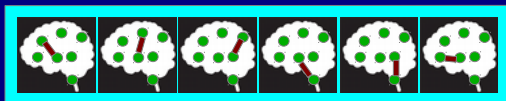
2) Probabilistic tracking



3) set of WM ROIs → set of simultaneous measures



4) Network-level test: multivariate model (MVM)



$\{FA_1, FA_2, FA_3, \dots\}$
 $\sim \text{var1} + \text{var2} + \text{var3} \dots$

5) WMC-level / ROI-level tests: follow-up GLM for each WMC

$FA_1 \sim \text{var1} + \text{var2} + \text{var3} \dots$



$FA_2 \sim \text{var1} + \text{var2} + \text{var3} \dots$

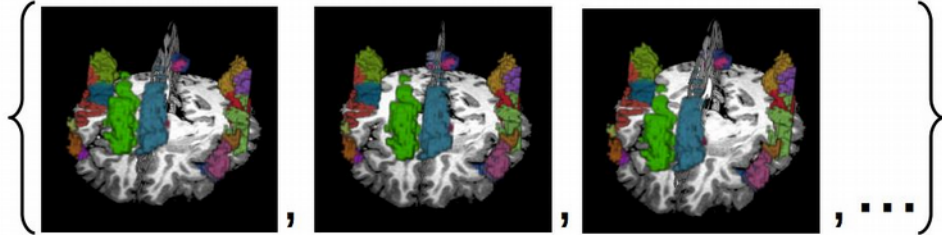


...

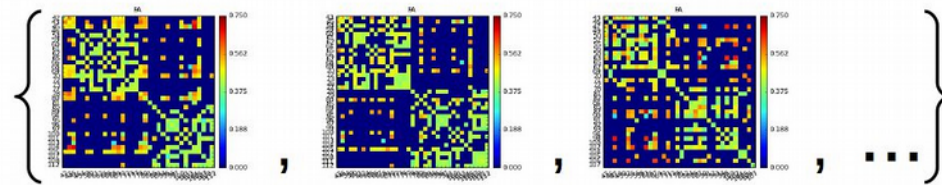
➡ AFNI's 3dMVM, written by G. Chen

Group Analysis: Summary

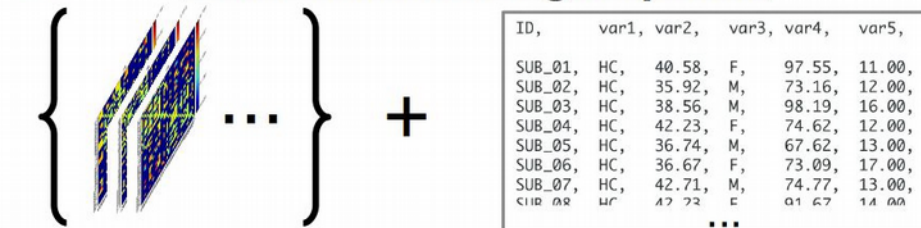
A. N networks of FMRI/DTI ROIs



B. N functional/structural matrices



C. Combined MRI + group data



D. User-defined model(s)

```
'measure1 ~ var1 + var2 + var3*var4 + ...'
'measure2 ~ var1 + var2 + var3*var4 + ...'
...
```

E. Network-level statistics for each model

ANOVA table of χ^2 , DF, and p-value:

```
# RESULTS: ANOVA table - FA
5 # Number of effects
# Chisq DF Pr(>Chisq)
4.9457216 1 2.615532e-02 # var1
0.8453055 1 3.578838e-01 # var2
0.6640459 1 4.151352e-01 # var3
0.8097606 1 3.681910e-01 # var4
2.1255675 1 1.448591e-01 # var3:var4
```

F. Set of ROI statistics for each model

Post hoc table of value, t-stat, DF and 2-sided p:

```
# RESULTS: Post hoc tests - FA
54 # Number of tests
# value t-stat DF 2-sided-P
-0.0044778181 -0.62834967 14 5.398911e-01 # 001_002-var1(+HC-IL)
-0.0002940607 -0.23287694 14 8.192272e-01 # 001_002-var2
0.0011186177 2.13603173 14 5.082097e-02 # 001_002-var3(+F-M)-var4
-0.0069573895 -1.12411575 14 2.798695e-01 # 001_002-var3(+F-M)
0.0004507261 1.54181323 14 1.454148e-01 # 001_002-var4
0.0130966286 1.51536073 14 1.519300e-01 # 003_004-var1(+HC-IL)
0.0010852927 0.70869270 14 4.901486e-01 # 003_004-var2
```

Helper functions

Combine data: **fat_mvm_prep.py**

+ make a data table combining:

- a CSV (~XLS) file of subject data with
- a set of 3dTrackID “*.grid” (or 3dNetcorr “*.netcc”) files

+ automatically determines matrix elements found across all subj (*some* missing data allowed with LME modeling)

Specify model + GLTs: **fat_mvm_scripter.py**

+ define a statistical model of MRI data from CSV columns

+ build a 3dMVM command for both

- the network-level model, and
- the follow-up GLTs (to investigate individual elements)

Example:
Group analysis of tracked networks
using multivariate statistics

from study:

*A DTI-Based Tractography Study of Effects
on Brain Structure Associated with
Prenatal Alcohol Exposure in Newborns,*
*Taylor, Jacobson, van der Kouwe, Molteno, Chen,
Wintermark, Alhamud, Jacobson, Meintjes (2015, HBM)*

Prenatal alcohol exposure (PAE)

- Alcohol is a teratogen, disrupting healthy embryonic and fetal development.
 - leads to various **Fetal Alcohol Spectrum Disorders (FASD)**
- FASD occurs in children whose pregnant mothers binge drank
 - e.g., ≥ 4 drinks/occasion and/or ≥ 14 drinks/wk
- Results in *poor*:
 - academic performance
 - language/math skills
 - impulse control
 - abstract reasoning
 - memory, attention and facial and skeletal dysmorphology



Fig. A:
Fetal Alcohol Syndrome, Diagnosis, Epidemiology, Prevention, and Treatment. (Institute of Medicine, 1996).

A



B



Fig. B:
Reprinted with permission from Clarren & Smith, (1978). Copyright 1978 by the New England Journal of Medicine, Massachusetts Medical Society.

12

C



D



Figs. C and D:
Reprinted with permission from Jones et al. (1973). Copyright 1973 by the Lancet Ltd.

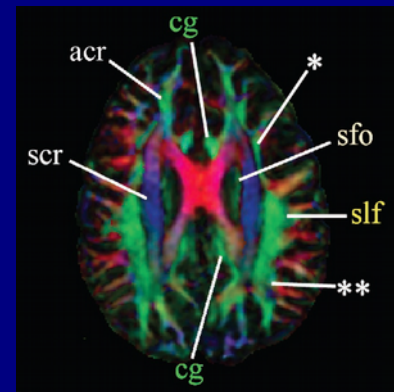
Goals of this study

To:

- 1) Use neuroimaging to compare structural brain development in newborns with PAE to that of HC newborns.
- 2) Quantitatively examine WM properties across the brain
- 3) Relate changes in (localized) WM properties with PAE, controlling for several confounding effects
→ examine several, and see which is/are (most) significant

Tools: diffusion tensor imaging (DTI) + tractography

- A) delineate similar WM ROIs across all subjects
- B) quantify structural properties (FA, MD, T1, ...)
- C) statistical modeling for comparisons
- *at whole brain, network and ROI levels*



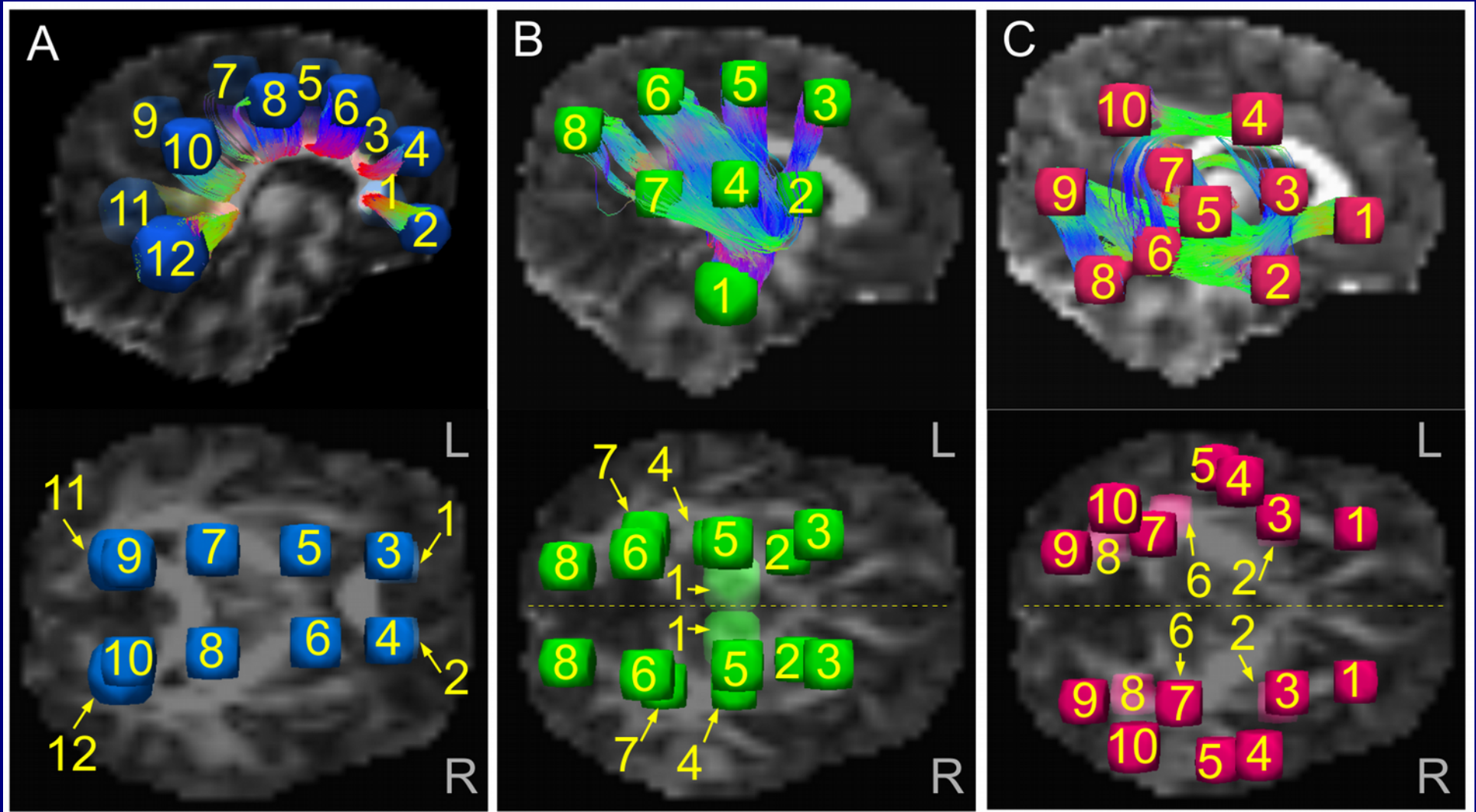
Setting up DTI-tractography

Location of targets for tractography: 5 WM networks.

CC and Cor. Rad.
(CCCR)

Projection
(L/R-PROJ)

Association
(L/R-ASSOC)



II) Results: network level

The questions:

- 1) which WM networks are affected by PAE?
- 2) which parameters show effects most strongly?

Answer using:

MVM for each network:

{set of DTI parameters} \sim
alcohol (frequency: binge/wk) +
infant age (wks since conception) +
infant sex (M/F) +
maternal age (yrs) +
maternal cigarette smoking (cig/day).

II) Results: network level

The questions:

- 1) which WM networks are affected by PAE?
- 2) which parameters show effects most strongly?

Parameters showing at least trends ($p < 0.1$) →

← Networks

Network	FA				MD				AD				PD			
	var.	β_{med}	$F(df_N, df_D)$	p	var.	β_{med}	$F(df_N, df_D)$	p	var.	β_{med}	$F(df_N, df_D)$	p	var.	β_{med}	$F(df_N, df_D)$	p
CCCR					alc	-0.70	8.6 (1, 14)	0.011*	alc	-0.72	14.0 (1, 14)	0.002**	cig	0.47	3.5 (1, 14)	0.083
L-PROJ	cig	0.12	4.2 (11, 4)	0.091	mat_age	0.56	5.5 (1, 14)	0.034*	mat_age	0.53	6.3 (1, 14)	0.025*	cig	0.52	4.0 (1, 14)	0.066
					alc	-0.41	3.9 (10, 140)	0.000***	alc	-0.52	4.1 (10, 140)	0.000***				
R-PROJ	age	0.33	8.6 (13, 2)	0.109	mat_age	0.37	4.4 (1, 14)	0.056	mat_age	0.44	6.5 (1, 14)	0.023*	cig	0.48	3.4 (1, 14)	0.085
					alc	-0.41	1.9 (12, 168)	0.035*	alc	-0.45	2.7 (12, 168)	0.002**				
					age	-0.41	5.8 (1, 14)	0.031*	age	-0.39	5.3 (1, 14)	0.038*				
L-ASSOC	mat_age	-0.16	9.2 (13, 2)	0.103	sex	-0.20	4.3 (1, 14)	0.056	sex	-0.39	5.9 (1, 14)	0.029*	cig	0.49	3.6 (1, 14)	0.080
					alc	-0.65	6.0 (7, 8)	0.011*	alc	-0.66	8.1 (1, 14)	0.013*				
					mat_age	0.44	3.8 (1, 14)	0.071	age	-0.16	2.5 (6, 84)	0.030*				
R-ASSOC	alc	0.23	1.8 (7, 98)	0.090	mat_age	0.44	3.8 (1, 14)	0.071	mat_age	0.43	4.7 (1, 14)	0.048*	cig	0.5	3.5 (1, 14)	0.082
					alc	-0.62	10.2 (1, 14)	0.007**	alc	-0.67	14.1 (1, 14)	0.002**				
									cig	-0.29	3.9 (1, 14)	0.068				

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

II) Results: network level

The questions:

- 1) which WM networks are affected by PAE?
- 2) which parameters show effects most strongly?

Parameters showing at least trends ($p < 0.1$) →

← Networks

	FA					MD					AD					PD			
Network	var.	β_{med}	$F(df_N, df_D)$	p		var.	β_{med}	$F(df_N, df_D)$	p		var.	β_{med}	$F(df_N, df_D)$	p		var.	β_{med}	$F(df_N, df_D)$	p
CCCR						alc	-0.70	8.6 (1, 14)	0.011*		alc	-0.72	14.0 (1, 14)	0.002**		cig	0.47	3.5 (1, 14)	0.083
						mat_age	0.56	5.5 (1, 14)	0.034*		cig	-0.27	2.5 (6, 9)	0.101					
											mat_age	0.53	6.3 (1, 14)	0.025*					
L-PROJ	cig	0.12	4.2 (11, 4)	0.091		alc	-0.41	3.9 (10, 140)	0.000***		alc	-0.52	4.1 (10, 140)	0.000***		cig	0.52	4.0 (1, 14)	0.066
						mat_age	0.37	4.4 (1, 14)	0.056		mat_age	0.44	6.5 (1, 14)	0.023*					
R-PROJ	age	0.33	8.6 (13, 2)	0.109		alc	-0.41	1.9 (12, 168)	0.035*		alc	-0.45	2.7 (12, 168)	0.002**		cig	0.48	3.4 (1, 14)	0.085
	mat_age	-0.16	9.2 (13, 2)	0.103		age	-0.41	5.8 (1, 14)	0.031*		age	-0.39	5.3 (1, 14)	0.038*					
						sex	-0.20	4.3 (1, 14)	0.056		sex	-0.39	5.9 (1, 14)	0.029*					
L-ASSOC						alc	-0.65	6.0 (7, 8)	0.011*		alc	-0.66	8.1 (1, 14)	0.013*		cig	0.49	3.6 (1, 14)	0.080
						mat_age	0.44	3.8 (1, 14)	0.071		age	-0.16	2.5 (6, 84)	0.030*					
											mat_age	0.43	4.7 (1, 14)	0.048*					
R-ASSOC	alc	0.23	1.8 (7, 98)	0.090		alc	-0.62	10.2 (1, 14)	0.007**		alc	-0.67	14.1 (1, 14)	0.002**		cig	0.5	3.5 (1, 14)	0.082
											cig	-0.29	3.9 (1, 14)	0.068					

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

→ Statistically significant alcohol exposure associations in ~every WM network

II) Results: network level

The questions:

- 1) which WM networks are affected by PAE?
- 2) which parameters show effects most strongly?

Parameters showing at least trends ($p < 0.1$) →

← Networks

Network	FA				MD				AD				PD			
	var.	β_{med}	$F(df_N, df_D)$	p	var.	β_{med}	$F(df_N, df_D)$	p	var.	β_{med}	$F(df_N, df_D)$	p	var.	β_{med}	$F(df_N, df_D)$	p
CCCR					alc	-0.70	8.6 (1, 14)	0.011*	alc	-0.72	14.0 (1, 14)	0.002**	cig	0.47	3.5 (1, 14)	0.083
L-PROJ					mat_age	0.56	5.5 (1, 14)	0.034*	mat_age	0.53	6.3 (1, 14)	0.025*				
	cig	0.12	4.2 (11, 4)	0.091	alc	-0.41	3.9 (10, 140)	0.000***	alc	-0.52	4.1 (10, 140)	0.000***	cig	0.52	4.0 (1, 14)	0.066
R-PROJ					mat_age	0.37	4.4 (1, 14)	0.056	mat_age	0.44	6.5 (1, 14)	0.023*				
	alc				alc	-0.41	1.9 (12, 168)	0.035*	alc	-0.45	2.7 (12, 168)	0.002**	cig	0.48	3.4 (1, 14)	0.085
	age	0.33	8.6 (13, 2)	0.109	age	-0.41	5.8 (1, 14)	0.031*	age	-0.39	5.3 (1, 14)	0.038*				
L-ASSOC	sex				sex	-0.20	4.3 (1, 14)	0.056	sex	-0.39	5.9 (1, 14)	0.029*				
	mat_age	-0.16	9.2 (13, 2)	0.103	alc	-0.65	6.0 (7, 8)	0.011*	alc	-0.66	8.1 (1, 14)	0.013*	cig	0.49	3.6 (1, 14)	0.080
					mat_age	0.44	3.8 (1, 14)	0.071	age	-0.16	2.5 (6, 84)	0.030*				
R-ASSOC					mat_age	0.43	4.7 (1, 14)	0.048*	mat_age	0.43	4.7 (1, 14)	0.048*				
	alc	0.23	1.8 (7, 98)	0.090	alc	-0.62	10.2 (1, 14)	0.007**	alc	-0.67	14.1 (1, 14)	0.002**	cig	0.5	3.5 (1, 14)	0.082
									cig	-0.29	3.9 (1, 14)	0.068				

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

→ Increased alcohol exposure:
decreased AD
(and decreased MD)

III) Results: ROI level

The question:

1) where are most significant AD-alcohol relations in each network?

Answer using:

Follow-up GLT for each WMC:

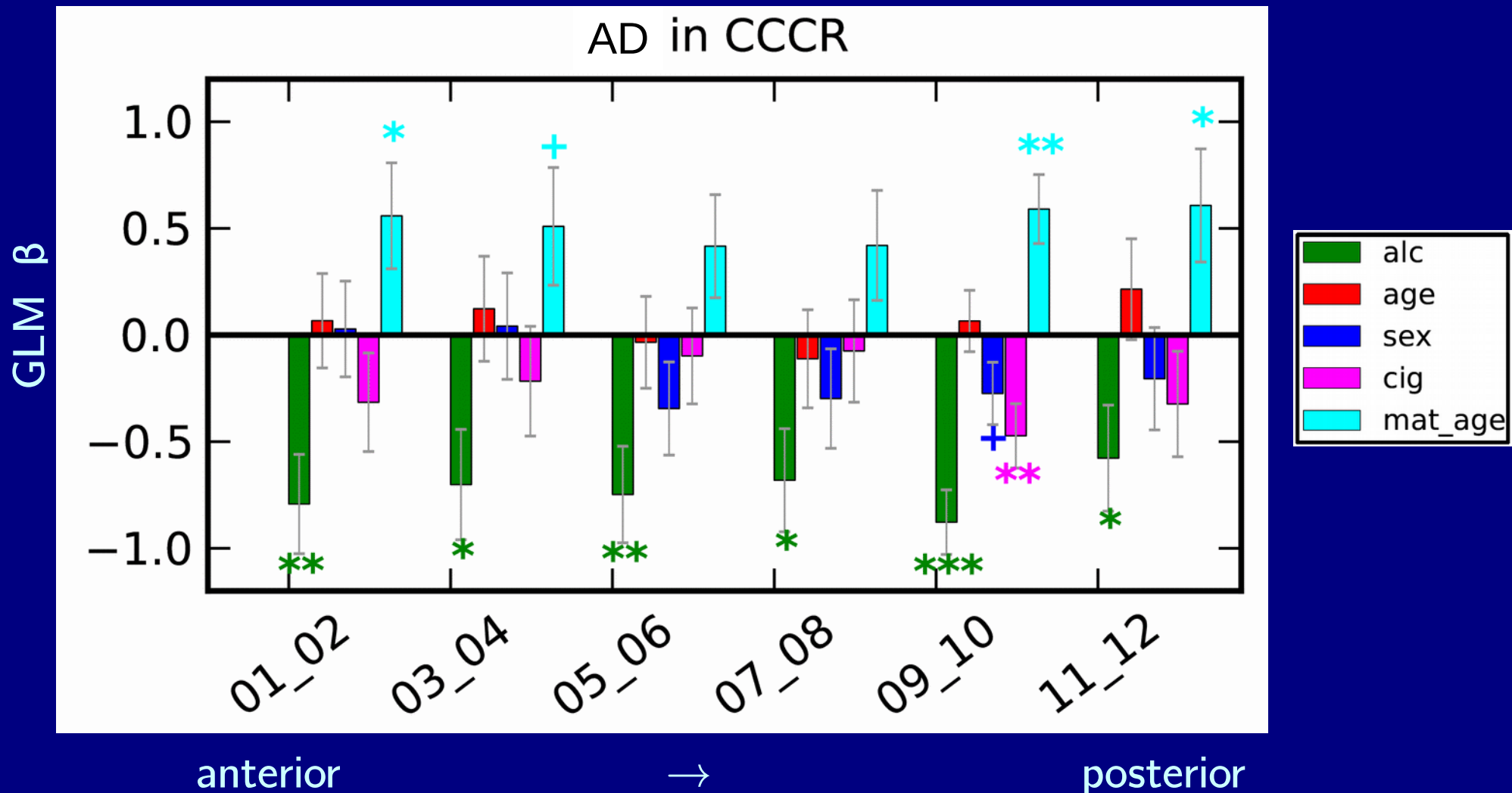
{set of DTI parameters} ~
alcohol (frequency: binge/wk) +
infant age (wks since conception) +
infant sex (M/F) +
maternal age (yrs) +
maternal cigarette smoking (cig/day).

III) Results: ROI level

The question:

1) where are most significant AD-alcohol relations in each network?

Transcallosal (CC and corona radiata)

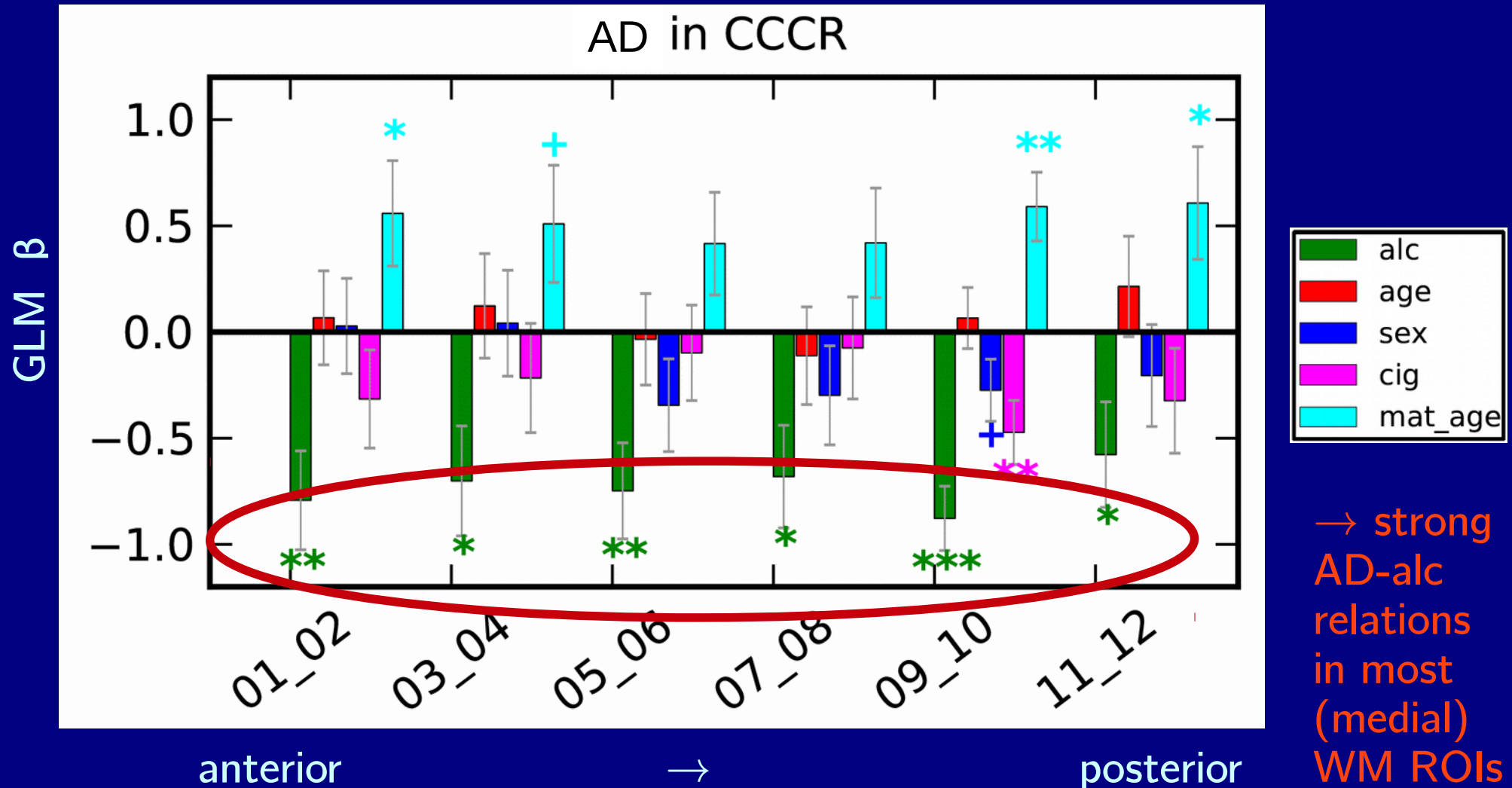


III) Results: ROI level

The question:

1) where are most significant AD-alcohol relations in each network?

Transcallosal (CC and corona radiata)



SUMMARY

- + Tracking allows one to compare and investigate properties first at a network level, and then “zoomed in” at WMC level
 - Same applies for FC matrices (e.g., from 3dNetCorr)
- + MVM modeling provides omnibus F -statistic for network level, and GLTs for follow-up
- + FATCAT functions help combine MRI data (*.grid or *.netcc files) with subject characteristics (*.xls -> *.cxv file)
- + Additional functions help specify the model for 3dMVM

