

# Population-Level Analysis: Hands-On

Gang Chen

Scientific and Statistical Computing Core  
National Institute of Mental Health  
National Institutes of Health  
U.S.A.



National Institute  
of Mental Health



# Prerequisites

- Directory for scripts & data: [AFNI\\_data6/GroupAna\\_cases/](#)
- R packages
  - ★ afex, phia, snow, lme4, gamm4, brms
  - ★ installing packages at the terminal
    - `rPkgsInstall -pkgs ALL`
    - `rPkgsInstall -pkgs 'gamm4'`
    - `rPkgsInstall -pkgs 'brms'`
  - ★ installing packages in R
    - `install.packages('gamm4')`

## Program list: population level

Spatial Unit	Program	Model
voxel, node, ROI  massively univariate	<a href="#">3dttest++</a>	$t$ -tests, GLM
	<a href="#">3dMEMA</a>	<a href="#">3dttest++</a> analog with $\beta + t$ as input
	<a href="#">3dMVM</a>	GLM, AN(C)OVA
	<a href="#">3dLMEr</a>	LME: hierarchical modeling
	<a href="#">3dMSS</a>	multilevel smoothing splines: nonlinearity
	<a href="#">3dISC</a>	inter-subject correlation: naturalistic data
	<a href="#">3dICC</a>	intra-class correlation: reliability
ROI	<a href="#">RBA</a>	region-based analysis: Bayesian modeling
	<a href="#">PTA</a>	profile trajectory analysis: nonlinearity
	<a href="#">TRR</a>	test-retest reliability

Others: 3dANOVA, 3dANOVA2, 3dANOVA3, 3dRegAna, 3dLME

# Writing scripts

- demo 3dMVM script

```
3dMVM    -prefix result.MVM -jobs 2          \  
        -wsVars 'Affect*Category'          \  
        ...  
Subj    Affect  Category      InputFile      \  
ss03    Neg     A             data/ss03.a.neg+tlrc  \  
ss03    Neg     B             data/ss03.b.neg+tlrc  \  
ss03    Pos     A             data/ss03.a.pos+tlrc  \  
ss03    Pos     B             data/ss03.b.pos+tlrc  \  
...
```

- Model specification: R notation
  - ★ population level:  $A*B = 1 + A + B + A:B$
  - ★ lower levels:  $(1|Subj)$ ,  $(1|A:Subj)$ ,  $(1|B:Subj)$
- tool for simple diagnosis:  
file\_tool -test -infile ...  
file\_tool -show\_file\_type -infile bad.txt -prefix ...

# Data table

- similar to long format in R
  - ★ columns: variables
    - 1st column (**Subj**): subject IDs
    - last column (**InputFile**): input file list (y)
    - other columns: explanatory variables
    - **factors**: at least 1 character
    - **quantitative variables**: declared with **-qVars**
  - ★ rows: variable info for each combination
  - ★ data table as part of the script
    - requiring **backslash** \ on the end of each line
    - requiring **quotes** surrounding sub-brick selector: '[Pos#0\_Coef]'
  - ★ data table as a separate file
    - important for many rows
    - same table can be used for different models
    - backslash on each line: not needed
    - **NO quotes** surrounding sub-brick selector
  - ★ **no missing** info allowed
  - ★ **rectangularity**: a table checker will be available

# Case 1: one-way within-individual ANOVA with missing data

- `tssh run.LMEr.txt > LMEr.log &`
- background info: MEG data
  - ★ 17 participants
  - ★ 3 conditions: Baseline, Ketamine, Placebo
  - ★ missing data: missing unrelated to task conditions?
- Effects of interest
  - ★ 3 condition effects
  - ★ 3 pairwise comparisons
- Analysis approaches
  - ★ 1-way within-subject ANOVA: losing 6 individuals (`3dMVM`)
  - ★ 3 pairwise *t*-tests: losing 2-6 individuals (`3dttest++`)
  - ★ 1 hierarchical model: `3dLMEr`
- Result visualization

Subj	Baseline	Ket	Placebo
S101	1	1	0
S102	1	1	1
S105	1	1	1
S107	1	1	1
S108	1	1	1
S109	1	1	1
S110	1	1	1
S111	1	1	0
S112	0	1	1
S113	1	1	1
S115	0	1	1
S116	1	1	0
S118	1	1	1
S120	1	1	1
S121	1	1	0
S122	1	1	1
S123	1	1	1

## Case 2: one-way between-individual ANOVA

- Data info
  - ★ COMT (catechol-O-methyl transferase) gene for schizophrenia: Val/Met (valine-to-methionine) polymorphism
  - ★ 31 individuals among 3 genotypic groups: Val/Val (12), Val/Met (10), Met/Met (9)
  - ★ demo with  $3 \times 3 \times 3$  voxels
- Effects of interest
  - ★ 3 group effects
  - ★ 3 pairwise comparisons
- Analysis approach
  - ★ Student's  $t$ -tests: `3dtttest++`
  - ★ 1-way between-individual ANOVA: `3dMVM`
    - executing script: `tcsh run.5_MVM_3grps.txt > MVM.log &`

## Case 3: 4-way ANOVA with mixed factors

- Data info:  $2 \times 2 \times 2 \times 2$  linguistic study
  - ★ 1 **between-individual factor**: group (young and old)
  - ★ 3 **within-individual factors**
    - task: perception and production
    - syllable: simple and complex
    - sequence: simple and complex
  - ★ demo with 20 nodes on surface
- Analysis approaches
  - ★ multiple Student's *t*-tests: `3dtttest++`
  - ★ ANOVA
    - **3dMVM**: script `run.MVM.txt`
  - ★ hierarchical model
    - population:  $A*B*C=1+A+B+C+A:B+A:C+B:C+A:B:C$
    - lower level: (1|A:Subj), (1|A:Subj), (1|B:Subj), (1|C:Subj), (1|A:B:Subj), (1|A:C:Subj), (1|B:C:Subj)
    - **3dLMEr**: script `run.LMEr.txt`
    - -model  
'Group\*Condition\*Image\*Clarity+(1|Subj)+(1|Condition:Subj)+(1|Image:Subj)+(1|Clarity:Subj)+(1|Condition:Clarity:Subj)+(1|Image:Clarity:Subj)'



## Case 4: 2-way ANOVA with within-individual factors

- Data info:  $3 \times 2$  factorial design
  - ★ 2 within-individual factors
    - category: animal (A), human (H), and tool (T)
    - emotion: negative (Neg), and neutral (Neu)
  - ★ demo with  $3 \times 3 \times 3$  voxels
- Analysis approaches
  - ★ multiple Student's  $t$ -tests: `3dtttest++`
  - ★ ANOVA with `3dMVM`: script `run.MVM.txt`
  - ★ hierarchical model
    - population:  $A*B=1+A+B+A:B$
    - lower level:  $(1|Subj)$ ,  $(1|A:Subj)$ ,  $(1|B:Subj)$
    - `3dLMER`: script `run.LMER.txt`
    - `-model 'Affect*Category+(1|Subj)+(1|Affect:Subj)+(1|Category:Subj)'`

## Case 5: Group comparisons with estimated HDRs

- Data info: selective attention study
  - ★ **between-individual factor** - group: 44 controls (CG) & 43 patients (PG)
  - ★ estimated HDRs under incongruent condition at the individual level: 12 time points; **TR:** 1.25 s; `stim_times ... tent(0,13.75,12)`
- Demo with  $3 \times 3 \times 3$  voxels
- Analysis approach: multilevel smooth splines
  - ★ **3dMSS:** `tcsch run.MSS.txt > MSS.log &`
- Result visualization
  - ★ HDR profile with uncertainty
  - ★ HDR profile comparison between two groups

## Case 6: profile tracking analysis

- Data info
  - ★ 16 participants
  - ★ 11 layers: VASO (vascular space occupancy) data
- Analysis approaches: multilevel smooth splines - PTA
  - ★ nonlinear modeling with PTA: `tssh run.pta1.txt`
  - ★ nonlinear modeling with PTA (with separate linear trend): `tssh run.pta2.txt`
- Checking results
- Result visualization

# Case 7: region-based analysis

- execute: `tcsh run.RBA2.txt > RBA2.diary &`
  - ★ requiring R package `brms`
- Data info: resting-state
  - ★ 124 children
  - ★ 21 regions of Z-score values; `ToMI.txt`
- Analysis approaches
  - ★ conventional mass univariate: one model per region
  - ★ hierarchical modeling `RBA`
- Result visualization: `run.plot.R`

