Group Analysis

File: **afni24_GroupAna.pdf**

Gang Chen

SSCC/NIMH/NIH

Program List

- **3dttest++** (GLM: one-, two-sample, paired *t*, between-subjects variables)
- **3dMVM** (generic AN(C)OVA)
- **3dLME** (sophisticated cases: missing data, within-subject covariates)
- **3dMEMA** (similar to 3dttest++: measurement errors)
- **3dANOVA** (one-way between-subject)
- **3dANOVA2** (one-way within-subject, 2-way between-subjects)
- **3dANOVA3** (2-way within-subject and mixed, 3-way between-subjects)
- **3dttest** (obsolete: one-sample, two-sample and paired *t*)
- **3dRegAna** (obsolete: regression/correlation, covariates)
- **GroupAna** (obsolete: up to four-way ANOVA)
- **3dICC** (intraclass correlation): prototype only
- **3dISC** (intersubject correlation): prototype only

Preview

- Concepts and terminology

- Group analysis approaches
  - GLM: 3dttest++, 3dMEMA
  - GLM, ANOVA, ANCOVA: 3dMVM
  - LME: 3dLME
  - Presumed vs. estimated HDR

- Miscellaneous
  - Issues with covariates
  - Intra-Class Correlation (ICC)
  - Inter-Subject Correlation (ISC)

Why Group Analysis?

- Reproducibility and generalization
  - Summarization
  - Generalization: from current results to population level
  - Typically 10 or more subjects per group
  - Individualized inferences: pre-surgical planning, lie detection, …

- One model combining both steps?
  - + Ideal: less information loss, more accurate inferences
  - - Historical
  - - Computationally unmanageable, and very hard to set up
  - - Data quality check at individual level

Simplest case

- BOLD responses from a group of 20 subjects
  - data: $(\beta_1, \beta_2, \ldots, \beta_{20})$=(1.13, 0.87, …, 0.72)
  - mean: 0.92
  - standard deviation: 0.40, 0.90
  - Do we have strong evidence for the effect?

- Modeling perspective
  - Simple GLM: one-sample *t*-test

$$\hat{\beta}_i = b + \epsilon_i, \epsilon_i \sim N(0, \sigma^2)$$

  - Statistical evidence - *t*-test: $\hat{b} / \left(\hat{\sigma} / n\right)$
  - summarization: ***b*** (dimensional), sd, and ***t*** (dimensionless)

Terminology

- Response/outcome variable: left-hand side of model
  - Regression β coefficients (plus measurement errors)
  - Structured: subjects, tasks, groups

- Explanatory variables: right-hand side of model
  - Categorical (factors) vs quantitative (covariates)
  - Fixed- vs random-effects: conventional statistics

- Models
  - Univariate GLM: Student's *t*-tests, regression, AN(C)OVA
  - Multivariate GLM: within-subject factors
  - LME: linear mixed-effects model
  - MEMA: mixed-effects multilevel analysis
  - BML (Bayesian multilevel model)

Terminology: categorical vs quantitative

- Factors
  - Number of levels: categories
  - Within-subject (repeated-measures): tasks, conditions
  - Between-subjects
    - patients/controls, genotypes, scanners/sites, handedness, …
    - Each subject nested within a group
  - Subjects: random-effects factor - measuring randomness
    - Of no interest: random samples from a population

- Quantitative variables
  - numeric or continuous
  - age, IQ, reaction time, brain volume, …
  - 3 usages of covariate
    - Quantitative
    - No interest: qualitative (scanner/site, groups) or quantitative
    - Explanatory variable

Terminology: fixed vs random

- Fixed-effects variables
  - Of research interest
    - Visual vs auditory, age, …
    - Unable to extend to something else
  - Modeled as constants, not random variables
    - Shared by all subjects
  - Not exchangeable/replaceable or extendable to something else

- Random-effects variables
  - Of research interest?
    - Subjects: random samples
    - Trials, regions?
  - Modeled as random variables: Gaussian distributions
  - Exchangeable, replaceable, generalizable

$$\hat{\beta}_i = b + \epsilon_i, \epsilon_i \sim N(0, \sigma^2)$$

- Differentiations blurred under BML

Terminology: main effects

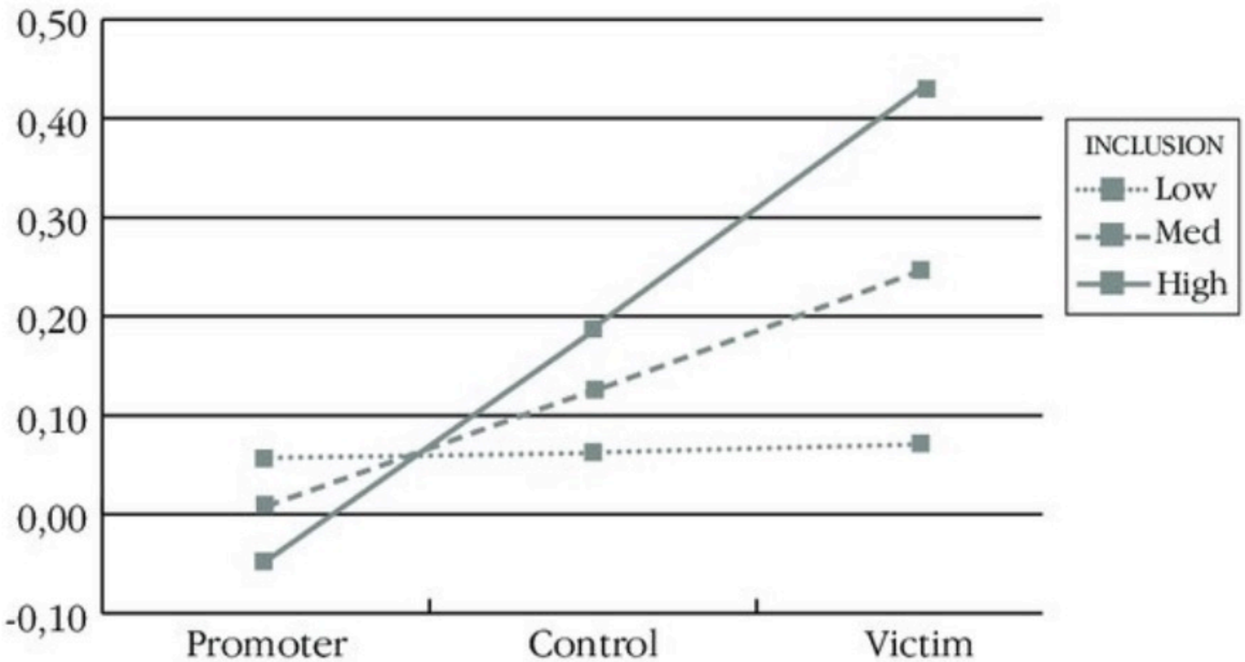
- Main effect for a fixed-effects factor
  - Omnibus: overall inference or summarization
    - Evidence for differences across 3 levels
    - Conventional ANOVA framework
    - *F*-statistic: not detailed enough
    - Further partitions: post hoc inferences via pairwise comparisons
    - *F*-statistic as a two-sided test?

      1) A > B, 2) A < B, 3) A $\neq$ B

Terminology: interactions

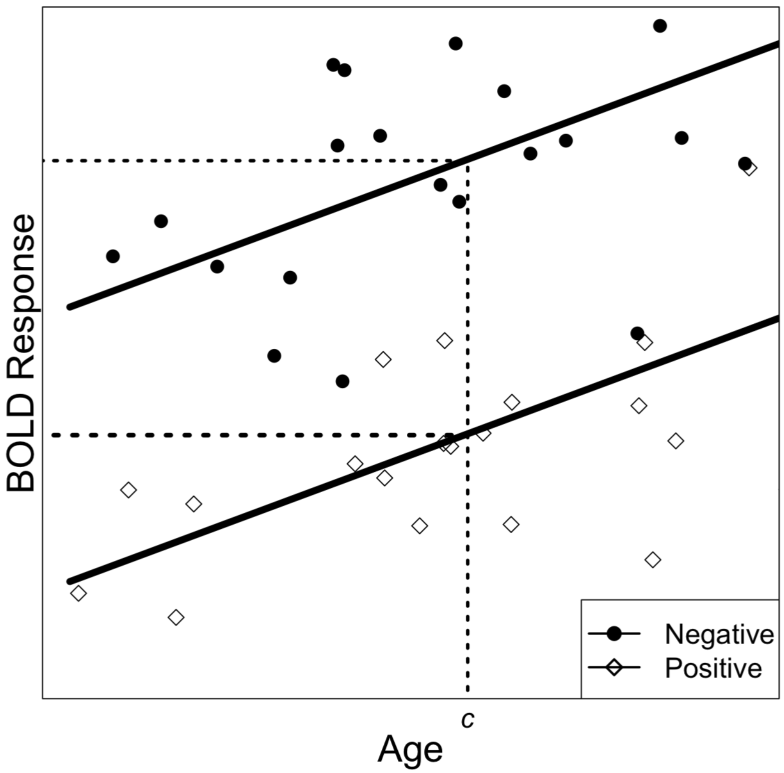
- Interaction effect between 2 or more factor
  - Omnibus: overall inference or summarization
    - Conventional ANOVA framework
    - *F*-statistic: not detailed enough
    - Further partitions: post hoc inferences via pairwise comparisons
  - 2 × 2 design: difference of difference
    - *F*-test for interaction = *t*-test of

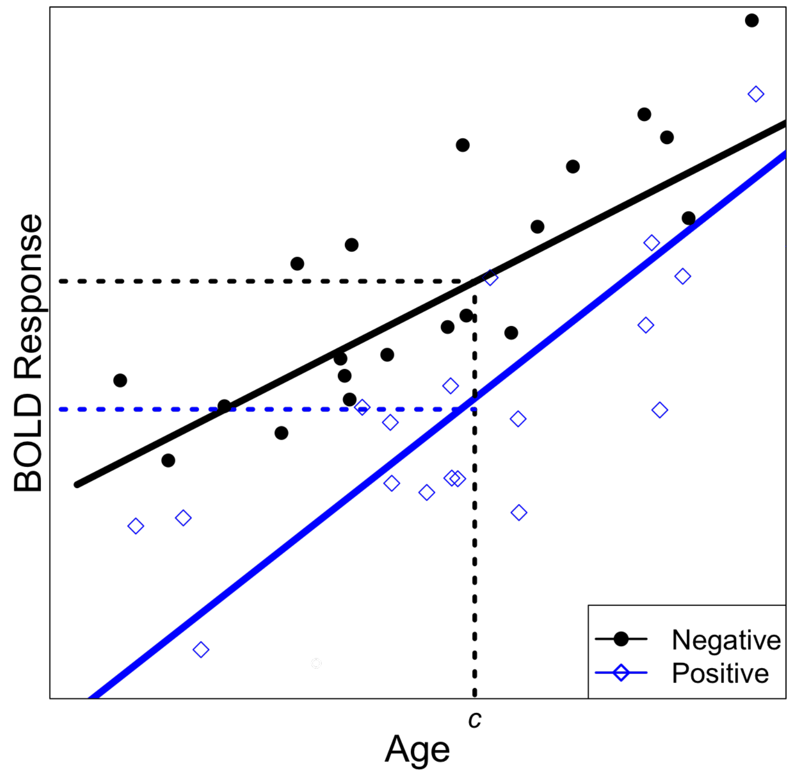
(A1B1 - A1B2) - (A2B1 - A2B2) or (A1B1 - A2B1) - (A1B2 - A2B2)



Terminology

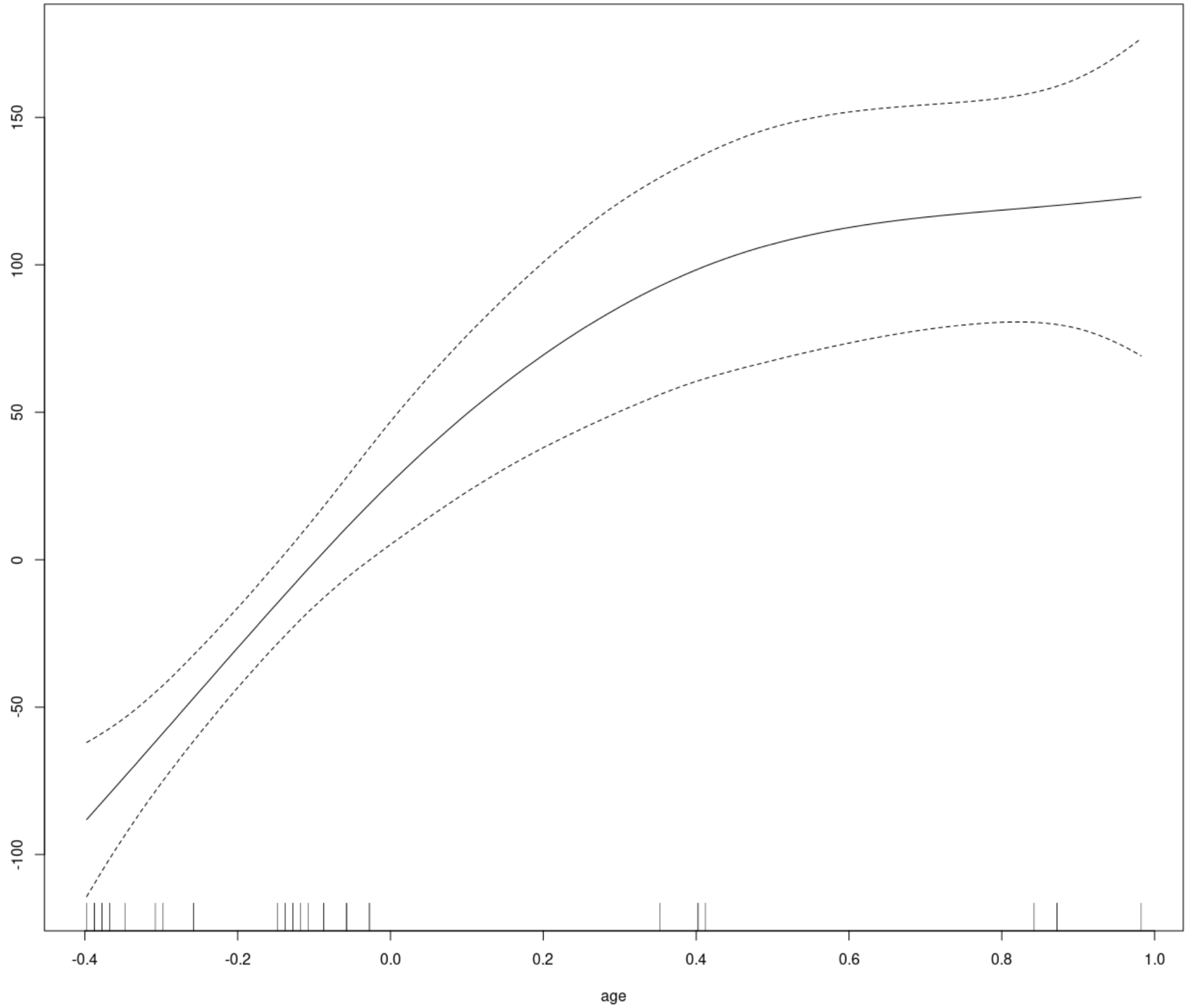
- Interaction effect involving a quantitative variable
  - By default: linearity (age, modulation, …)
    - Controlling: misconception - covariate out?
    - Effect of interest
  - Interaction between a factor and a quantitative variable





Terminology

- Interaction effect involving a quantitative variable
  - Validity of linearity
    - Nonlinear: difficult! Polynomials? Theory-driven?



Example: 2 × 3 Mixed ANCOVA

- Explanatory variables
  - Factor A (Group): 2 levels (patient and control)
  - Factor B (Condition): 3 levels (pos, neg, neu)
  - Factor S (Subject): 15 ASD children and 15 healthy controls
  - Quantitative covariate: Age

- Piecemeal: multiple *t*-tests – too tedious
  - Group comparison + age effect
  - Pairwise comparisons among three conditions
    - Assumption: same age effect across conditions
  - Difficulties with *t*-tests
    - Main effect of Condition: 3 levels plus age?
    - Interaction between Group and Condition
    - Age effect across three conditions?

Classical ANOVA: 2 × 3 Mixed ANOVA

- Factor A (Group): 2 levels (patient and control)
- Factor B (Condition): 3 levels (pos, neg, neu)
- Factor S (Subject): 15 ASD children and 15 healthy controls
- Covariate (Age): **cannot** be modeled; **no** correction for sphericity violation

$$F_{(a-1,a(n-1))}(A) = \frac{MSA}{MSS(A)},$$

$$F_{(b-1,a(b-1)(n-1))}(B) = \frac{MSB}{MSE},$$

$$F_{((a-1)(b-1),a(b-1)(n-1))}(AB) = \frac{MSAB}{MSE}$$

Different denominator

where

$$MSA = \frac{SSA}{a-1} = \frac{1}{a-1}\left(\frac{1}{bn}\sum_{j=1}^{a} Y_{.j.}^2 - \frac{1}{abn}Y_{...}^2\right),$$

$$MSB = \frac{SSB}{b-1} = \frac{1}{b-1}\left(\frac{1}{an}\sum_{k=1}^{b} Y_{..k}^2 - \frac{1}{abn}Y_{...}^2\right),$$

$$MSAB = \frac{SSAB}{(a-1)(b-1)} = \frac{1}{(a-1)(b-1)}\left(\frac{1}{n}\sum_{j=1}^{a}\sum_{k=1}^{b} Y_{.jk} - \frac{1}{bn}\sum_{j=1}^{a} Y_{.j.}^2 - \frac{1}{an}\sum_{k=1}^{b} Y_{..k}^2 + \frac{1}{abn}Y_{...}^2\right),$$

$$MSS(A) = \frac{SSS(A)}{a(n-1)} = \frac{1}{a(n-1)}\left(\frac{1}{b}\sum_{i=1}^{n}\sum_{j=1}^{a} Y_{ij.}^2 - \frac{1}{bn}\sum_{j=1}^{a} Y_{.j.}^2\right),$$

$$MSE = \frac{1}{a(b-1)(n-1)}\left(\sum_{i=1}^{n}\sum_{j=1}^{a}\sum_{k=1}^{b} Y_{ijk}^2 - \frac{1}{n}\sum_{j=1}^{a}\sum_{k=1}^{b} Y_{.jk} - \frac{1}{b}\sum_{i=1}^{n}\sum_{j=1}^{a} Y_{ij.}^2 + \frac{1}{bn}\sum_{j=1}^{a} Y_{.j.}^2 + \frac{1}{abn}Y_{...}^2\right)$$

3dANOVA3 –type 5 (equal # of subjects across groups)

Univariate GLM: 2 x 3 mixed ANOVA

- Group: 2 levels (patient and control)
- Condition: 3 levels (pos, neg, neu)
- Subject: 3 ASD children and 3 healthy controls

Difficult to incorporate covariates

- Broken orthogonality of matrix

No correction for sphericity violation

$$
\begin{array}{c}
\text{Subj}\\ 1\\ 1\\ 1\\ 2\\ 2\\ 2\\ 3\\ 3\\ 3\\ 4\\ 4\\ 4\\ 5\\ 5\\ 5\\ 6\\ 6\\ 6
\end{array}
\quad
\underbrace{\begin{pmatrix}
\beta_{11}\\ \beta_{12}\\ \beta_{13}\\ \beta_{21}\\ \beta_{22}\\ \beta_{23}\\ \beta_{31}\\ \beta_{32}\\ \beta_{33}\\ \beta_{41}\\ \beta_{42}\\ \beta_{43}\\ \beta_{51}\\ \beta_{52}\\ \beta_{53}\\ \beta_{61}\\ \beta_{62}\\ \beta_{63}
\end{pmatrix}}_{b}
=
\underbrace{\begin{array}{c}
\begin{array}{cccccccccc} X_0 & X_1 & X_2 & X_3 & X_4 & X_5 & X_6 & X_7 & X_8 & X_9 \end{array}\\
\begin{pmatrix}
1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0\\
1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0\\
1 & 1 & -1 & -1 & -1 & -1 & 1 & 0 & 0 & 0\\
1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0\\
1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0\\
1 & 1 & -1 & -1 & -1 & -1 & 0 & 1 & 0 & 0\\
1 & 1 & 1 & 0 & 1 & 0 & -1 & -1 & 0 & 0\\
1 & 1 & 0 & 1 & 0 & 1 & -1 & -1 & 0 & 0\\
1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & 0 & 0\\
1 & -1 & 1 & 0 & -1 & 0 & 0 & 0 & 1 & 0\\
1 & -1 & 0 & 1 & 0 & -1 & 0 & 0 & 1 & 0\\
1 & -1 & -1 & -1 & 1 & 1 & 0 & 0 & 1 & 0\\
1 & -1 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 1\\
1 & -1 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 1\\
1 & -1 & -1 & -1 & 1 & 1 & 0 & 0 & 0 & 1\\
1 & -1 & 1 & 0 & -1 & 0 & 0 & 0 & -1 & -1\\
1 & -1 & 0 & 1 & 0 & -1 & 0 & 0 & -1 & -1\\
1 & -1 & -1 & -1 & 1 & 1 & 0 & 0 & -1 & -1
\end{pmatrix}
\end{array}}_{X}
\underbrace{\begin{pmatrix}
\alpha_0\\ \alpha_1\\ \alpha_2\\ \alpha_3\\ \alpha_4\\ \alpha_5\\ \alpha_6\\ \alpha_7\\ \alpha_8\\ \alpha_9
\end{pmatrix}}_{a}
+
\underbrace{\begin{pmatrix}
\delta_{11}\\ \delta_{12}\\ \delta_{13}\\ \delta_{21}\\ \delta_{22}\\ \delta_{23}\\ \delta_{31}\\ \delta_{32}\\ \delta_{33}\\ \delta_{41}\\ \delta_{42}\\ \delta_{43}\\ \delta_{51}\\ \delta_{52}\\ \delta_{53}\\ \delta_{61}\\ \delta_{62}\\ \delta_{63}
\end{pmatrix}}_{d}
$$

Univariate GLM: problematic implementations

Two-way mixed ANOVA

Between-subjects Factor A (Group): 2 levels (patient, control)

Within-subject Factor B (Condition): 3 levels (pos, neg, neu)

1) Omnibus tests

$$F_A = \frac{MSA}{MSA(C)}, \quad F_B = \frac{MSB}{MSE}, \quad F_{AB} = \frac{MSAB}{MSE}$$

Correct

$$F_A = \frac{MSA}{MSE}, \quad F_B = \frac{MSB}{MSE}, \quad F_{AB} = \frac{MSAB}{MSE}$$

Incorrect

2) Post hoc tests (contrasts)

- **Incorrect** *t*-tests for factor A due to incorrect denominator
- **Incorrect** *t*-tests for factor B or interaction effect AB when weights do not add up to 0

Univariate GLM: **problematic implementations**

Two-way repeated-measures ANOVA

Within-subjects Factor A (Object): 2 levels (house, face)

Within-subject Factor B (Condition): 3 levels (pos, neg, neu)

1) Omnibus tests

Correct

$$F_A = \frac{MSA}{MSAC}, \quad F_B = \frac{MSB}{MSBC}, \quad F_{AB} = \frac{MSAB}{MSE}$$

Incorrect

$$F_A = \frac{MSA}{MSE}, \quad F_B = \frac{MSB}{MSE}, \quad F_{AB} = \frac{MSAB}{MSE}$$

2) Post hoc tests (contrasts)

- **Incorrect** *t*-tests for both factors A and B due to incorrect denominator
- **Incorrect** *t*-tests for interaction effect AB if weights don't add up to 0

Better Approach: Multivariate GLM

- Group: 2 levels (patient and control)
- Condition: 3 levels (pos, neg, neu)
- Subject: 3 ASD children and 3 healthy controls
- Age: quantitative covariate

$$\boldsymbol{B}_{n\times m} = \boldsymbol{X}_{n\times q}\,\boldsymbol{A}_{q\times m} + \boldsymbol{D}_{n\times m}$$

$$\begin{array}{c} Subj \\ \begin{matrix}1\\2\\3\\4\\5\\6\end{matrix}\end{array}
\begin{array}{c} \begin{matrix}Pos & Neg & Neu\end{matrix} \\ \underbrace{\begin{pmatrix}\beta_{11} & \beta_{12} & \beta_{13}\\ \beta_{21} & \beta_{22} & \beta_{23}\\ \beta_{31} & \beta_{32} & \beta_{33}\\ \beta_{41} & \beta_{42} & \beta_{43}\\ \beta_{51} & \beta_{52} & \beta_{53}\\ \beta_{61} & \beta_{62} & \beta_{63}\end{pmatrix}}_{B}\end{array}
=
\begin{array}{c} \begin{matrix}Int & Grp & Age\end{matrix} \\ \underbrace{\begin{pmatrix}1 & 1 & -6\\ 1 & 1 & 10\\ 1 & 1 & 4\\ 1 & -1 & -4\\ 1 & -1 & -1\\ 1 & -1 & -3\end{pmatrix}}_{X}\end{array}
\begin{array}{c} \begin{matrix}Pos & Neg & Neu\end{matrix} \\ \underbrace{\begin{pmatrix}\alpha_{01} & \alpha_{02} & \alpha_{03}\\ \alpha_{11} & \alpha_{12} & \alpha_{13}\\ \alpha_{21} & \alpha_{22} & \alpha_{23}\end{pmatrix}}_{A}\end{array}
+
\begin{array}{c} \begin{matrix}Pos & Neg & Neu\end{matrix} \\ \underbrace{\begin{pmatrix}\delta_{11} & \delta_{12} & \delta_{13}\\ \delta_{21} & \delta_{22} & \delta_{23}\\ \delta_{31} & \delta_{32} & \delta_{33}\\ \delta_{41} & \delta_{42} & \delta_{43}\\ \delta_{51} & \delta_{52} & \delta_{53}\\ \delta_{61} & \delta_{62} & \delta_{63}\end{pmatrix}}_{D}\end{array}
\begin{array}{c} Subj \\ \begin{matrix}1\\2\\3\\4\\5\\6\end{matrix}\end{array}$$

MVM Implementation in AFNI

- Program **3dMVM**
  - No dummy coding needed!
  - **Symbolic coding** for variables and post hoc testing

Variable types

Post hoc tests

```
3dMVM  -prefix       OutputFile   -jobs 8      -SC                  \
       -bsVars       'Grp*Age'    -wsVars      'Cond'   -qVars 'Age' \
       -num_glt 4                                                    \
       -gltLabel 1   Pat_Pos      -gltCode 1   'Grp : 1*Pat Cond : 1*Pos'        \
       -gltLabel 2   Ctl_Pos-Neg  -gltCode 2   'Grp : 1*Ctl Cond : 1*Pos -1*Neg' \
       -gltLabel 3   GrpD_Pos-Neg -gltCode 3   'Grp : 1*Ctl -1*Pat Cond : 1*Pos -1*Neg' \
       -gltLabel 4   Pat_Age      -gltCode 4   'Grp : 1*Pat Age :'              \
       -dataTable                                                    \
       Subj          Grp          Age          Cond     InputFile    \
       S1            Ctl          23           Pos      S1_Pos.nii   \
       S1            Ctl          23           Neg      S1_Neg.nii   \
       S1            Ctl          23           Neu      S1_Neu.nii   \
       ...
       S50           Pat          19           Pos      S50_Pos.nii  \
       S50           Pat          19           Neg      S50_Neg.nii  \
       S50           Pat          19           Neu      S50_Neu.nii
```

Data layout

Improvement 1: precision information

- Conventional approach: βs as response variable
  - Assumptions
    - no measurement errors
    - all subjects have same precision
  - All subjects are treated equally
- More precise method: βs plus precision
  - *t*-statistic contains precision
  - βs and their *t*-stats as input
  - βs weighted based on precision
  - Only available for GLM types: 3dMEMA
  - Regions with substantial cross-subject variability
- Best approach: combining all subjects in one big model
  - Currently not feasible

One group: Example

- **3dttest++**: β as input only

```
3dttest++ -prefix Vis -mask mask+tlrc -zskip        \
 -setA 'FP+tlrc[Vrel#0_Coef]'                        \
       'FR+tlrc[Vrel#0_Coef]'                        \
       ……
       'GM+tlrc[Vrel#0_Coef]'
```

Voxel value = 0 ➔ treated it as missing

- **3dMEMA**: β and t-statistic as input

```
3dMEMA -prefix VisMEMA -mask mask+tlrc -setA Vis          \
 FP 'FP+tlrc[Vrel#0_Coef]' 'FP+tlrc[Vrel#0_Tstat]'        \
 FR 'FR+tlrc[Vrel#0_Coef]' 'FR+tlrc[Vrel#0_Tstat]'        \
       ……
 GM 'GM+tlrc[Vrel#0_Coef]' 'GM+tlrc[Vrel#0_Tstat]'        \
 -missing_data 0
```

Voxel value = 0 ➔ treated it as missing

Paired comparison: Example

- **3dttest++**: comparing two conditions

```
3dttest++ -prefix Vis_Aud          \
  -mask mask+tlrc -paired -zskip \
  -setA 'FP+tlrc[Vrel#0_Coef]'    \
        'FR+tlrc[Vrel#0_Coef]'    \
        ……
        'GM+tlrc[Vrel#0_Coef]'    \
  -setB 'FP+tlrc[Arel#0_Coef]'    \
        'FR+tlrc[Arel#0_Coef]'    \
        ……
        'GM+tlrc[Arel#0_Coef]'
```

Paired Comparison: Example

- 3dMEMA: accounting for differential accuracy
  - Contrast as input

```
3dMEMA -prefix Vis_Aud_MEMA                                    \
 -mask mask+tlrc -missing_data 0                               \
 -setA Vis-Aud                                                 \
FP 'FP+tlrc[Vrel-Arel#0_Coef]' 'FP+tlrc[Vrel-Arel#0_Tstat]' \
FR 'FR+tlrc[Vrel-Arel#0_Coef]' 'FR+tlrc[Vrel-Arel#0_Tstat]‘ \
 ……
GM 'GM+tlrc[Vrel-Arel#0_Coef]''GM+tlrc[Vrel-Arel#0_Tstat]'
```

Improvement 2: more accurate HDR

- Conventional approach $f(t) = t^q e^{-t}/(q^q e^{-q})$ (q=4)
  - Presumed curve (empirical and approximate): BLOCK(d,1)
  - Fixing HDR shape and capturing magnitude with one number
  - Simple and straightforward: one β per effect
  - Not ideal: HDR varies across regions, tasks/conditions, groups, subjects
- More accurate HDR modeling
  - Data driven (no assumptions about HDR shape): TENTzero, CSPLINzero
  - Estimating both shape and magnitude with multiple effect estimates
  - More complicated: multiple βs per task/condition
  - More challenging: how to make inferences? H_0: β_1=0, β_2=0, …, β_k=0
- Middle
  - Adjust major HDR curve with 2/3 auxiliary functions: SPMG2/3
  - Focus: magnitude (β) associated with major HDR curve

Improvement 2: more accurate HDR

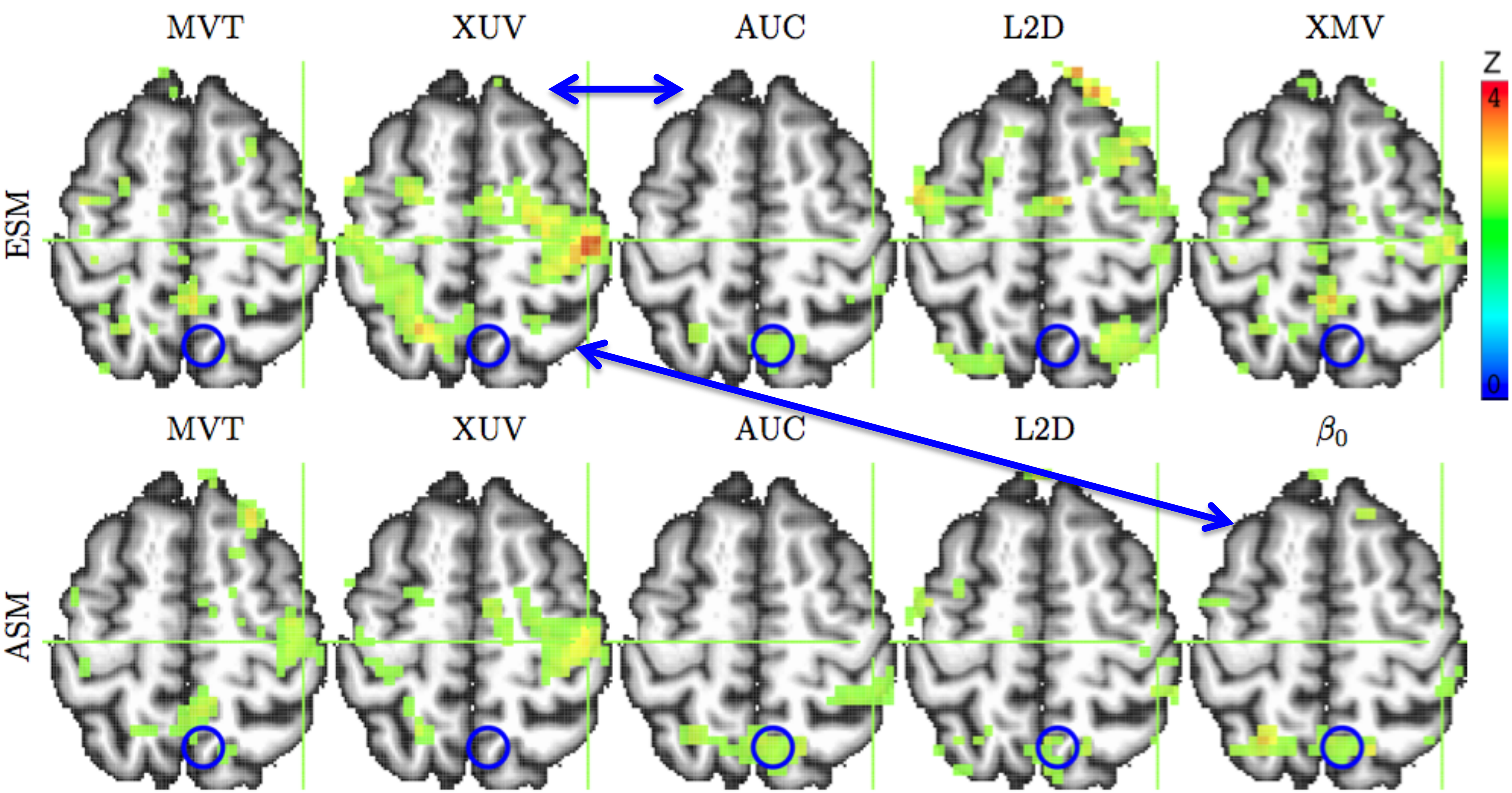
- Group analysis with HDR estimates: TENTzero, CSPLINzero
  - NHST: H_0: β_1=0, β_2=0, …, β_k=0
  - Area under curve (AUC) approach
    - Reduce to one number: use area as magnitude approximation
    - Ignore **shape** subtleties
    - Shape information loss: (undershoot, peak location/width)
  - Better approach: maintaining shape integrity
    - Take individual βs to group analysis (MVM)
    - One group with one condition: 3dLME
    - Other scenarios: treat βs as levels of a factor (e.g., Time) - 3dMVM

    ** Task or group effect: *F*-stat for interaction between task group and Time, complemented with main effect for task/group (AUC)

Chen et al. (2015). Detecting the subtle shape differences in hemodynamic responses at the group level. Front. Neurosci., 26 October 2015.

Improvement 2: more accurate HDR

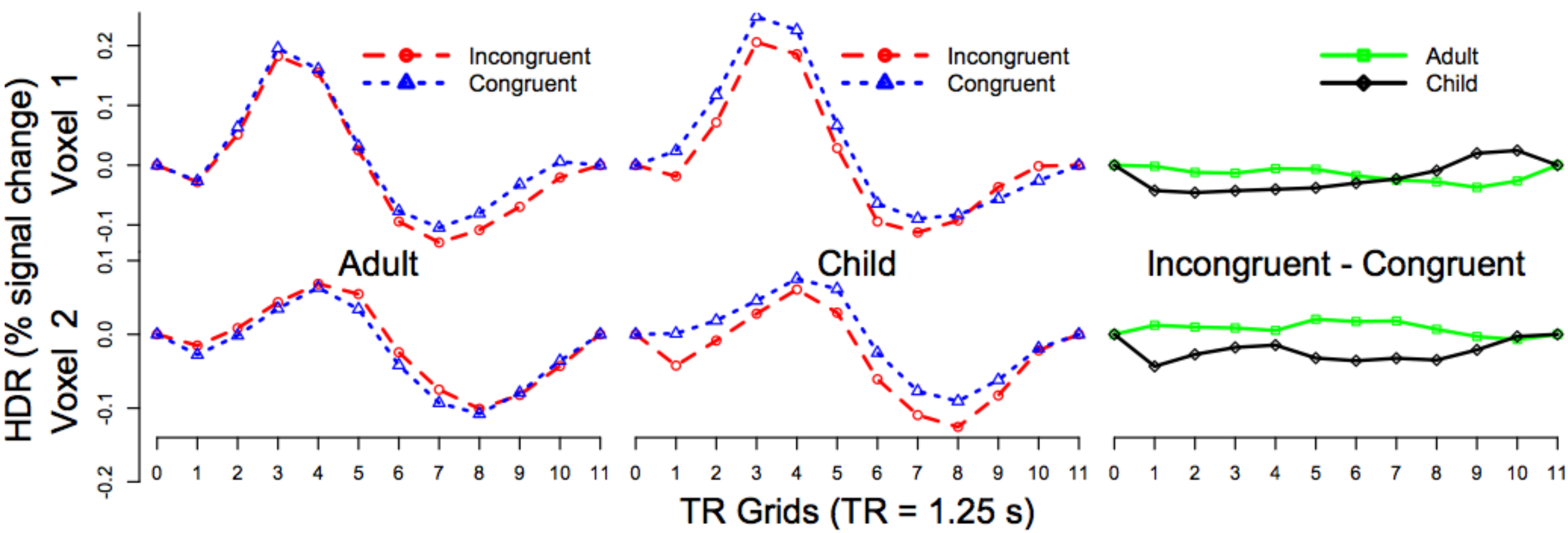
- 2 groups (children, adults), 2 conditions (congruent, incongruent), 1 quantitative covariate (age)
- 2 methods: HRF modeled by 10 (tents) and 3 (SPMG3) bases
- Effect of interaction: interaction group:condition – 3dMVM



Improvement 2: more accurate HDR

- Advantages of ESM over FSM
  - More likely to detect HDR shape subtleties
  - Visual verification of HDR signature shape (vs. relying significance testing: *p*-values)

Study: Adults/Children with Congruent/Incongruent stimuli (2 × 2)



Dealing with quantitative variables

- Reasons to consider a covariate
  - Effect of interest
  - Model improvement: accounting for data variability
- Frameworks
  - ANCOVA: between-subjects factor (e.g., group) + quantitative variable
  - Broader frameworks: regression, GLM, MVM, LME, BML
  - Assumptions: linearity, homogeneity of slopes (interaction)
- Interpretations
  - Effect of interest: slope, rate, marginal effect
  - Regress/covariate out x? head motion at individual level
  - "Controlling x at …", "holding x constant": centering

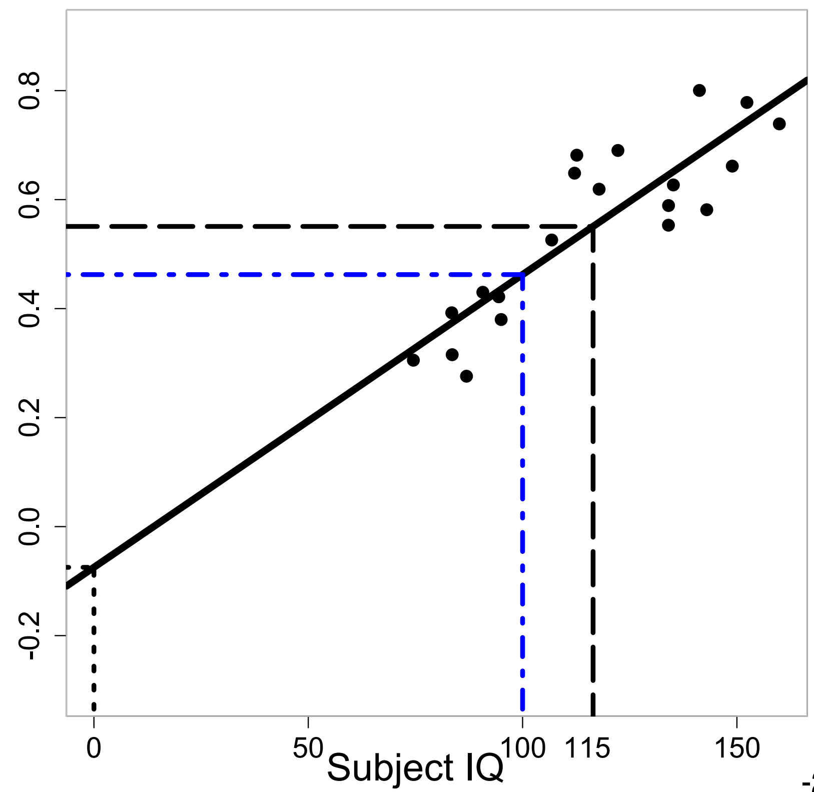
Quantitative variables: centering

- Model

$$\hat{\beta}_i = \alpha_0 + \alpha_1 * x_{1i} + \alpha_2 * x_{2i} + \epsilon_i$$

  - α_1, α_2 - slope
  - α_0 – intercept: group effect when ***x*=0**
    - Not necessarily meaningful
    - Linearity may not hold
    - Centering for interpretability
    - Mean or median centering?

- When a factor is involved
  - Complicated: within-level or grand centering



https://afni.nimh.nih.gov/pub/dist/doc/htmldoc/STATISTICS/center.html

IntraClass Correlation (ICC)

- Reliability (consistency, agreement/reproducibility) across two or more measurements of a condition/task
  - sessions, scanners, sites, studies, twins
  - Classic example (Shrout and Fleiss, 1979): *n* targets are rated by *k* raters
  - Relationship with Pearson correlation
    - Pearson correlation: two different types of measure: e.g., BOLD response vs. RT
    - ICC: **same** measurement
  - Modeling frameworks: ANOVA, LME
  - 3 types ICC: ICC(1,1), **ICC(2,1)**, **ICC(3,1)** – one-, two-way random- and mixed-effects ANOVA

- Whole-brain voxel-level ICC
  - ICC(2,1): 3dLME –ICC or 3dLME –ICCb
  - 3dICC: ICC(1,1), ICC(2,1) and ICC(3,1)

Chen et al. (2017), Human Brain Mapping 39(3) DOI:10.1002/hbm.23909

Naturalistic scanning

- Subjects view a natural scene during scanning
  - Visuoauditory movie clip (e.g., http://studyforrest.org/)
  - Music, speech, games, …
- Duration: a few minutes or more
- Close to naturalistic settings: minimally manipulated
- Effect of interest: intersubject correlation (ISC) – 3dTcorrelate
  - Extent of synchronization/entrainment
- Whole-brain voxel-wise analysis: 3dISC

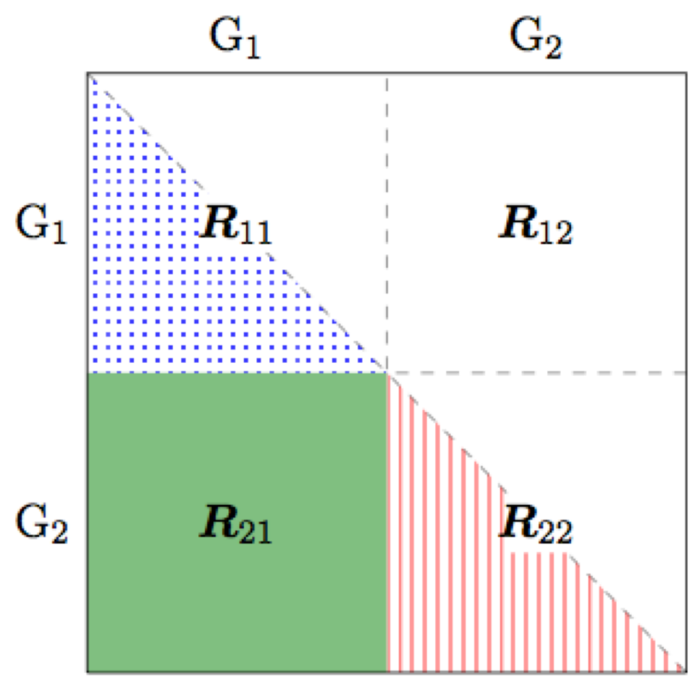
Hasson et al., 2004. Intersubject synchronization of cortical activity during natural vision. Science 303:1634-1640.

ISC group analysis

- Voxel-wise ISC matrix (usually Fisher-transformed)
    - One group

$$
\boldsymbol{R}^{(n)} = \begin{array}{c c}
 & \begin{array}{ccccc} S_1 & S_2 & S_3 & \cdots & S_n \end{array} \\
\begin{array}{c} S_1 \\ S_2 \\ S_3 \\ \vdots \\ S_n \end{array} &
\begin{pmatrix}
1 & r_{12} & r_{13} & \cdots & r_{1n} \\
r_{21} & 1 & r_{23} & \cdots & r_{2n} \\
r_{31} & r_{32} & 1 & \cdots & r_{3n} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
r_{n1} & r_{n2} & r_{n3} & \cdots & 1
\end{pmatrix}
\end{array}
\qquad
\boldsymbol{Z}^{(n)} = \begin{array}{c c}
 & \begin{array}{ccccc} S_1 & S_2 & S_3 & \cdots & S_n \end{array} \\
\begin{array}{c} S_1 \\ S_2 \\ S_3 \\ \vdots \\ S_n \end{array} &
\begin{pmatrix}
- & z_{12} & z_{13} & \cdots & z_{1n} \\
z_{21} & - & z_{23} & \cdots & z_{2n} \\
z_{31} & z_{32} & - & \dots & z_{3n} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
z_{n1} & z_{n2} & z_{n3} & \cdots & -
\end{pmatrix}
\end{array}
$$

    - Two groups
        - Within-group ISC: R11, R22
        - Inter-group ISC: R21
        - 3 group comparisons: R11 vs R22, R11 vs R21, R22 vs R21



Complexity of ISC analysis

- 2 ISC values associated with a common subject are correlated with each other: 5 subjects, 10 ISC values
- $\rho \neq 0$ characterizes non-independent relationship

$$
\begin{array}{c}
\\
Z_{21}\\ Z_{31}\\ Z_{41}\\ Z_{51}\\ Z_{32}\\ Z_{42}\\ Z_{52}\\ Z_{43}\\ Z_{53}\\ Z_{54}
\end{array}
\begin{array}{c}
\begin{matrix} Z_{21} & Z_{31} & Z_{41} & Z_{51} & Z_{32} & Z_{42} & Z_{52} & Z_{43} & Z_{53} & Z_{54} \end{matrix}\\
\begin{pmatrix}
1 & \rho & \rho & \rho & \rho & \rho & \rho & 0 & 0 & 0\\
\rho & 1 & \rho & \rho & \rho & 0 & 0 & \rho & \rho & 0\\
\rho & \rho & 1 & \boxed{\rho} & 0 & \rho & 0 & \rho & 0 & \rho\\
\rho & \rho & \rho & 1 & 0 & 0 & \rho & 0 & \rho & \rho\\
\rho & \rho & 0 & 0 & 1 & \rho & \rho & \rho & \rho & 0\\
\rho & 0 & \rho & 0 & \rho & 1 & \rho & \rho & 0 & \rho\\
\rho & 0 & 0 & \rho & \rho & \rho & 1 & 0 & \rho & \rho\\
0 & \rho & \boxed{\rho} & 0 & \rho & \rho & 0 & 1 & \rho & \rho\\
0 & \rho & 0 & \rho & \rho & 0 & \rho & \rho & 1 & \rho\\
0 & 0 & \rho & \rho & 0 & \rho & \rho & \rho & \rho & 1
\end{pmatrix}
\end{array}
$$

- **Challenge**: how to handle this irregular correlation matrix?

ISC: LME approach

- Modeling via effect partitioning: crossed random-effects LME

$$z_{ij} = b_0 + \theta_i + \theta_j + \epsilon_{ij}, \quad i \neq j$$

$$\theta_i, \theta_j \overset{iid}{\sim} G(0, \zeta^2) \text{ and } \epsilon_{ij} \overset{iid}{\sim} G(0, \eta^2)$$

cross-subject **within-subject**

- Charactering the relatedness among ISCs via LME

$$\rho = Corr(z_{ij}, z_{jl}) = \frac{Cov(z_{ij}, z_{jl})}{\sqrt{Var(z_{ij})Var(z_{jl})}} = \frac{\zeta^2}{2\zeta^2 + \eta^2}$$

$$0 \leq \rho = \frac{\zeta^2}{2\zeta^2 + \eta^2} = \frac{\zeta^2}{\sigma^2} \leq 0.5$$

Chen et al, 2016. Untangling the Relatedness among Correlations, Part II: Inter-Subject Correlation Group Analysis through Linear Mixed-Effects Modeling. Neuroimage. NeuroImage 147:825-840

Summary

- Concepts and terminology
- Group analysis approaches
  - GLM: 3dttest++, 3dMEMA
  - GLM, ANOVA, ANCOVA: 3dMVM
  - LME: 3dLME
  - Presumed vs. estimated HDR
- Miscellaneous
  - Issues with covariates
  - Intra-Class Correlation (ICC)
  - Inter-Subject Correlation (ISC)

Program List

- **3dttest++** (GLM: one-, two-sample, paired *t*, between-subjects variables)
- **3dMVM** (generic AN(C)OVA)
- **3dLME** (sophisticated cases: missing data, within-subject covariates)
- **3dMEMA** (similar to 3dttest++: measurement errors)
- **3dANOVA** (one-way between-subject)
- **3dANOVA2** (one-way within-subject, 2-way between-subjects)
- **3dANOVA3** (2-way within-subject and mixed, 3-way between-subjects)
- **3dttest** (obsolete: one-sample, two-sample and paired *t*)
- **3dRegAna** (obsolete: regression/correlation, covariates)
- **GroupAna** (obsolete: up to four-way ANOVA)
- **3dICC** (intraclass correlation): prototype only
- **3dISC** (intersubject correlation): prototype only