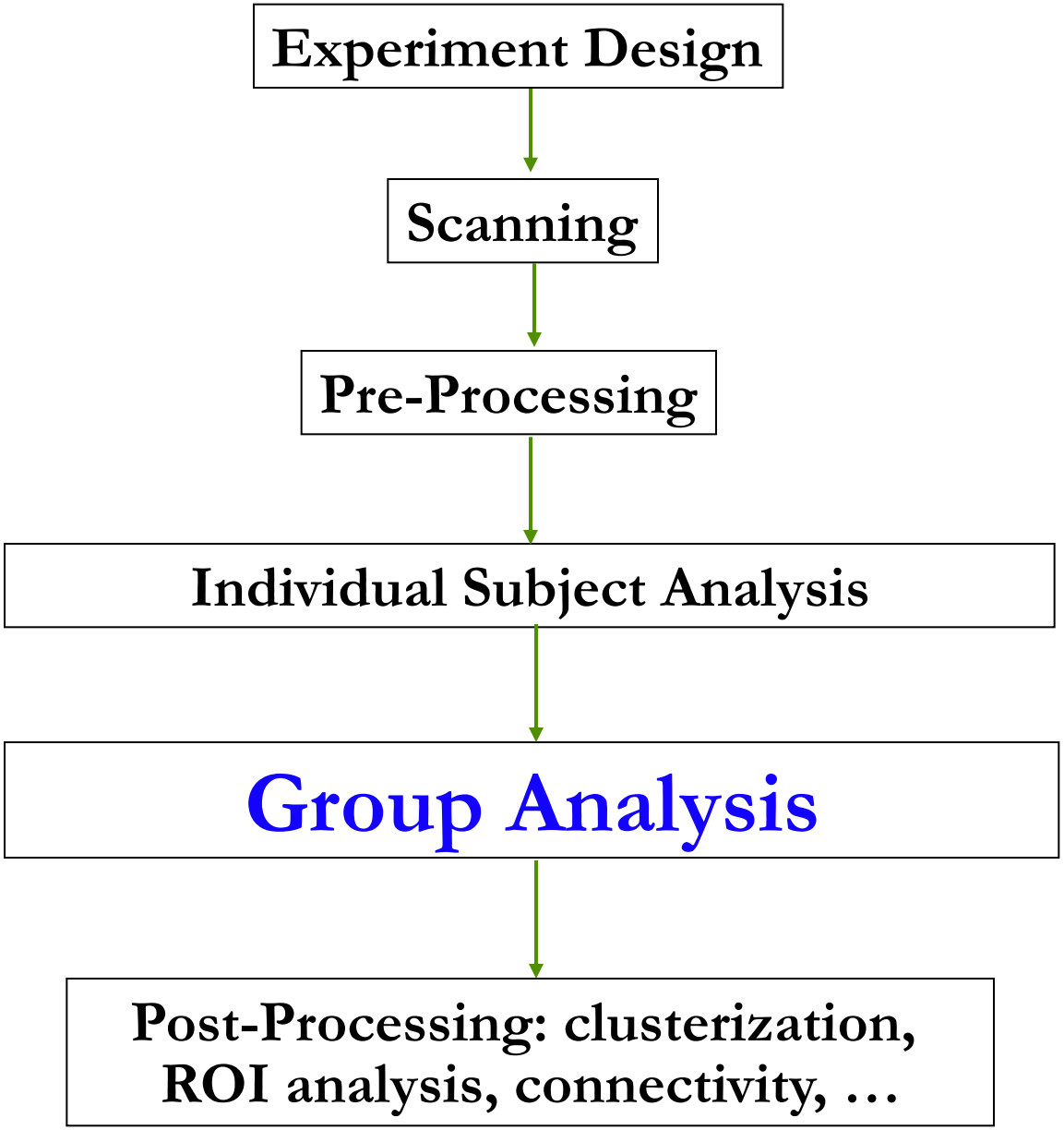
Group Analysis

File: **GroupAna.pdf**

Gang Chen

SSCC/NIMH/NIH/HHS

FMRI Study Pipeline



Preview

- Introduction: basic concepts
  - Why do we need to do group analysis?
  - Factor, quantitative covariates, main effect, interaction, …
- Group analysis approaches
  - *t*-test: 3dttest++ (3dttest), 3dMEMA
  - Regression: 3dttest++, 3dMEMA, 3RegAna
  - ANOVA: 3dANOVAx, 3dMVM, GroupAna
  - ANCOVA or GLM: 3dttest++, 3dMEMA, 3dMVM, 3dLME
  - Impact & consequence of SFM, SAM, and SEM
- Miscellaneous
  - Centering for covariates
  - Issues regarding result reporting
  - Intra-Class Correlation (ICC)
  - Nonparametric approach and fixed-effects analysis

Why Group Analysis?

- Evolution of FMRI studies
  - Early days: no need for group analysis
    - Seed-based correlation for one subject was revolutionary
  - Now: torture brain/data enough, and hope nature will confess!
    - Many ways to manipulate the brain (and data)
- Reproducibility and generalization
  - Science strives for generality: summarizing subject results
  - Typically 10 or more subjects per group
  - Exceptions: pre-surgical planning, lie detection, …
- Why not one analysis with a mega model for all subjects?
  - Computationally unmanageable
  - Heterogeneity in data or experiment design across subjects
  - Model quality check at individual subject level

Toy example of group analysis

- Responses from a group of subjects under one condition
  - What we have: $(\beta_1, \beta_2, \ldots, \beta_{10})=(1.13, 0.87, \ldots, 0.72)$
- Centroid: average $(\beta_1+\beta_2+\ldots+\beta_{10})/10 = 0.92$ is not enough
  - Variation/reliability measure: diversity, spread, deviation
- Model building
  - Subject *i*'s response = group average + deviation of subject *i*: simple model GLM (one-sample *t*-test)

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

  - If individual responses are consistent, ϵ_i should be small
  - How small (*p*-value)?
    - *t*-test: significance measure = $\frac{\hat{b}}{\hat{\sigma}/n}$
- 2 measures: *b* (dimensional) and *t* (dimensionless)

Group Analysis Caveats

- Conventional: voxel-wise (brain) or node-wise (surface)
  - Proper model to account for cross-and within-subject variability
- Results: two components (on afni: OLay + Thr)
  - Effect estimates: have unit and physical meaning
  - Their significance (response to house **significantly** > face)
    - Very unfortunately p-values solely focused in FMRI!
- Statistical significance (p-value) becomes obsession
  - Published papers: Big and tall parents (violent men, engineers) have more sons, beautiful parents (nurses) have more daughters
  - Statistical significance is not the same as practical importance
- Statistically insignificant but the effect magnitude is suggestive
  - Sample size
  - Alignment

Group Analysis Caveats

- Conventional: voxel-wise (brain) or node-wise (surface)
  - Prerequisite: reasonable alignment to some template
  - Limitations: alignment could be suboptimal or even poor
    - Different folding patterns across subjects: better alignment could help
    - Different cytoarchitectonic (or functional) locations across subjects: alignment won't help!
    - Impact on conjunction vs. selectivity
- Alternative (won't discuss): ROI-based approach
  - Half data for functional localizers, and half for ROI analysis
  - Easier: whole brain reduced to one or a few numbers per subject
  - Model building and tuning possible
  - Most AFNI 3d programs also handle ROI input (1D)

Terminology: Explanatory variables

- Response/Outcome variable (HDR): regression β coefficients
- Factor: categorical, qualitative, nominal or discrete variable
  - Categorization of conditions/tasks
    - Within-subject (repeated-measures) factor
  - Subject-grouping: Group of subjects (sex, normal/patients)
    - Between-subjects factor
    - Gender, patients/controls, genotypes, …
  - Subject: random factor measuring deviations
    - Of no interest, but served as random samples from a population
- Quantitative (numeric or continuous) covariate
  - Three usages of 'covariate'
    - Quantitative
    - Variable of no interest: qualitative (scanner, sex, handedness) or quantitative
    - Explanatory variable (regressor, independent variable, or predictor)
  - Examples: age, IQ, reaction time, *etc*.

Terminology: Fixed effects

- Fixed-effects factor: categorical (qualitative or discrete) variable
  - Treated as a fixed variable (constant to be estimated) in the model
    - Categorization of conditions/tasks (modality: visual/auditory)
      - Within-subject (repeated-measures) factor: 3 emotions
    - Subject-grouping: Group of subjects (gender, normal/patients)
      - Between-subject factor
  - All levels of a factor are of interest
    - main effect, contrasts among levels
  - Fixed in the sense of statistical inferences
    - Apply only to the specific levels of the factor
      - Categories: human, tool
    - Don't extend to other potential levels that might have been included
      - Inferences on human and tool categories can't be generated to animal
- Fixed-effects variable: quantitative covariate

Terminology: Random effects

- Random factor/effect
  - Random variable in the model: exclusively subject in FMRI
    - average + effects uniquely attributable to each subject: *e.g.* $N(\mu, \tau^2)$
    - Requires enough number of subjects
  - Each individual subject effect is of NO interest
    - Group response = 0.92%, subject 1 = 1.13%, random effect = 0.21%
  - Random in the sense
    - Subjects as random samples (representations) from a population
    - Inferences can be generalized to a hypothetical population
- A generic model: decomposing each subject's response $\boldsymbol{y}_i = X_i\boldsymbol{\beta} + Z_i\boldsymbol{b}_i + \boldsymbol{\epsilon}_i$
  - Fixed (population) effects: universal constants (immutable): $\boldsymbol{\beta}$
  - Random effects: individual subject's deviation from the population (personality: durable): $\boldsymbol{b}_i$
  - Residuals: noise (evanescent): $\boldsymbol{\epsilon}_i$

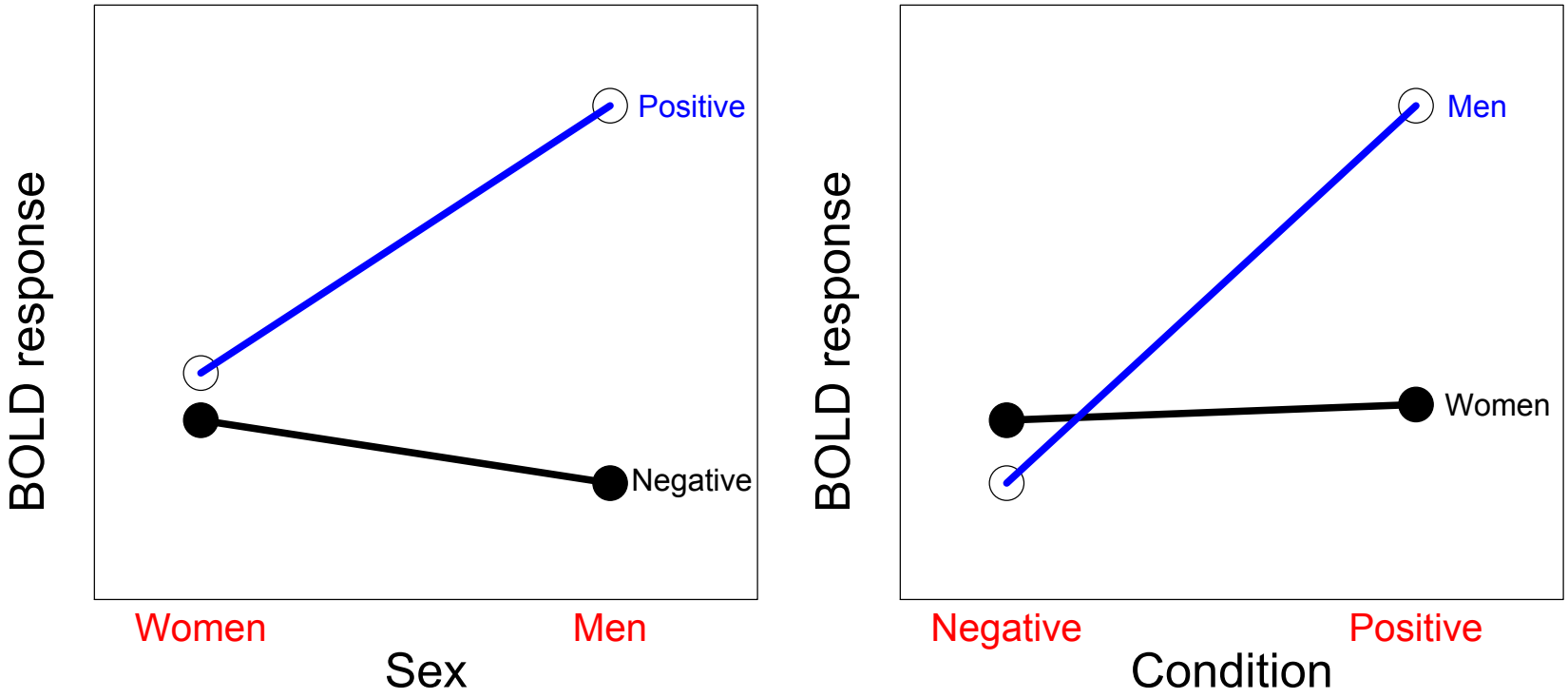
Terminology: Omnibus tests - main effect and interaction

- Main effect: any difference across levels of a factor?
- Interactions: with ≥ 2 factors, interaction may exist
  - 2 × 2 design: *F*-test for interaction between A and B = *t*-test of

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

    - *t* is better than *F*: a positive *t* shows

      A1B1 - A1B2 > A2B1 - A2B2 and A1B1 - A2B1 > A1B2 - A2B2

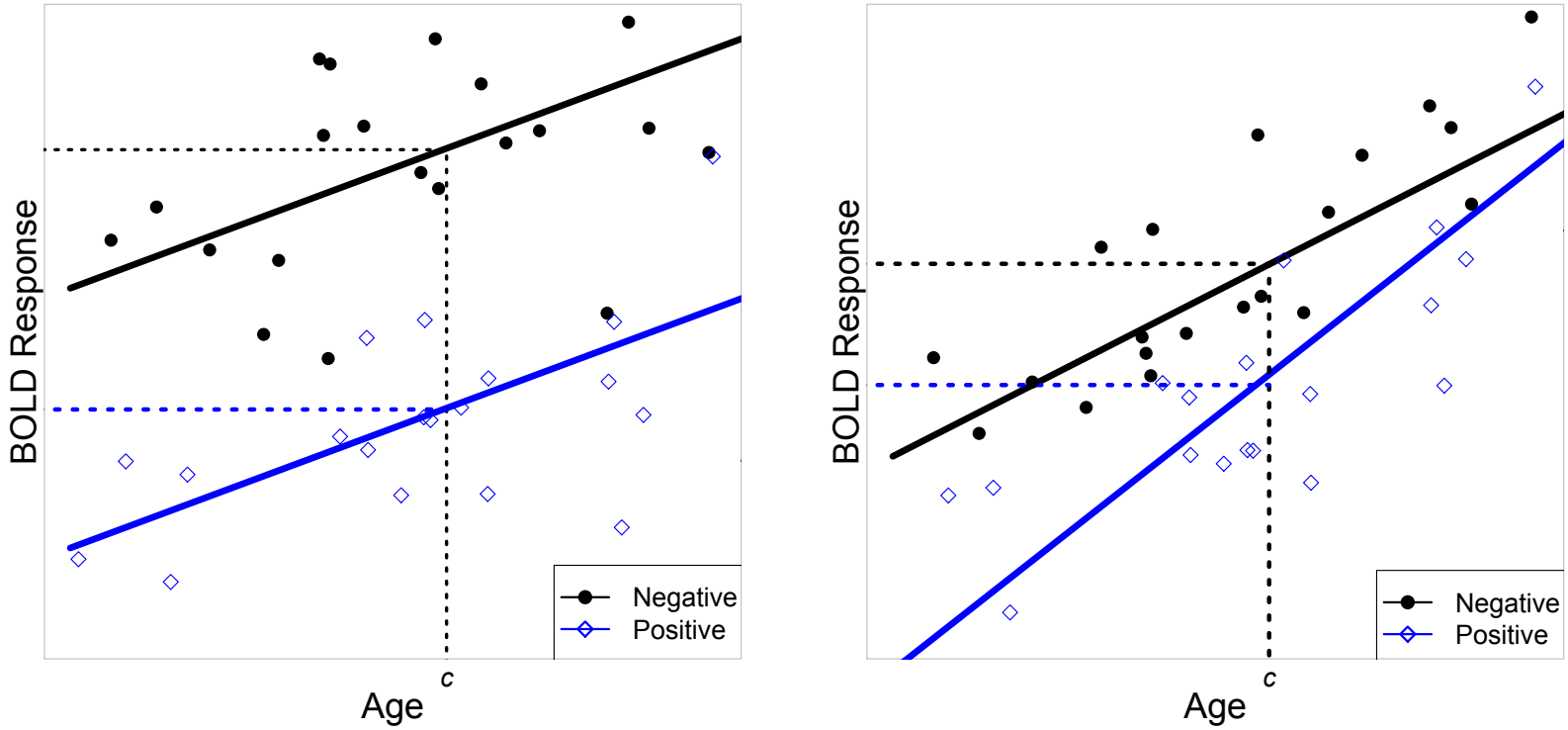


Terminology: Interaction

- Interactions: ≥ 2 factors
  - May become very difficult to sort out!
    - ≥ 3 levels in a factor
    - ≥ 3 factors
  - Solutions: reduction
    - Pairwise comparison
    - Plotting: ROI (Figures don't lie, but liars do figure. Mark Twain)
  - Requires sophisticated modeling
    - AN(C)OVA: 3dANOVAx, 3dMVM, 3dLME
- Interactions: quantitative covariates
  - In addition to linear effects, may have nonlinearity: $x_1 * x_2$, or x^2

Terminology: Interaction

- Interaction: between a factor and a quantitative covariate



  - Throw in an explanatory variable in a model as a nuisance regressor (additive effect) may not be enough
    - Model building/tuning: Potential interactions with other explanatory variables?
    - Of scientific interest (e.g., gender difference)

Models at Group Level

- Conventional approach: taking β (or linear combination of multiple β's) only for group analysis
  - Assumption: all subjects have same precision (reliability, standard error, confidence interval) about β
  - All subjects are treated equally
  - Student *t*-test: paired, one- and two-sample: not random-effects models in strict sense as usually claimed
  - AN(C)OVA, GLM, LME
- Alternative: taking both effect estimates and *t*-statistics
  - *t*-statistic contains precision information about effect estimates
  - Each subject is weighted based on precision of effect estimate
- All models are some sorts of linear model
  - *t*-test, AN(C)OVA, LME, MEMA
  - Partition each subject's effect into multiple components

Group Analysis in NeuroImaging: why big models?

- Various group analysis approaches
  - Student's ***t*-test**: one-, two-sample, and paired
  - **ANOVA**: one or more categorical explanatory variables (factors)
  - **GLM**: AN(C)OVA
  - **LME**: linear mixed-effects modeling
- *t*-tests not always practical or feasible
  - Tedious when layout is too complex
  - Main effects and interactions: desirable
  - When quantitative covariates are involved
- Advantages of big models: AN(C)OVA, GLM, LME
  - All tests in one analysis (vs. piecemeal *t*-tests)
  - Omnibus *F*-statistics
  - Power gain: combining subjects across groups

Piecemeal *t*-tests: 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
- Multiple *t*-tests
  - Group comparison + age effect
  - Pairwise comparisons among three conditions
    - Cannot control for age effect
  - Effects that cannot be analyzed
    - Main effect of Condition
    - Interaction between Group and Condition
    - Age effect across three conditions

Classical ANOVA: 2 × 3 Mixed ANCOVA

- 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

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

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

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

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

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

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

No correction for sphericity violation

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

b = X a + d

Univariate GLM: popular in neuroimaging

- Advantages: more *flexible* than the method of sums of squares
  - No limit on the the number of explanatory variables (in principle)
  - Easy to handle unbalanced designs
  - Covariates can be modeled when no within-subject factors present
- Disadvantages: costs paid for the flexibility
  - Intricate dummy coding
  - Tedious pairing for numerator and denominator of *F*-stat
    - Proper denominator SS
    - Can't generalize (in practice) to any number of explanatory variables
    - Susceptible to invalid formulations and problematic post hoc tests
  - **Cannot** handle covariates in the presence of within-subject factors
  - **No** direct approach to correcting for sphericity violation
    - **Unrealistic** assumption: **same** variance-covariance structure
- **Problematic**: When residual SS is adopted for all tests
  - *F*-stat: valid only for highest order interaction of within-subject factors
  - Most post hoc tests are inappropriate

Our 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

$$B_{n\times m} = X_{n\times q}\, A_{q\times m} + D_{n\times m}$$

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

Why taking β values for group analysis?

- Statistics (t, F)
  - Dimensionless
  - No physical meaning
  - Sensitive to sample size (#trials, #subjects) & signal-to-noise ratio
    - Are t-values of 4 and 100 (or p-values of 0.05 and 10^{-8}) really informative? The HDR of the latter is not 25 times larger than the former?
  - Distributional consideration
- β values
  - Physical meaning: measuring HDR magnitude: % signal change
- β values + their t-statistics
  - More accurate approach: 3dMEMA
  - Mostly about the same as the conventional approach
  - Not always practical

Road Map: Choosing a program?

- Starting with HDR estimated via shape-fixed method (SFM)
  - One β per condition per subject
  - It could be significantly underpowered (more later)
- Two perspectives
  - Data structure
  - Ultimate goal: list all the tests you want to perform
    - Possible to avoid a big model
    - Use a piecemeal approach with 3dttest++ or 3dMEMA
- Most analyses can be done with 3dMVM and 3dLME
  - Computationally inefficient
  - Last resort: not recommended if alternatives are available

Road Map: Student's *t*-tests

- 3dttest++ (3dttest) and 3dMEMA
- Not for *F*-tests except for ones with 1 DF for numerator
  - All factors are of two levels, e.g., 2 x 2, or 2 x 2 x 2
- Scenarios
  - One-, two-sample, paired
  - Multiple regression: one group + one or more quantitative variables
  - ANCOVA: two groups + one or more quantitative variables
  - ANOVA through dummy coding: all factors (between- or within-subject) are of two levels
  - AN(C)OVA: multiple between-subjects factors + one or more quantitative variables
  - One group against a whole brain constant: 3dttest -base1 C
  - One group against a voxel-wise constant: 3dttest -base1_dset

Road Map: Between-subjects ANOVA

- ✧ One-way between-subjects ANOVA
  - 3dANOVA
  - Two groups: 3dttest++, 3dMEMA (OK with > 2 groups too)
- ✧ Two-way between-subjects ANOVA
  - Equal #subjects across groups: 3dANOVA2 –type 1
  - Unequal #subjects across groups: 3dMVM
  - 2 x 2 design: 3dttest++, 3dMEMA (OK with > 2 groups too)
- ✧ Three-way between-subjects ANOVA
  - 3dANOVA3 –type 1
  - Unequal #subjects across groups: 3dMVM
  - 2 x 2 design: 3dttest++, 3dMEMA (OK with > 2 groups too)
- ✧ *N*-way between-subjects ANOVA
  - 3dMVM

Road Map: With-subject ANOVA

- ✧ One-way within-subject ANOVA
  - 3dANOVA2 –type 3
  - Two conditions: 3dttest++, 3dMEMA
- ✧ Two-way within-subject ANOVA
  - 3dANOVA3 –type 4
  - 2 x 2 design: 3dttest++, 3dMEMA
- ✧ *N*-way within-subject ANOVA
  - 3dMVM

Road Map: Mixed-type ANOVA and others

- One between- and one within-subject factor
  - Equal #subjects across groups: 3dANOVA3 -type 5
  - Unequal #subjects across groups: 3dMVM
  - 2 x 2 design: 3dttest++, 3dMEMA
- Other scenarios
  - Multi-way ANOVA: 3dMVM
  - Multi-way ANCOVA (between-subjects covariates only): 3dMVM
  - HDR estimated with multiple basis functions: 3dLME, 3dMVM
  - Missing data: 3dLME
  - Within-subject covariates: 3dLME
  - Subjects genetically related: 3dLME
  - Trend analysis: 3dLME

One-Sample Case

- One group of subjects ($n \geq 10$)
  - One condition (visual or auditory) effect
  - Linear combination of multiple effects (visual vs. auditory)
- Null hypothesis H_0: average effect = 0
  - Rejecting H_0 is of interest!
- Results
  - Average effect at group level (OLay)
  - Significance: *t*-statistic (Thr - Two-tailed by default)
- Approaches
  - uber_ttest.py (gen_group_command.py), 3dttest++ (3dttest), 3dMEMA
- Special cases
  - H_0: group effect = c (constant): 3dttest –base1 c…
  - H_0: group effect = c (voxelwise constant): 3dttest –base1_dset …

One-Sample Case: Example

- 3dttest++: taking β only for group analysis

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

- 3dMEMA: taking β and t-statistic for group analysis

```
3dMEMA -prefix VisGroupMEMA -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
```

Two-Sample Case

- Two groups of subjects ($n \geq 10$): males and females
  - One condition (visual or auditory) effect
  - Linear combination of multiple effects (visual vs. auditory)
  - Example: Gender difference in emotion effect?
- Null hypothesis H_0: Group1 = Group2
  - Results
    - Group difference in average effect
    - Significance: *t*-statistic - Two-tailed by default
- Approaches
  - uber_ttest.py, 3dttest++, 3dMEMA
  - One-way between-subjects ANOVA
    - 3dANOVA: can also obtain individual group test

Paired Case

- One groups of subjects ($n \geq 10$)
  - 2 conditions (visual or auditory): no missing data allowed (3dLME)
- Null hypothesis H_0: Condition1 = Condition2
  - Results
    - Average difference at group level
    - Significance: *t*-statistic (two-tailed by default)
- Approaches
  - uber_ttest.py, 3dttest++ (3dttest), 3dMEMA
  - One-way within-subject (repeated-measures) ANOVA
    - 3dANOVA2 –type 3: can also obtain individual condition test
  - Missing data (3dLME): only 10 among 20 subjects have both
- Essentially equivalent to one-sample case: use contrast as input

Paired Case: Example

- 3dttest++: comparing two conditions

```
3dttest++ -prefix Vis_Aud          \
  -mask mask+tlrc -paired          \
  -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 Case: Example

- 3dMEMA: comparing two conditions using subject-level response magnitudes and estimates of error levels
  - Contrast has to come from each subject

```
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]'
```

One-Way Between-Subjects ANOVA

- Two or more groups of subjects ($n \geq 10$)
  - One condition or linear combination of multiple conditions
  - Example: visual, auditory, or visual vs. auditory
- Null hypothesis H_0: Group1 = Group2
  - Results
    - Average group difference
    - Significance: *t*- and *F*-statistic (two-tailed by default)
- Approaches
  - 3dANOVA
  - > 2 groups: pair-group contrasts - 3dttest++ (3dttest), 3dMEMA
  - Dummy coding: 3dttest++, 3dMEMA
  - 3dMVM (not recommended)

Multiple-Way Between-Subjects ANOVA

- Two or more subject-grouping factors: factorial
  - One condition or linear combination of multiple conditions
  - Example: gender, control/patient, genotype, handedness, …
- Testing main effects, interactions, single group, group comparisons
  - Significance: *t*- (two-tailed by default) and *F*-statistic
- Approaches
  - Factorial design (imbalance not allowed): two-way (3dANOVA2 –type 1), three-way (3dANOVA3 –type 1)
  - 3dMVM: no limit on number of factors (imbalance allowed)
  - All factors have two levels: uber_ttest.py, 3dttest++, 3dMEMA
  - Using group coding with 3dttest++, 3dMEMA: imbalance allowed

One-Way Within-Subject ANOVA

- Also called one-way repeated-measures: one group of subject ($n \geq 10$)
  - Two or more conditions: extension to paired *t*-test
  - Example: happy, sad, neutral
- Main effect, simple effects, contrasts, general linear tests,
  - Significance: *t*- (two-tailed by default) and F-statistic
- Approaches
  - 3dANOVA2 -type 3 (two-way ANOVA with one random factor)
  - With two conditions, equivalent to paired case with 3dttest++ (3dttest), 3dMEMA
  - With more than two conditions, can break into pairwise comparisons with 3dttest++, 3dMEMA

One-Way Within-Subject ANOVA

- Example: visual vs. auditory condition

```
3dANOVA2 -type 3 -alevels 2 -blevels 10     \
 -prefix Vis_Aud -mask mask+tlrc            \
 -amean 1 Vis -amean 2 Aud -adiff 1 2 V-A   \
  -dset 1 1   'FP+tlrc[Vrel#0_Coef]'    \
  -dset 1 2   'FR+tlrc[Vrel#0_Coef]'    \
        ......
  -dset 1 10 'GM+tlrc[Vrel#0_Coef]'     \
  -dset 2 1   'FP+tlrc[Arel#0_Coef]'    \
  -dset 2 2   'FR+tlrc[Arel#0_Coef]'    \
        ......
  -dset 2 10 'GM+tlrc[Arel#0_Coef]'
```

Two-Way Within-Subject ANOVA

- Factorial design; also known as two-way repeated-measures
  - 2 within-subject factors
  - Example: emotion and category (visual/auditory)
- Testing main effects, interactions, simple effects, contrasts
  - Significance: *t*- (two-tailed by default) and F-statistic
- Approaches
  - 3dANOVA3 –type 4 (three-way ANOVA with one random factor)
  - All factors have 2 levels (2x2): uber_ttest.py, 3dttest++, 3dMEMA
  - Missing data?
    - Break into t-tests: uber_ttest.py, 3dttest++ (3dttest), 3dMEMA
    - 3dLME

Two-Way Mixed ANOVA

- Factorial design
  - One between-subjects and one within-subject factor
  - Example: gender (male and female) and emotion (happy, sad, neutral)
- Testing main effects, interactions, simple effects, contrasts
  - Significance: *t*- (two-tailed by default) and *F*-statistic
- Approaches
  - 3dANOVA3 –type 5 (three-way ANOVA with one random factor)
  - If all factors have 2 levels (2x2): 3dttest++, 3dMEMA
  - Missing data?
    - Unequal number of subjects across groups: 3dMVM, GroupAna
    - Break into *t*-tests: uber_ttest.py, 3dttest++ (3dttest), 3dMEMA
    - 3dLME

Univariate GLM: popular in neuroimaging

- Advantages: more *flexible* than the method of sums of squares
  - No limit on the the number of explanatory variables (in principle)
  - Easy to handle unbalanced designs
  - Covariates can be modeled when no within-subject factors present
- Disadvantages: costs paid for the flexibility
  - Intricate dummy coding
  - Tedious pairing for numerator and denominator of *F*-stat
    - Proper denominator SS
    - Can't generalize (in practice) to any number of explanatory variables
    - Susceptible to invalid formulations and problematic post hoc tests
  - **Cannot** handle covariates in the presence of within-subject factors
  - **No** direct approach to correcting for sphericity violation
    - **Unrealistic** assumption: **same** variance-covariance structure
- **Problematic**: When residual SS is adopted for all tests
  - *F*-stat: valid only for highest order interaction of within-subject factors
  - Most post hoc tests are inappropriate

MVM Implementation in **AFNI**

✧ Program 3dMVM

- No tedious and error-prone 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

Group analysis with multiple basis functions

- Shape-fixed method (SFM)
- Shape-estimated method (SEM) via basis functions: TENTzero, TENT, CSPLINzero, CSPLIN
  - Area under the curve (AUC) approach
    - Ignore subtle shape difference
    - Focus on the response magnitude measured by AUC
    - Potential issues: Shape information lost; Undershoot may cause trouble
  - Better approach: maintaining shape information
    - Take individual β values to group analysis
- Shape-adjusted method (SAM) via SPMG2/3
  - Only take the major component to group level
  - Reconstruct HDR, and take the effect estimates

Group analysis with multiple basis functions

- Analysis with effect estimates at consecutive time grids
  - Used to be considered very difficult
  - Extra variable, Time = t_0, t_1, …, t_k
  - One group of subjects under one condition
    - Accurate hypothesis H_0: $\beta_1=0$, $\beta_2=0$, …, $\beta_k=0$ (NOT $\beta_1=\beta_2=\ldots=\beta_k$)
      - Testing the centroid (multivariate testing)
      - 3dLME
    - Approximate hypothesis H_0: $\beta_1=\beta_2=\ldots=\beta_k$ (main effect)
      - 3dMVM
    - Result: F-statistic for H_0 and t-statistic for each time grid

Group analysis with multiple basis functions

- Multiple groups (or conditions) under one condition (or group)
  - Accurate hypothesis: $\beta_1^{(1)} - \beta_1^{(2)} = 0, \beta_2^{(1)} - \beta_2^{(2)} = 0, ..., \beta_k^{(1)} - \beta_k^{(2)} = 0$
    - 2 conditions: 3dLME
  - Approximate hypothesis: $\beta_1^{(1)} = \beta_1^{(2)}, \beta_2^{(1)} = \beta_2^{(2)}, ..., \beta_k^{(1)} = \beta_k^{(2)}$
    - Interaction
    - Multiple groups: 3dANOVA3 –type 5 (two-way mixed ANOVA: equal #subjects), or 3dMVM
    - Multiple conditions: 3dANOVA3 –type 4
  - Focus: do these groups/conditions have different response shape?
    - *F*-statistic for the interaction between Time and Group/Condition
    - *F*-statistic for main effect of Group: group/condition difference of AUC
    - *F*-statistic for main effect of Time: HDR effect across groups/conditions
- Other scenarios: factor, quantitative variables
  - 3dMVM

Correlation analysis

- Correlation between brain response and behavioral measures

$$\hat{\beta}_i = \alpha_0 + \alpha_1 * x_i + \epsilon_i$$

  - Difference between correlation and regression?
    - Essentially the same
    - When explanatory and response variable are standardized, the regression coefficient = correlation coefficient
  - Two approaches
    - Standardization
    - Convert *t*-statistic to *r* (or determination coefficient)

$$R^2 = t^2/(t^2 + DF)$$

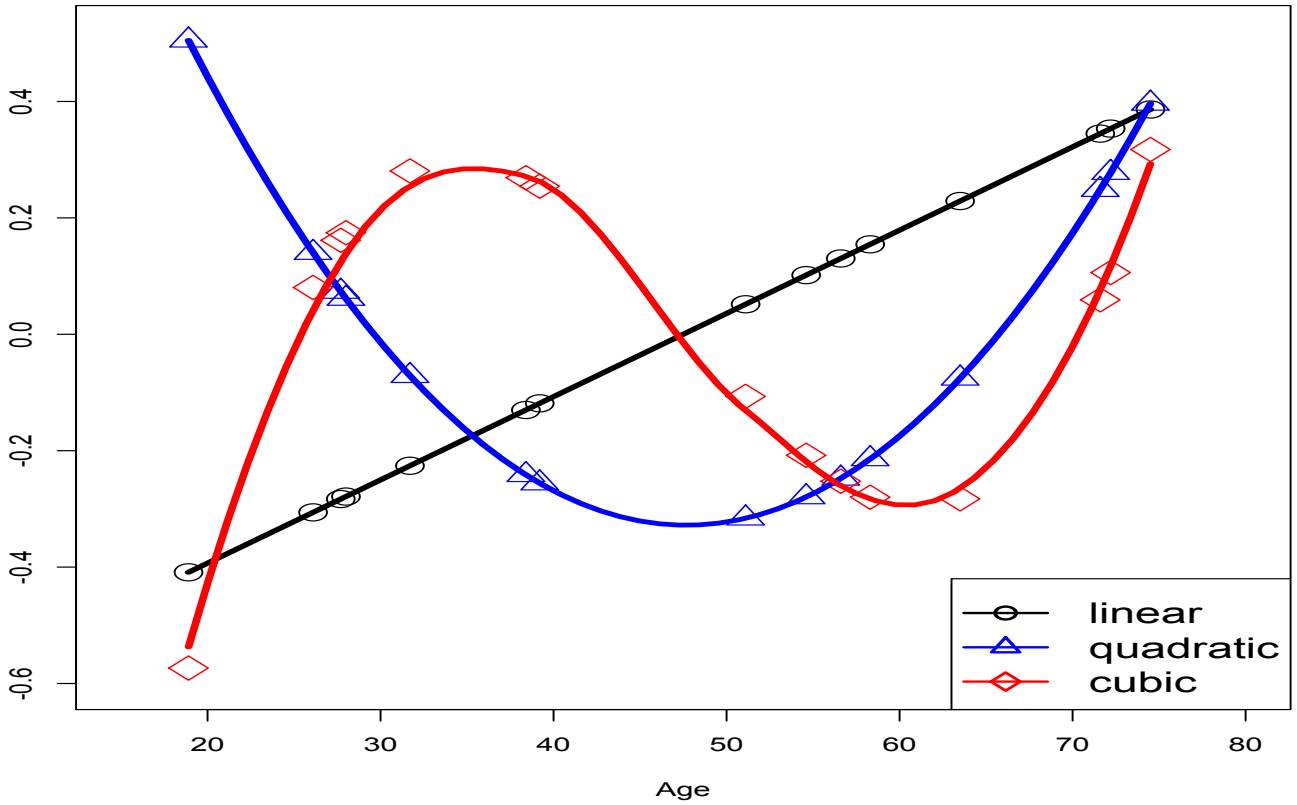
    - Programs: 3dttest++, 3dMEMA, 3dMVM, 3dRegAna

Trend analysis

- Correlation between brain response and some gradation
  - Linear, quadratic, or higher-order effects
    - Habituation or attenuation effect across time (trials)
    - Between-subjects: Age, IQ
      - Fixed effect
    - Within-subject measures: morphed images
      - Random effects involved: 3dLME
  - Modeling: weights based on gradation
    - Equally-spaced: coefficients from orthogonal polynomials
    - With 6 equally-spaced levels, *e.g.*, 0, 20, 40, 60, 80, 100%,
      - Linear: -5 -3 -1 1 3 5
      - Quadratic: 5 -1 -4 -4 -1 5
      - Cubic: -5 7 4 -4 -7 5

Trend analysis

- Correlation between brain response and some gradation
  - Modeling: weights based on gradation
    - Not equally-spaced: constructed from, *e.g.*, poly() in R
    - Ages of 15 subjects: 31.7 38.4 51.1 72.2 27.7 71.6 74.5 56.6 54.6 18.9 28.0 26.1 58.3 39.2 63.5



Trend analysis: summary

- Cross-trials trend: AM2 with weights
- Modeling with within-subject trend
  - Run GLT with appropriate weights at individual levels
- Modeling with within-subject trend: 3 approaches
  - Set up GLT weights among factor levels at group level 3dANOVA2/3, 3dMVM, 3dLME: best with equally-spaced with even number of levels
  - Set up the weights as the values of a variable
    - Needs to account for deviation of each subject
    - 3dLME
  - Run trend analysis at individual level (*i.e.*, -gltsym), and then take the trend effect estimates to group level
    - Simpler than the other two approaches

Group analysis with quantitative variables

- Covariate: 3 usages
  - Quantitative (vs. categorical) variable
    - Age, IQ, behavioral measures, …
  - Of no interest to the investigator
    - Age, IQ, sex, handedness, scanner,…
  - Any explanatory variables in a model
- Variable selection
  - Infinite candidates: relying on prior information
  - Typical choices: age, IQ, RT, …
  - RT: individual vs. group level
    - Amplitude modulation: cross-trial variability at individual level
    - Group level: variability across subjects

Group analysis with quantitative variables

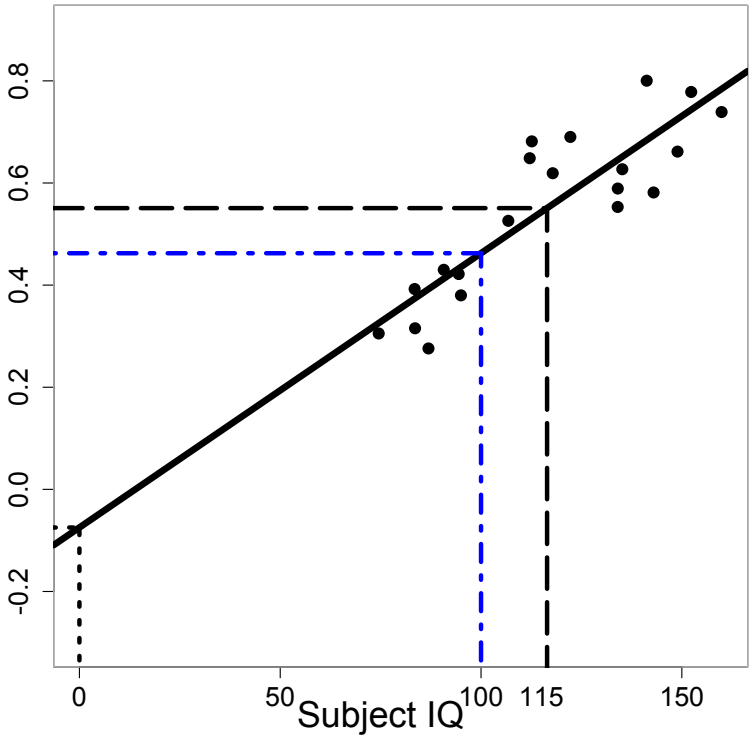
- Conventional framework
  - ANCOVA: one between-subjects factor (e.g., sex) + one quantitative variable (e.g., age)
    - Extension to ANOVA: GLM
    - Homogeneity of slopes
- Broader framework
  - Any modeling approaches involving quantitative variables
    - Regression, GLM, MVM, LME
    - Trend analysis, correlation analysis

Quantitative variables: subtleties

- Regression: one group of subjects + quantitative variables

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

  - Interpretation of effects
    - α_1 - slope (change rate, marginal effect): effect per unit of x
    - α_0 – intercept: group effect while x=0
      - Not necessarily meaningful
      - Linearity may not hold
      - Solution: centering - crucial
      - for interpretability
      - Mean centering?

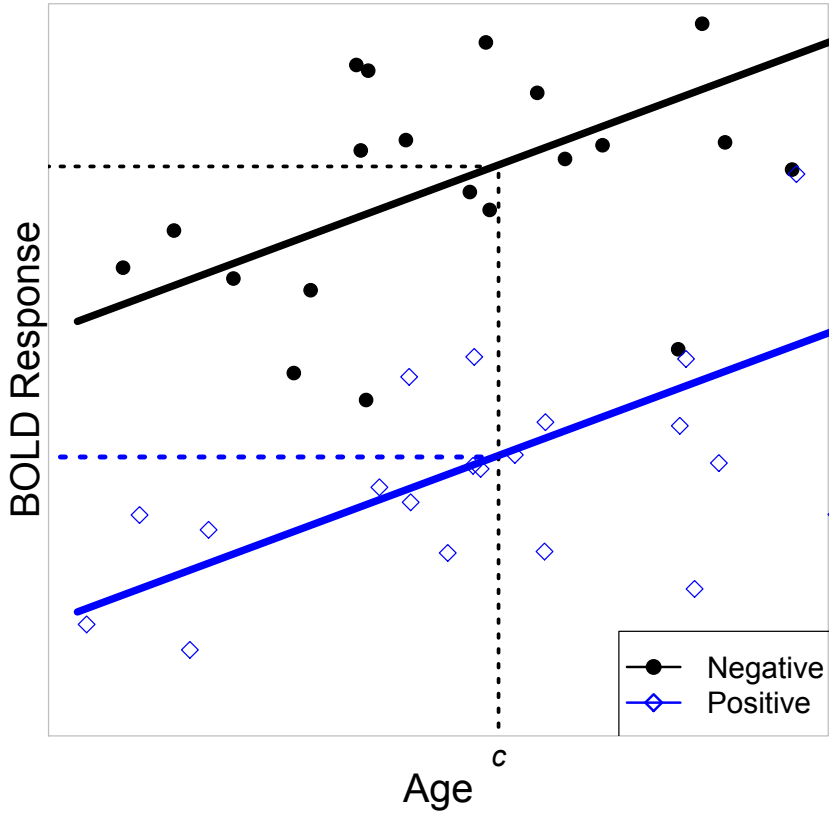


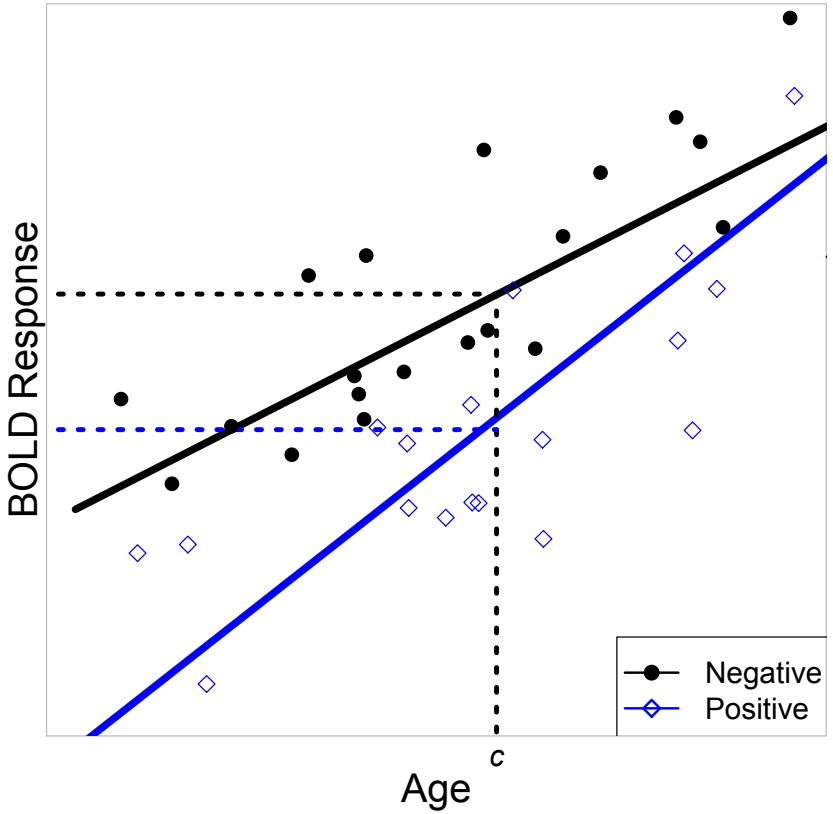
Quantitative variables: subtleties

- Trickier scenarios with two or more groups

$$\hat{\beta}_i = \alpha_0 + \alpha_1 * x_{1i} + \alpha_2 * x_{2i} + \alpha_3 * x_{3i} + \epsilon_{ij}$$

  - Interpretation of effects
    - Slope: Interaction! Same or different slope?
    - α_0 – same or different center?

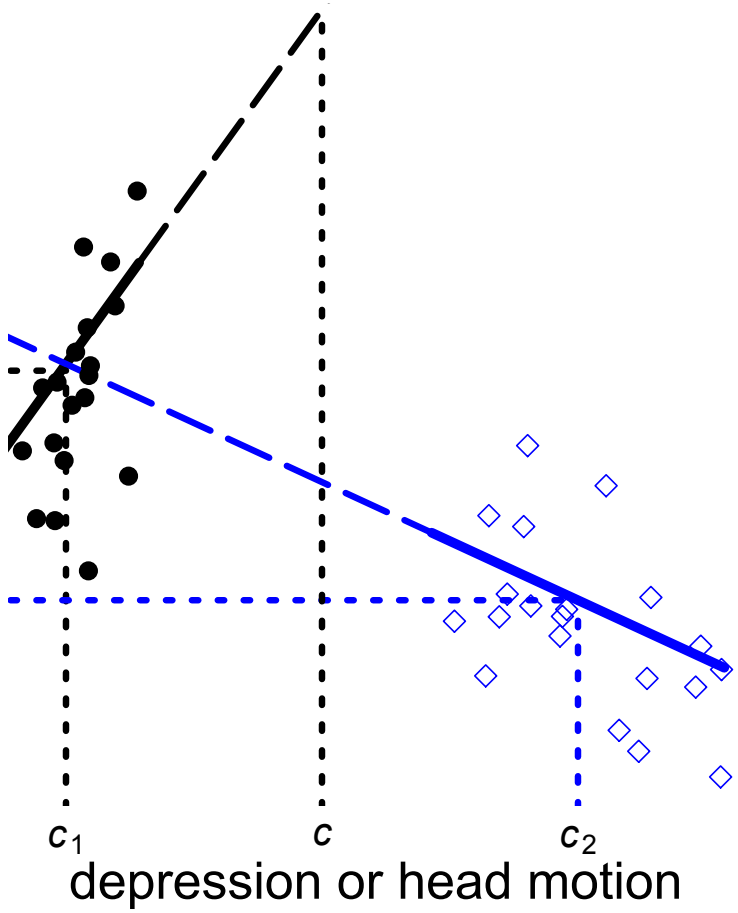




Quantitative variables: subtleties

- Trickiest scenario with two or more groups

$$\hat{\beta}_i = \alpha_0 + \alpha_1 * x_{1i} + \alpha_2 * x_{2i} + \alpha_3 * x_{3i} + \epsilon_{ij}$$

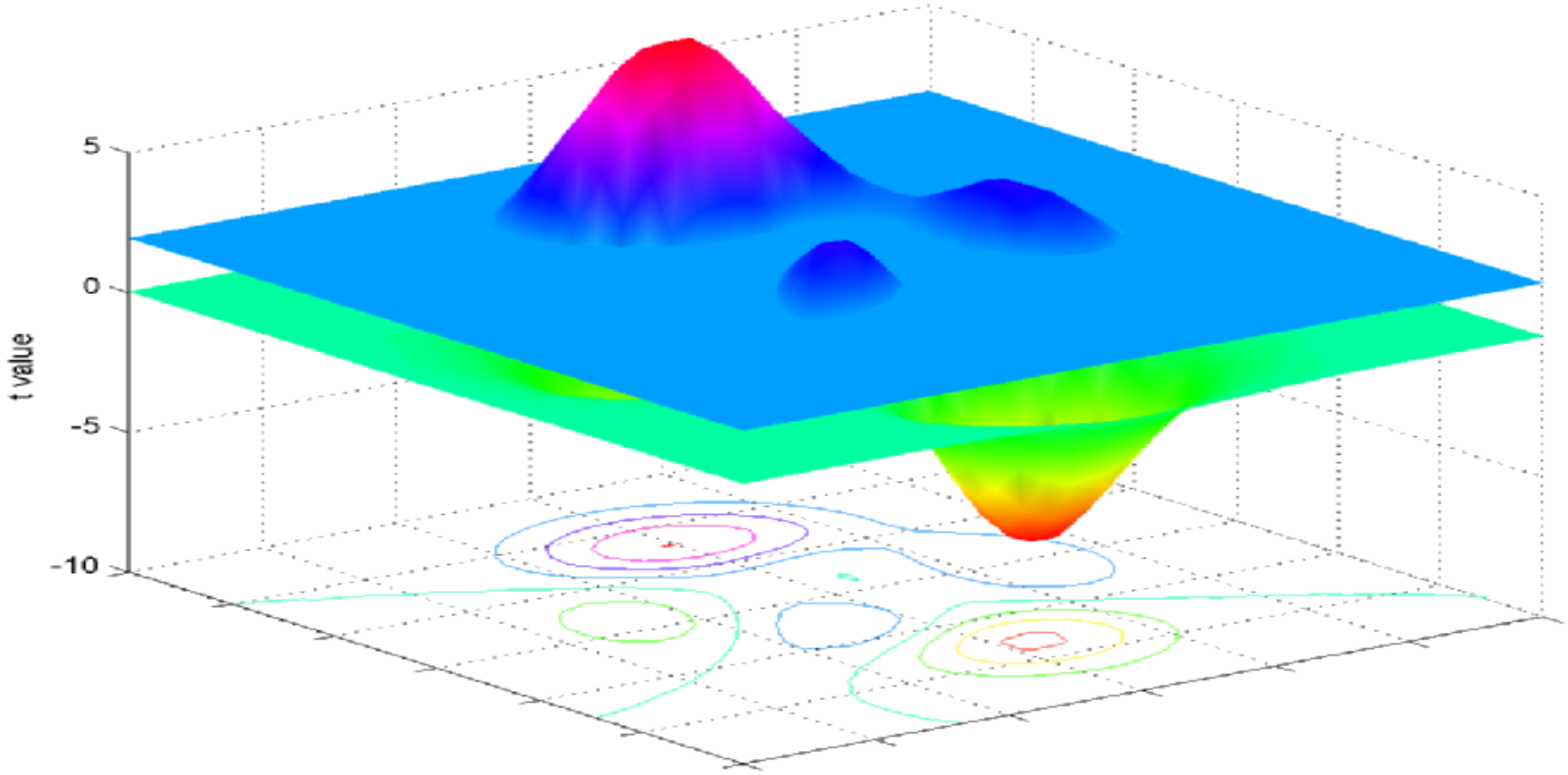


- More details: **http://afni.nimh.nih.gov/sscc/gangc/centering.html**

Why should we report response magnitudes?

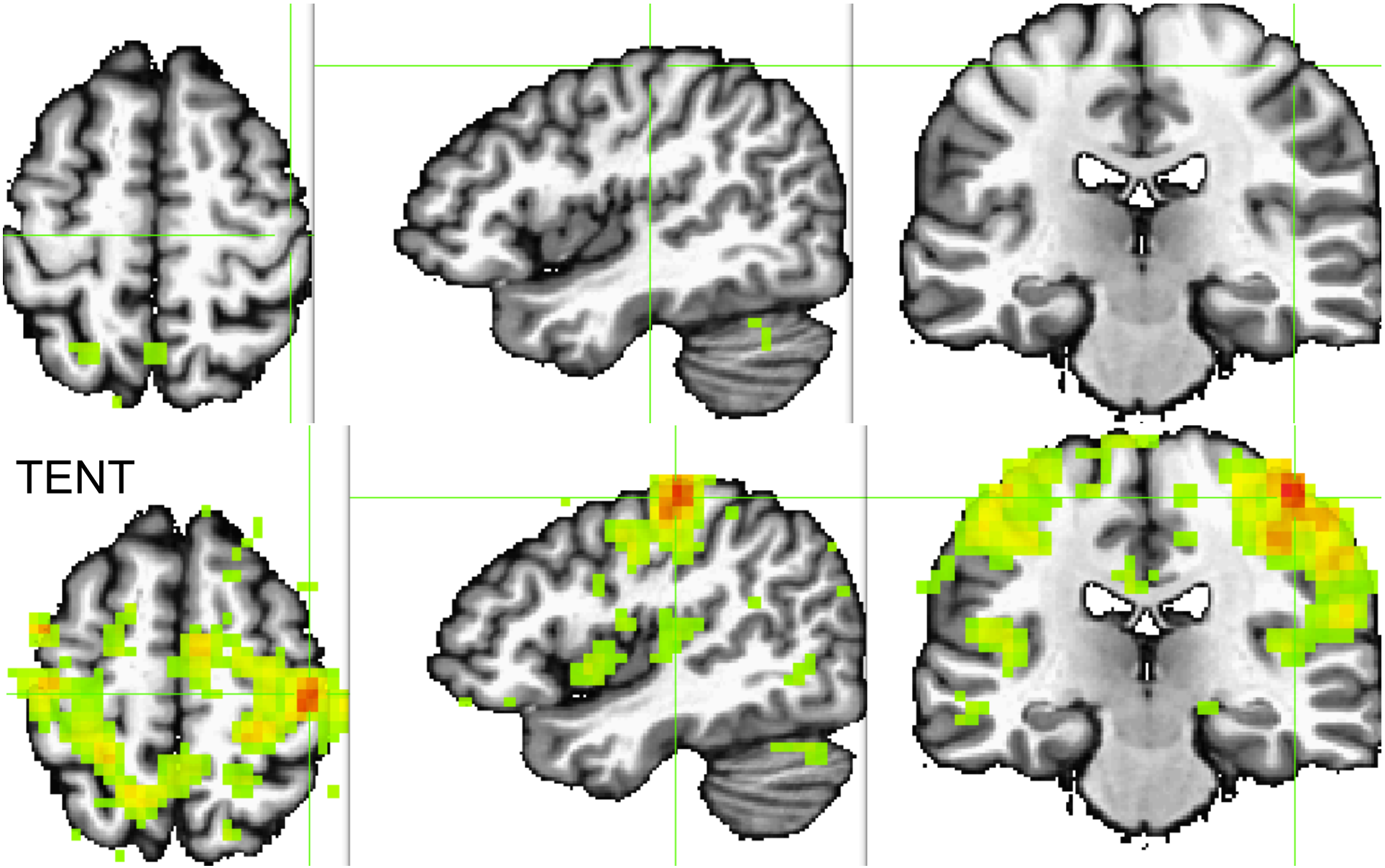
- **Unacceptable** in some fields if only significance is reported
  - Neuroimaging: an exception currently!
- Obsession in FMRI about *p*-value!
  - Colored blobs of *t*-values
  - Peak voxel selected based on peak *t*-value
- Science is about reproducibility
  - Response amplitude should be of primacy focus
  - Statistics are only for thresholding
    - No physical dimension
    - Once surviving threshold, specific values are not informative

- **Basics: Null hypothesis significance testing (NHST)**
  - Should science be based on a dichotomous or binary inference?
    - If a cluster fails to survive for thresholding, there is no value?
    - SVC: Band-Aid solution

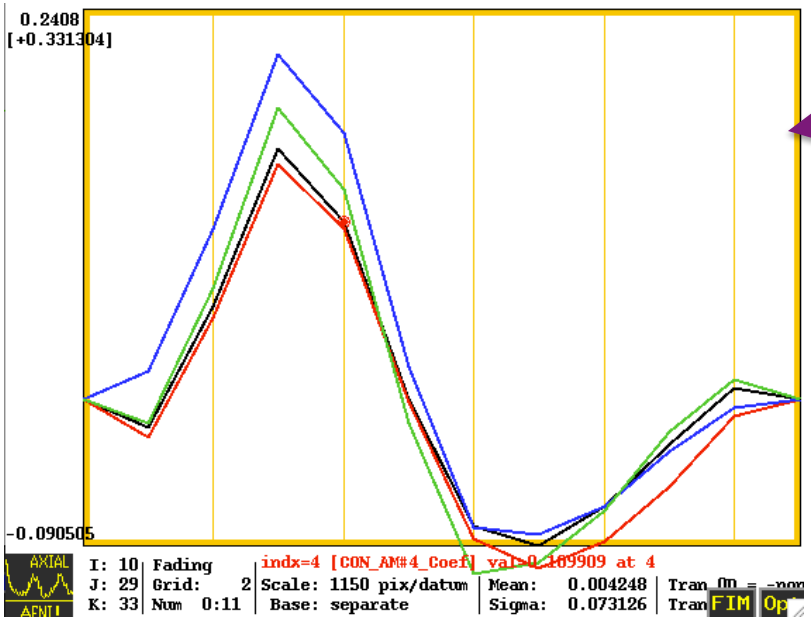


Modeling strategy & results: an example

SPMG3: 1st β (canonical HDR) [voxel-wise p=0.01]

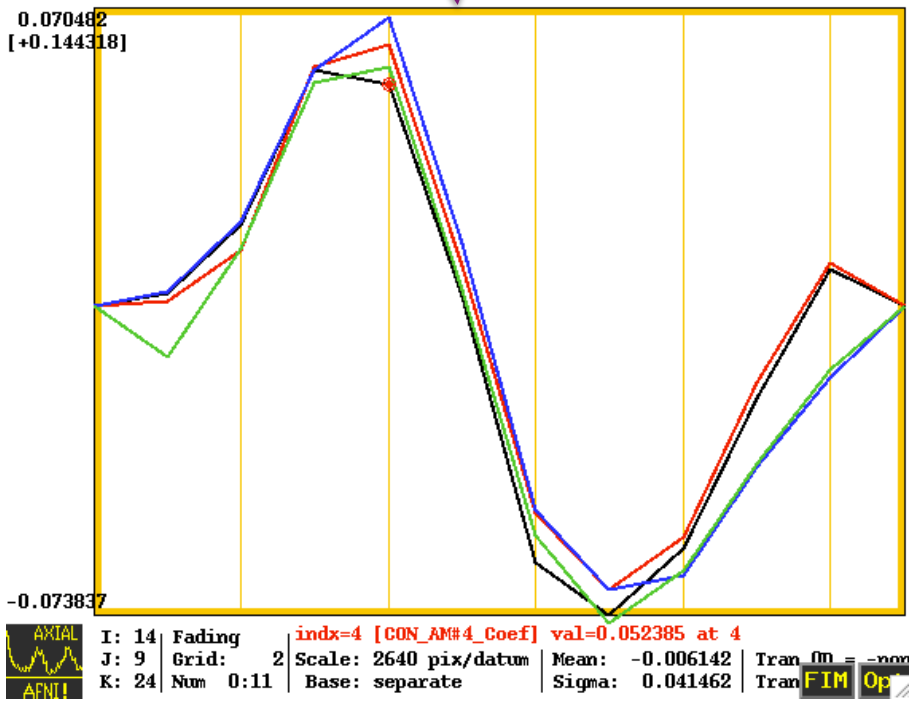


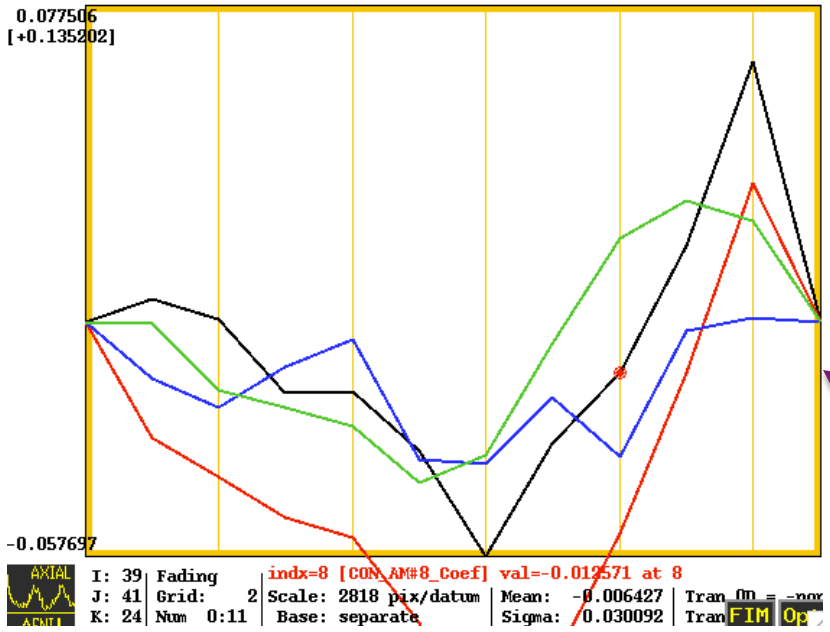
Is *p*-value everything? An example



$p=10^{-7}$

$p=0.5$





$p=10^{-4}$

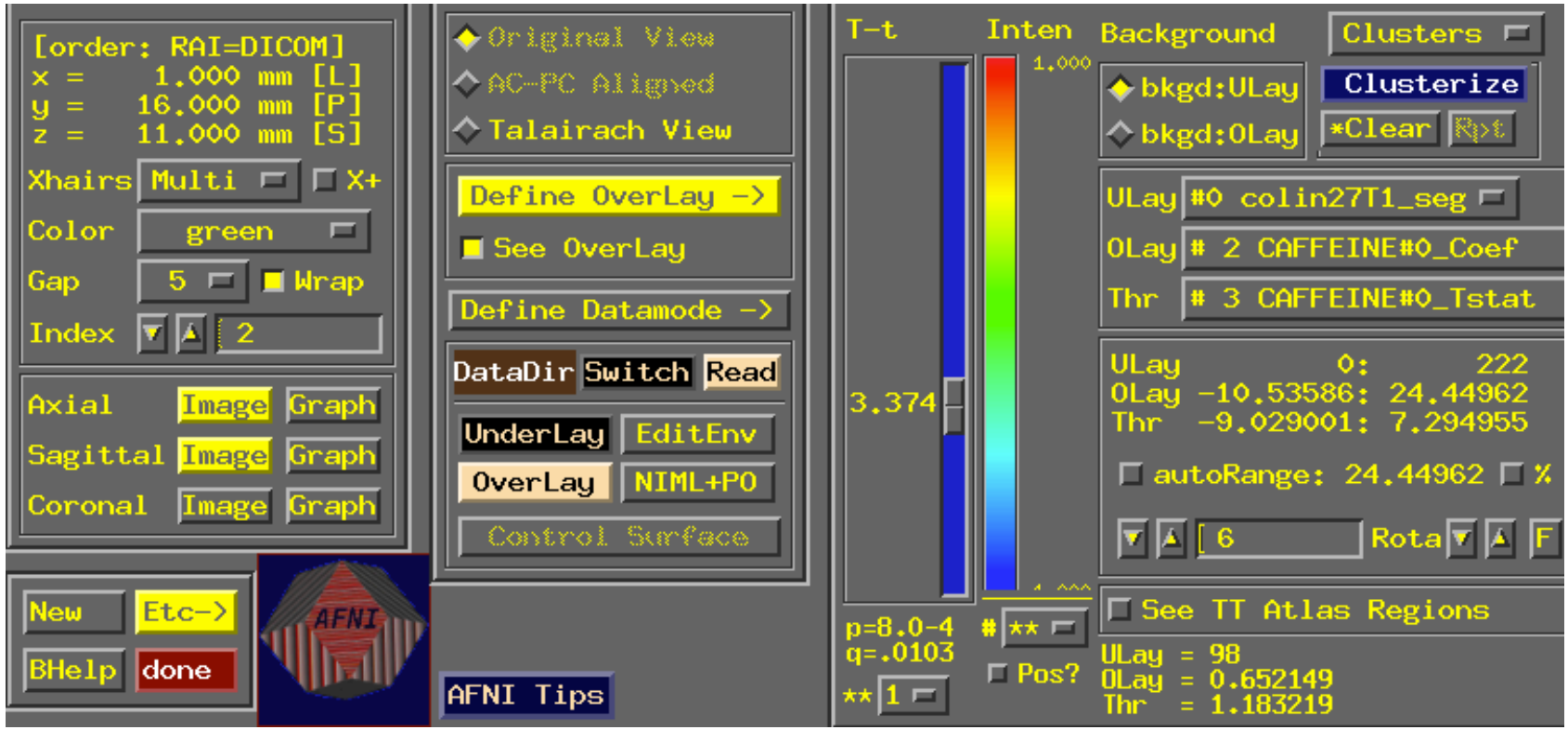
Advantages of SEM

- Multiple basis functions
  - TENTzero, TENT, CSPLINzero, CSPLIN
  - Similar to FIR in SPM, but FIR does not allow non-TR-synchronized modeling
- Higher statistical power than SFM and SAM
  - More likely identifying activations
- Extra support for true positives (TP) with HDR signature shape
  - Unavailable from SFM and SAM
- Crucial evidence if significance is marginal: false negatives (FP)
- Avoiding false positives (FP)
- Works best for event-related experiments
  - Useful for block designs: habituation, attenuation,…

How rigorous about corrections?

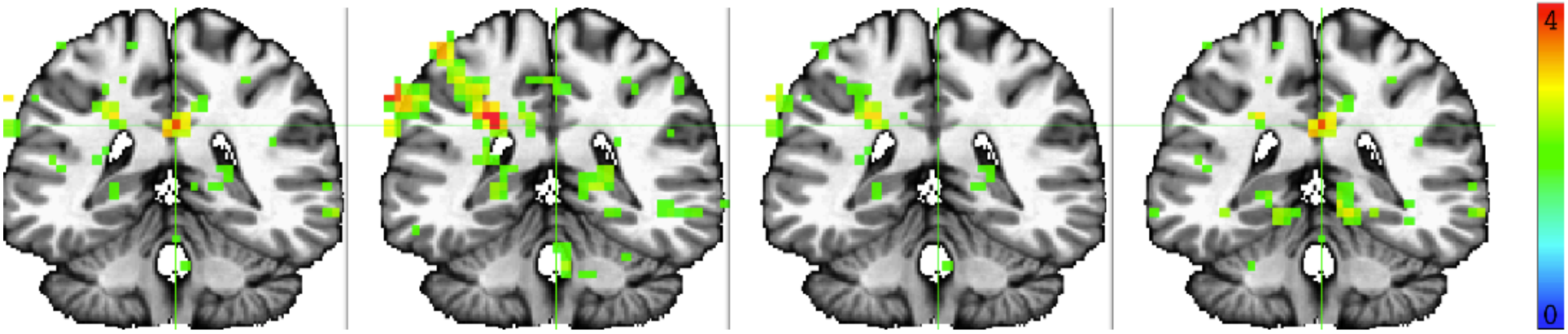
- Two types of correction
  - Multiple testing correction n(MTC): same test across brain
    - FWE, FDR, SVC(?)
    - People (esp. reviewers) worship this!
  - Multiple comparisons correction (MCC): different tests
    - Happy vs. Sad, Happy vs. Neutral, Sad vs. Neutral
    - Two one-sided *t*-tests: p-value is ½ of two-sided test!
    - How far do you want to go?
      - Tests in one study
      - Tests in all FMRI or all scientific studies?
    - Nobody cares the issue in FMRI
- Many reasons for correction failure
  - Region size, number of subjects, alignment quality, substantial cross-subject variability (anxiety disorder, depression, …)

Presenting response magnitudes



Presenting response magnitudes

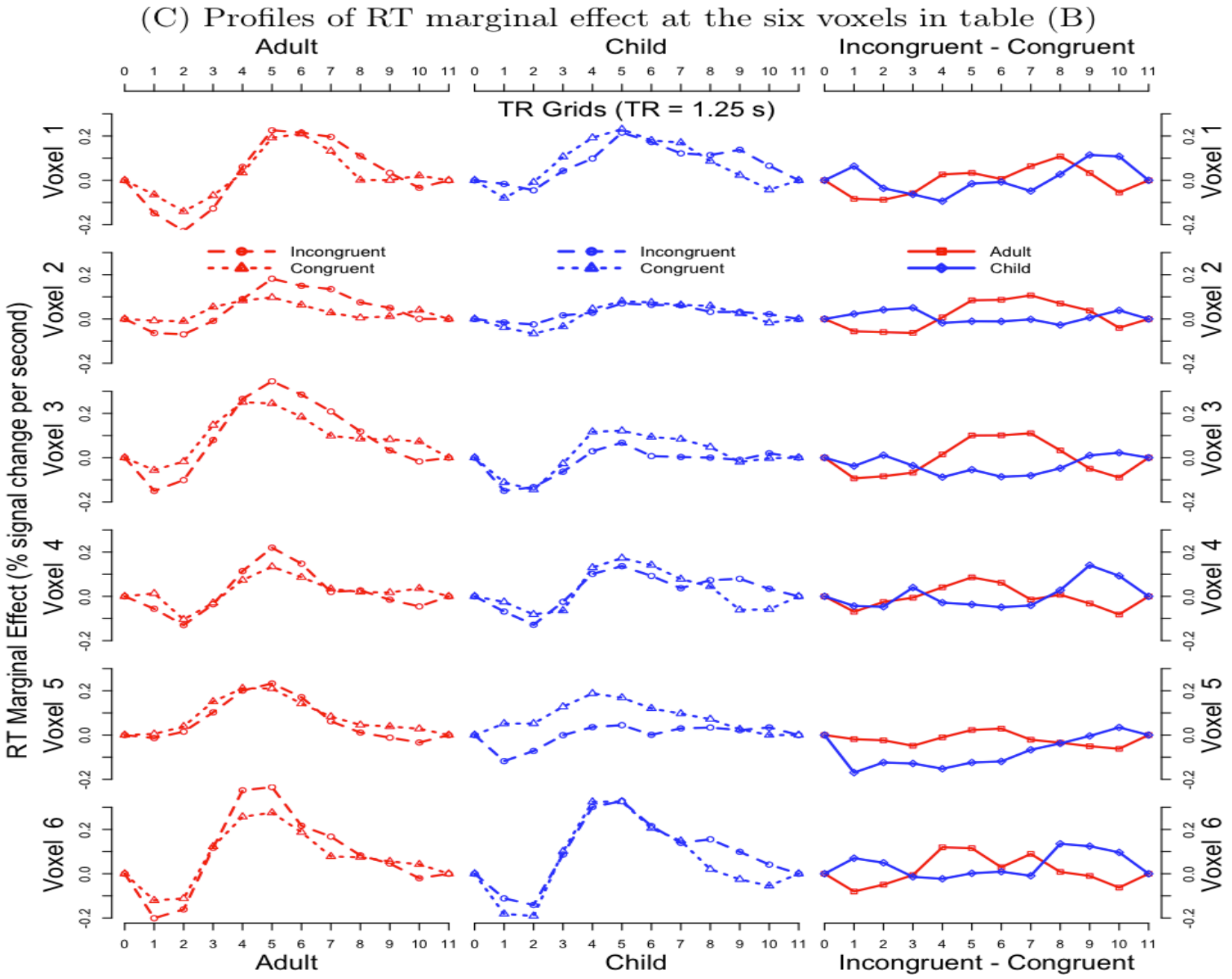
(A) Coronal view of interaction effect of Group:Condition:Time



(B) Sphericity scenarios at six representative voxels

Voxel		Sphericity			**UVT-UC**	**UVT-SC**	**MVT-WS**	**HT**
No.	coordinates	Mauchly p-value	ϵ_{GG}	ϵ_{HF}	p-value	p-value	p-value	taking
1	-2 36 27	0	0.32	0.35	0.28	0.31	0.00021	MVT-WS
2	-33 -5 42	0	0.42	0.46	3.8×10^{-6}	8.4×10^{-4}	1.6×10^{-4}	MVT-WS
3	-50 -16 24	0	0.45	0.50	1.6×10^{-4}	0.0041	0.14	MVT-WS
4	-5 -20 23	8.7×10^{-6}	0.68	0.79	1.8×10^{-5}	0.0001	0.008	UVT-SC
5	37 68 20	0	0.30	0.32	0.012	0.074	0.15	MVT-WS
6	-36 -16 7	0	0.53	0.60	1.8×10^{-5}	5.3×10^{-4}	0.0019	UVT-SC

Presenting response magnitudes



IntraClass Correlation (ICC)

- Reliability (consistency, reproducibility) of signal: extent to which the levels of a factor are related to each other
  - Example – 3 sources of variability: conditions, sites, subjects
  - Traditional approach: random-effects ANOVAs
  - LME approach

$$\hat{\beta}_{ijk} = \alpha_0 + \alpha_1 * x_k + b_i + c_j + d_k + \epsilon_{ijk},$$
$$b_i \sim N(0, \tau_1^2), c_j \sim N(0, \tau_2^2), d_k \sim N(0, \tau_3^2), \epsilon_{ijk} \sim N(0, \sigma^2)$$

$$ICC_l = \frac{\tau_l^2}{\tau_l^2 + \tau_2^2 + \tau_3^2 + \sigma^2}, l = 1, 2, 3$$

  - 3dICC_REML, 3dLME

Group Analysis: Non-Parametric Approach

- Parametric approach
  - Enough number of subjects $n > 10$
  - Random effects of subjects: usually Gaussian distribution
  - Individual and group analyses: separate

- Non-parametric approach
  - Moderate number of subjects: $4 < n < 10$
  - No assumption of data distribution (e.g., normality)
  - Statistics based on ranking or permutation
  - Individual and group analyses: separate

Group Analysis: Fixed-Effects Analysis

- When to consider?
  - LME approach
  - Group level: a few subjects: $n < 6$
  - Individual level: combining multiple runs/sessions
- Case study: difficult to generalize to whole population
- Model $\beta_i = b + \varepsilon_i$, $\varepsilon_i \sim N(0, \sigma_i^2)$, σ_i^2: within-subject variability
  - Fixed in the sense that cross-subject variability is not considered
- Direct fixed-effects analysis (**`3dDeconvolve/3dREMLfit`**)
  - Combine data from all subjects and then run regression
- Fixed-effects meta-analysis (**`3dcalc`**) : weighted least squares
  - $\beta = \sum w_i \beta_i / \sum w_i$, $w_i = t_i / \beta_i$ = weight for ith subject
  - $t = \beta \sqrt{\sum w_i}$

Non-Parametric Analysis

- Ranking-based: roughly equivalent to permutation tests
  - **3dWilcoxon** (~ paired *t*-test)
  - **3dFriedman** (~ one-way within-subject with **3dANOVA2**)
  - **3dMannWhitney** (~ two-sample *t*-test)
  - **3dKruskalWallis** (~ between-subjects with **3dANOVA**)
- Pros: Less sensitive to outliers (more robust)
- Cons
  - Multiple testing correction **limited** to FDR (**3dFDR**)
  - Less flexible than parametric tests
    - Can't handle complicated designs with more > 1 fixed-effects factor
    - Can't handle covariates
- Permutation approach?

Group Analysis Program List

- **3dttest++** (one-sample, two-sample and paired *t*) + covariates (voxel-wise)
- **3dMEMA** (R package for mixed-effects analysis, t-tests plus covariates)
- **3ddot** (correlation between two sets)
- **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)
- **3dMVM** (AN(C)OVA, and within-subject MAN(C)OVA)
- **3dLME** (R package for sophisticated cases)
- **3dttest** (mostly obsolete: one-sample, two-sample and paired *t*)
- **3dRegAna** (obsolete: regression/correlation, covariates)
- **GroupAna** (mostly obsolete: Matlab package for up to four-way ANOVA)

FMRI Group Analysis Comparison

		AFNI	SPM	FSL
t-test (one-, two-sample, paired)		3dttest++, 3dMEMA	Yes	FLAME1, FLAME1+2
One categorical variable: one-way ANOVA		3dANOVA/2/3, GroupAna	Only **one** WS factor: full and flexible factorial design	Only **one** within-subject factor: GLM in FEAT
Multi-way AN(C)OVA		3dANOVA2/3, GroupAna, 3dMVM	---	---
Between-subject covariate		3dttest++, 3dMEMA, 3dMVM	Partially	Partially
Sophisticated situations	Covariate + within-subject factor	3dLME	---	---
	Subject adjustment in trend analysis			
	Basis functions			
	Missing data			

Overview

- Basic concepts
  - Why do we need to do group analysis?
  - Factor, quantitative covariates, main effect, interaction, …
- Various group analysis approaches
  - Regression (*t*-test): 3dttest++, 3dMEMA, 3dttest, 3RegAna
  - AN(C)OVA: 3dANOVAx, 3dMVM, GroupAna
  - Quantitative covariates: 3dttest++, 3dMEMA, 3dMVM, 3dLME
  - Impact & consequence of SFM, SAM, and SEM
- Miscellaneous
  - Issues regarding result reporting
  - Intra-Class Correlation (ICC)
  - Nonparametric approach and fixed-effects analysis
- No routine statistical questions, only questionable routines!