Notizia dell'AFNI **Do We Have to Deal with Multiple Comparisons in Neuroimaging?**

Poster #: T546

Gang Chen[1], Yaqiong Xiao[2], Paul A Taylor[1], Justin Rajendra[1], Fengji Geng[2], Tracy Riggins[2], Elizabeth Redcay[2] , Robert W Cox[1]
[1]Scientific and Statistical Computing Core, NIMH / NIH / DHHS, USA; [2]Department of Psychology, University of Maryland, USA
Contact: gangchen@mail.nih.gov

Multiple comparisons in neuroimaging

Standard approach: massively univariate modeling

- As many models as voxels
- **Assumption**: no information shared across voxels

… which has a penalty for multiplicity

- "Correction" via neighborhood leverage
  - Cluster size
  - Permutation

… and several challenges

- Excessive penalty?
- Discrimination against small regions
- Arbitrariness: artificial dichotomization
- Spatial ambiguity
- Vulnerable to *p*-hacking

Food for thought

Sources of problem for conventional approach

- Too many models
- False assumption

Potential improvements

- One model (not maaany initially followed by some correction)
- Information shared across regions

Goals

- No more multiple comparisons issue
- More efficient
- No discrimination
- No spatial ambiguity
- Highlighting instead of hiding
- Less vulnerable to data manipulation

Demo dataset

Resting-state FMRI

- Subjects: *n* = 124
- Individual level: seed-based correlation (seed: right temporo-parietal junction)

Conventional group analysis

- Whole brain analysis for effect of behavior (theory of mind index):

$$y_{ij} = a_j + b_j x_i + \epsilon_{ij}, i = 1, 2, ..., n$$

- Surviving clusters: up to 4

voxel p	cluster threshold	surviving ROIs	ROIs
0.001	28	2	R PCC, PCC/PrC
0.005	66	4	R PCC, PCC/PrC., L IPL, L TPJ
0.01	106	4	R PCC, PCC/PrC., L IPL, L TPJ
0.05	467	4	R PCC, PCC/PrC., L IPL, L TPJ

Conventional region-based group analysis

- Group analysis with 21 predefined ROIs: 21 GLMs

$$y_{ij} = a_j + b_j x_i + \epsilon_{ij}, i = 1, 2, ..., n$$

a_j, b_j: freely vary from -∞ to +∞

- Correction for multiplicity: Bonferroni *too costly*

New method: Bayesian multilevel (BML) modeling

Integrative modeling

- One (beautiful) model

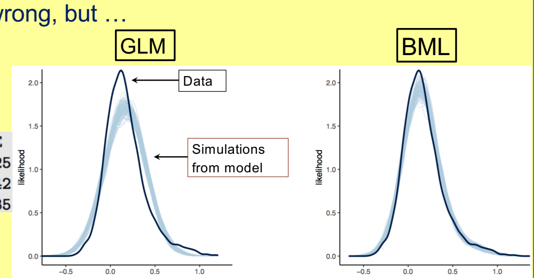
$$y_{ij} = b_0 + b_1 x_i + \pi_i + \xi_{0j} + \xi_{1j} x_i + \epsilon_{ij},\ i = 1, 2, ..., n,\ j = 1, 2, ..., r$$

- Assumption: Gaussian across regions
  - Information shared and regularized among regions: partial pooling
  - Not fully trusting effects estimated individually

Good model?

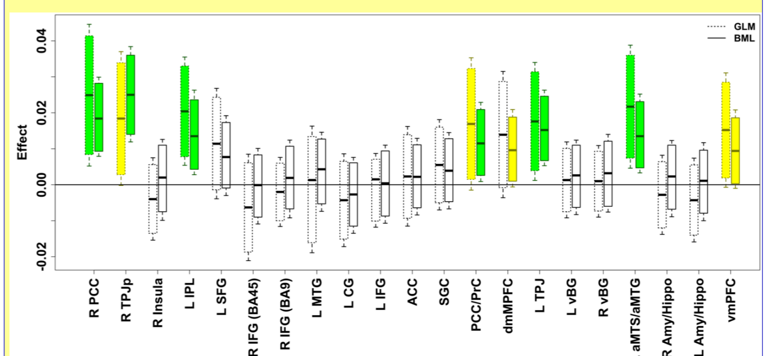
- All models are wrong, but …

	LOOIC	SE
GLM	-300.39	98.25
BML	-2247.06	86.42
GLM - BML	1946.67	96.35



Results comparisons

- GLM (no correction) vs BML



Conclusions

Multiplicity in neuroimaging: result of inefficient modeling

- Information waste: incorrect assumption
- Correction: excessive penalization

Improved approach via Bayesian multilevel modeling (BML)

- Information sharing: more efficient
- Full results reporting: crucial for reproducibility and future studies

This new approach/program available in AFNI: RBA (= "region-based analysis")

Reference

Chen et al., 2019. Handling Multiplicity in Neuroimaging through Bayesian Lenses with Multilevel Modeling. Neuroinformatics. https://rdcu.be/bhhJp

Acknowledgements: This research supported by the NIMH & NINDS Intramural Research Programs of the NIH.