Reliability Measure of NeuroImaging

IntraClass Correlation (ICC): Which Type? How to Compute?

Gang Chen[1], Simone Haller[2], Katharina Kircanski[2], Joel Stoddard[3],
Daniel S. Pine[4], Ellen Leibenluft[2], Melissa Brotman[2], Paul Taylor[1], Robert Cox[1]

[1]Scientific and Statistical Computing Core, NIMH / NIH / DHHS, USA; [2]Section on Mood Dysregulation and Neuroscience, Emotion and Development Branch, NIMH / NIH / DHHS, USA; [3]Division of Child and Adolescent Psychiatry, Department of Psychiatry, University of Colorado School of Medicine, USA; [4]Section on Development and Affective Neuroscience, Emotion and Development Branch, NIMH / NIH / DHHS, USA

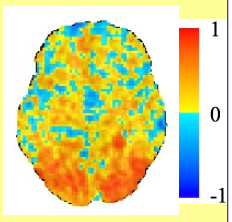
Contact: gangchen@mail.nih.gov http://afni.nimh.nih.gov

Introduction: IntraClass Correlation (ICC)

✧ **Three types**
- **ICC(1,1)**: genetic relatedness for twins
- **ICC(2,1)**: absolute agreement across sessions, scanners, …
  ⇨ **currently most popular**
- **ICC(3,1)**: consistency across sessions, scanners, sites, …

✧ **Conventional model: ANOVA**
- Variance partitioning
- ICC formulated through ratio of Mean Squares (MS)

✧ **Current challenges** for investigators
- **Types**: which one to adopt?
- **Uninterpretable results**: negative or zero ICCs may occur ⇨
- **Precision information** of effect estimates: not considered
- **Missing data**: abandoned
- **Covariates**: can't be included



✧ **These challenges** are addressed with a series of 4 models.

Mixed-Effects Model #1: LME

✧ **LME: ANOVA framed as a special mixed-effects model**
- **ICC(1,1)**

$$y_{ij} = b_0 + \lambda_j + \epsilon_{ij} \qquad \rho_1 = \frac{\sigma_\lambda^2}{\sigma_\lambda^2 + \sigma_\epsilon^2}$$

- **ICC(2,1)**

$$y_{ij} = b_0 + \pi_i + \lambda_j + \epsilon_{ij} \qquad \rho_2 = \frac{\sigma_\lambda^2}{\sigma_\pi^2 + \sigma_\lambda^2 + \sigma_\epsilon^2}$$

- **ICC(3,1)**

$$y_{ij} = b_0 + b_i + \lambda_j + \epsilon_{ij} \qquad \rho_3 = \frac{\sigma_\lambda^2}{\sigma_\lambda^2 + \sigma_\epsilon^2}$$

+ No negative ICCs
+ Type selection: ICC(3,1) is preferred
+ Missing data: easily handled
+ Covariates: can easily be included in model

- Zero ICCs possible: unrealistic
- Precision: not considered

Mixed-Effects Model #2: BME

✧ **Bayesian**: avoid 0 ICC via a weakly informative prior: **Gamma density function** with shape and rate parameters fixed at 2 and 0.5, respectively.

+ Realistic ICCs

- Precision: not considered

Poster # 4135 stand-by time:
Thur. June 29, 12:45 PM- 14:45 PM

Scan me!

Mixed-Effects Model #3: MME

✧ **Incorporate precision into LME**
- Each effect estimate contains measurement error: standard deviation
- Model parameters estimated via weighting

+ More robust

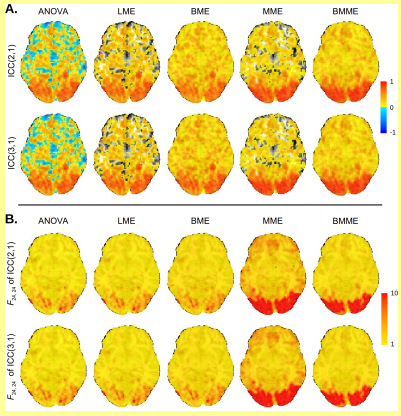
- Zero ICCs still possible: unrealistic

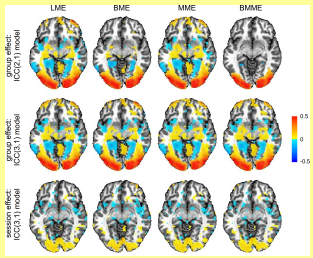
Mixed-Effects Model #4: BMME

✧ **Bayesian: insert weakly informative prior to MME**
+ Avoids 0 ICCs

Performance Comparisons with Real Data

✧ **25 Subjects, 2 Sessions**





Thresholded at two-tailed *p* of 0.1

Conclusions & Recommendations

✧ **Model capabilities in handling issues**

Issues	negative ICC	zero ICC	missing data	confounding effects	sampling error	type selection
ANOVA	✗	✗	✗	✗	✗	✗
LME	✓	✗	✓	✓	✗	✓
BME	✓	✓	✓	✓	✗	✓
MME	✓	✗	✓	✓	✓	✓
BMME	✓	✓	✓	✓	✓	✓

✧ **Recommendation for models**
- **If precision is available, use**: MME, BMME
- **If precision is unavailabe, use**: LME, BME
- **Incorporate potential covariates**: age, sex, …

✧ **Recommendation for ICC type between ICC(2,1) and ICC(3,1)**
- **ICC(3,1)** instead of **ICC(2,1)**: accurate characterization of the data by exploring and understanding potential differences across sessions/scanners/sites

✧ **Result reporting**
- Indicate type and model
- Report group mean and covariate effects in addition to ICC

✧ **Tool - program 3dICC publicly available in AFNI**
- All levels: voxel-wise whole brain, network, ROI / voxel

New

Acknowledgements

The research was supported by the NIMH & NINDS Intramural Research Programs of the NIH. We are indebted to Wolfgang Viechtbauer and Vincent Dorie for their generous help.