Optimizing Lifespan Normative Models for Regional Subcortical Volume, Cortical Thickness, and Surface Area

Ruiyang Ge¹, Yuetong Yu¹, Yi Xuan Qi¹, Yunan Fan², Faye New³, Shalaila Haas², Paul Thompson³, Sophia Frangou¹,²

¹Department of Psychiatry, University of British Columbia, Canada; ²Department of Psychiatry, Icahn School of Medicine at Mount Sinai, USA; ³Imaging Genetics Center, Mark and Mary Stevens Neuroimaging and Informatics Institute, Keck School of Medicine, University of Southern California, USA

BACKGROUND

- Normative brain modeling has recently been used to derive individual-level deviation in quantitative neuroimaging phenotypes. ¹,²,³
- Choice of algorithms lacks empirical support from direct comparative benchmarking.

OBJECTIVES

- Evaluate commonly used normative modeling algorithms in building normative neuroanatomical models
- Establish the minimum sample size required for reliable model performance
- Present normative neuroanatomical models of high generalizability & an open-access web portal to establish the minimum sample size required for reliable model performance

METHODS

1. Model optimization sample & 3 generalizability samples
- 150 FreeSurfer² measures
- 14 subcortical volumes
- 68 cortical thickness measures
- 68 surface area measures

DATA ANALYSIS PROCEDURES

Each FreeSurfer measure of the model optimization sample are stratified by sex, and split into training (80%) and testing (20%) sub-samples for each anatomical measure separately.

RESULTS & CONCLUSIONS

- Results of female participants are presented in this section, and all conclusions can be generalized to males.
- Optimal algorithm: FRP algorithm with its best predictive performance and reasonable computational efficacy
  1. Predictive performance, MAE (mean absolute error): FPR = GPR = GAMLSS = OLSR = BLR
  2. CPU time: OLSR = 10ms, FPR = 500ms, BLR = GAMLSS = 5,000ms, and GPR = 30minutes.
- Optimal model: FRP algorithm with the nonlinear fractional polynomials of age and linear global measure of the corresponding anatomical measure.
- Samples below 1,000 are likely to yield unreliable normative models.
- Collectively, a sample size above 3,000 is a decent size to generate a robust normative model.
- The pre-trained model for each anatomical measure has very high generalizability for all 3 samples with different races/ethnicities.
  1. Black sample vs. model training sample
  2. South Asian sample vs. model training sample
  3. East Asian sample vs. model training sample

- We present an open-access web portal with:
  1. Pre-trained sex-specific normative neuroanatomical models
  2. Functionality to generate individual-level normalised deviations of user data
  3. Full scripts for generating normative models in any user-specified T1-weighted MRI datasets

REFERENCE