Combining/fusing genetics and brain imaging data to make inferences and predictions

OUTLINE:

  I. Ways to approach multimodal data analysis

  II. The use of matrix factorization to combine high-dimensional data

  III. Making inferences about disease and/or variables of interest

  IV. Common pitfalls

  V. Summary

TAKE-HOME MESSAGES/WHAT YOU SHOULD KNOW:

1. An understanding of the potential benefits of combining multimodal imaging and genetics data.
2. A review of the various approaches for analyzing multiple data sets
3. Prediction and regression approaches to high-dimensional data sets

SUGGESTED READING:

Calhoun VD, Liu J, Adalı T. A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data. NeuroImage 2009; 45(S1), 163-171.
J. Chen, V. D. Calhoun, G. D. Pearlson, S. Ehrlich, J. Turner, B. C. Ho, T. Wassink, A. Michael, and J. Liu, "Multifaceted genomic risk for brain function in schizophrenia," NeuroImage, vol. 61, pp. 866-875, 2012
J. Chen, V. D. Calhoun, G. D. Pearlson, N. Perrone-Bizzozero, J. Sui, J. A. Turner, J. Bustillo, S. Ehrlich, S. Sponheim, J. Canive, B. C. Ho, and J. Liu, "Guided Exploration of Genomic Risk for Gray Matter Abnormalities in Schizophrenia Using Parallel Independent Component Analysis with Reference," NeuroImage.
J. Sui, T. Adalı, Q. Yu, and V. D. Calhoun, "A Review of Multivariate Methods for Multimodal Fusion of Brain Imaging Data," Journal of Neuroscience Methods, vol. 204, pp. 68-81, 2012.

Overview:
MROmics can be described as the combination of imaging and genetic data in the context of studying links among (usually) very high-dimensional data with the goal of identifying putative biomarkers, intermediate phenotypes, or associations with some variable of interest in the healthy or diseased brain. In this talk we will review some of the available approaches, with an emphasis on data-driven approaches. We will also address the critical issue of how to make reliable/robust predictions about categorical or continuous variable (e.g. diagnosis, treatment response, cognitive function) in the context of models with a high number of parameters.