Mass Univariate and Multivariate Approaches to Understanding Genetic Variation in the Brain

Thomas Nichols
Department of Statistics,
Warwick Manufacturing Group
University of Warwick

Joint with

Giovanni Montana
Kings College

Tian Ge, Jianfeng Feng
Computer Science, Warwick

Steve Smith
Oxford

Derek Hibbean, Paul Thompson
University of California Los Angeles
Outline

• Multivariate Imaging Genetics
  • Sparse Reduced Rank Regression

• Mass Univariate
  – RFT + Fast Permutation
  – Accelerated Heritability Inference
  – Heritability summaries for ranking
## Imaging Genetics Menu

<table>
<thead>
<tr>
<th>Imaging</th>
<th>Candidate ROI</th>
<th>Many ROI</th>
<th>Voxelwise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genetics</td>
<td><img src="image" alt="Candidate SNP" /></td>
<td><img src="image" alt="Candidate Gene" /></td>
<td><img src="image" alt="Genome-wide SNP" /></td>
</tr>
<tr>
<td>Candidate Gene</td>
<td></td>
<td>[Potkin et al. 2009] 1 BOLD ROI 317, 503 SNPs</td>
<td></td>
</tr>
<tr>
<td>Genome-wide SNP</td>
<td>[Stein et al. 2010] 31,622 voxels 448,293 SNPs</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Jason Stein/Andy Saykin/Bertrand Thirion)
Possible Mass-Univariate Analyses

- Full cross analysis
  - Massive multiple testing problem!

- Candidate SNP
  - Full image result
  - Must have right SNP

- Voxel/Region QTL
  - Whole genome association
  - Must have right ROI

- 500,000 SNPs
- 100,000 voxels
- \( \approx 10^{10} \) tests!

- 500,000 SNPs
- 100,000 voxels
- \( \approx 10^{6} \) tests

- 500,000 SNPs
- 100,000 voxels
- \( \approx 10^{5} \) tests
Multivariate Regression

- **Silly...**
  - If $N > N_G$, fit equivalent to $N_V$ univariate models fit independently
  - Much redundancy in $C$
    - $\text{rank}(C) \leq \min(N_V, N_G) \ll N_V \cdot N_G$

$$\begin{align*}
Y & = X C + E \\
N \times N_V & = N \times N_G \\
C & = N_G \times N_V
\end{align*}$$

$N$  # subjects
$N_V$  # voxels/ROIs
$N_G$  # genes/SNPs
Reduced Rank Regression

\[
Y = XB + A
\]

Images

\(Y\) \(N \times N_V\)

Genotypes

\(X\) \(N \times N_G\)

Genotype Coefficients

\(A\) \(r \times N_V\)

B & A each rank \(r\)

\(N \times r\)

Error

\(E\) \(N \times N_V\)

- Fix rank \(r\)
- Approximate

\[C \approx BA\]

\(N\) \# subjects
\(N_V\) \# voxels/ROIs
\(N_G\) \# genes/SNPs
Sparse Reduced Rank Regression

\[ Y = X B + A N \times NV + E N \times NV \]

• Fix rank \( r \)
• Approximate \( C \approx B A \)
• Enforce sparsity

Sparse Reduced Rank Regression - Estimation

- **RRR**
  - \( Y = X A B + E \)
  - For fixed rank \( r \), find \( A \) & \( B \) that minimize
    \[
    M = \text{tr} \left\{ (Y - XBA) \Gamma (Y - XBA)' \right\}
    \]
    for some \( NV \times NV \) matrix \( \Gamma \), e.g. \( \Gamma = I \)

- **SRRR**
  - For rank 1, find \( a \) & \( b \) that minimize
    \[
    M = \text{tr} \left\{ (Y - Xba') \Gamma (Y - Xba')' \right\}
    + \lambda_a ||a||_1 + \lambda_b ||b||_1
    \]
  - Then subtract \( Xba' \) from the data, and repeat
  - Need to specify final rank \( r \), \( \lambda_a \) & \( \lambda_b \)
    - Can set \( \lambda_a \) & \( \lambda_b \) in terms of \#|a|>0 & \#|b|>0
SRRR Power: Multivariate vs. Mass-Univariate

- **Setting**
  - N=1000 subjects
  - 110 ROI’s, 6 associated
    - Moderate effect size
  - 10 causal SNPs, removed
  - Rank 3 model

- **Power \( \geq 2 \times \) relative to Mass-Univariate

- Absolute power still tiny
  - \( \approx 70\% \) with 2000 SNPs
Real Data – Vounou et al. (2012)

- Phenotype (TBM) Feature Selection
  - Find voxels most predictive of AD vs NC
  - Reduces voxels from 1,650,857 to 11,349

SRR Tuning Parameters

• Objective function
  \[ M = \text{tr} \left\{ (Y - Xba') \Gamma (Y - Xba')' \right\} \]
  \[ + \lambda_a \|a\|_1 + \lambda_b \|b\|_1 \]

• Need to specify tuning parameters
  – Rank \( r \)
  – \( \lambda_a (\#|a|>0) \)
  – \( \lambda_b (\#|b|>0) \)

• Stability Selection  
  Meinshausen & Bühlmann (2010)
  – Randomly sub-sample \( \frac{1}{2} \) your data
  – Fit model for range of tuning parameters
  – Repeat, noting how often each SNP selected
  – Keep those selected \( \geq 50\% \)
Real Data – Vounou et al. (2012)

- SNP inference by Stability Selection

Real Data – Vounou et al. (2012)
SNPs Found

- APOE-ε4 (1.0)
  - Well known and replicated risk factor (10-fold risk!)
- TOMM40 (0.96)
  - Close to APOE gene, recently linked to AD
- BZW1 (0.8)
  - No prior implication, but expressed in brain
  - Also differentially expressed in mouse model of amyotrophic lateral sclerosis (ALS)
- PDZD2 (0.65)
  - Interacts with CST3, implicated in late-onset AD
- YES1 (0.5) 3 SNPs
  - Possible link to AD suggested in the literature
Multivariate Conclusions

- Intuition and simulations suggest multivariate more sensitive
- Sparse RRR
  - Parsimonious, but no P-values
  - Requires alternate inference methods
- Active area of work!
Mass-Univariate Neuroimaging Genetics

• Why bother?
  – Sub optimal, no?

• Multivariate methods struggle with specificity
  – Effect narrowed to some combination of …
    • SNPs
    • Voxels

• Mass-Univariate infers on single voxels/clusters

• Can use 20 years of imaging statistics results
Whole Brain, Whole Genome: “vGWAS” & “vGeneWAS”

• Stein et al. (2010) “vGWAS”
  – Quantitative trait regression at each voxel
  – Run ‘gene-wise’
    • Use PLINK on each and every voxel
    • Bonferroni with effective N over SNPs
    • FDR over brain voxels

• Hibar et al. (2011) “vGeneWAS”
  – PCA Regression on genes
    • F-test on PC’s that account for 95% of SNP variance

Previous Work: vGWAS & vGeneWAS

• Application
  – Tensor-Based Morphometry of ADNI sample
    • 448,293 SNPs (vGWAS) 18,044 genes (vGeneWAS)
    • 31,622 4mm³ voxels, 740 NC+MCI+AD subjs.

• Findings
  – vGWAS: min $P_{FDR} = 0.5$
  – vGenWAS: min $P_{FDR} = 0.3$
  – But suggests vGeneWAS more sensitive

Previous Work’s Limitations

• Works gene-wise
  – Doesn’t use expected spatial structure of genetic-anatomical associations
    • Hibar et al. constructs an omnibus cluster test, but doesn’t localize to specific genes

• Gene-combining
  – PCA regression best to detect linear combination of several SNPs
  – Can’t capture interactions between SNPs

• FDR over the brain
  – Valid under positive dependency, but conservative
    • Robust to smoothness, but doesn’t use info on smoothness
Advancing vGWAS/vGeneWAS

• Work image-wise
  – Can put 20 years of statistical imaging tools to work

• Use Random Field Theory
  – Inferences calibrated to image smoothness

• Permutation with parametric tail fitting
  – When RFT fails, need a “fall back”

• SNP-combining
  – “Kernel-machine” sensitive to epistatic effects

Least Squares Kernel Machines for SNPs within Genes (1)

- Nonparametric regression
  \( G_i \) is \( S \) SNPs for subj \( i \)
  \( h() \) is a some function

\[
y_i(v) = \mathbf{x}_i^T \beta(v) + h^v(G_i) + e_i(v)
\]

- Mixed model version
  Elements of \( K \) set by similarity of SNP’s between subjects

\[
y = X\beta + h + e, \quad h \sim \text{N}(0, \tau K)
\]

\[
k(G_j, G_k) = \frac{1}{2S} \sum_{s=1}^{S} \text{IBS}(G_{j,s}, G_{k,s})
\]

- Amazingly, (score) test for \( H_0: \tau = 0 \) requires no iterative estimation.
  - Beta-hat here from OLS

\[
Q_\tau(\hat{\beta}, \hat{\sigma}^2) = \frac{1}{2\hat{\sigma}^2} (y - X\hat{\beta})^T K (y - X\hat{\beta})
\]

Method Outline

• Run No-SNP GLM
  – Demographic-variables only
  – Fit needed for LSKM; FWHM needed for RFT

• Single-locus (1-SNP) analysis
  – Run $N_{SNP}$ GLMs, 1 per SNP

• Multi-locus (1-100 SNP) analysis
  – Run $N_{gene}$ kernel machine tests, 1 per gene

• Voxel-wise
  – RFT, $P_{FWE}^{brain}$ for each peak

• Cluster-wise
  – Use fast permutation, $P_{FWE}^{brain}$ for each cluster

• Genome-wise Bonferroni correction
  – $P_{FWE}^{brain,SNP}$ or $P_{FWE}^{brain,gene}$
Evaluations: RFT Accuracy

- 10,000 Gaussian simulations
- Peak inference OK
  - Cluster inference valid, but conservative
Evaluations: RFT Accuracy

• 10,000 null permutations of real data
  – Apply RFT each time
• Peak inference OK
  – Cluster inference now invalid! (See also [Silver, et al 2010])

Voxel-wise RFT 😊
Cluster-wise RFT 😞
Evaluations: Permutation Tail Fitting

- 1,000,000 permutation gold standard (blue)
- Fit based on independent 100,000 perms (black)
  - 1,000 samples of 100,000 (gray)
Real Data Result

• Same data as Hibar et al.
  – ADNI TBM, 448,293 SNPs, 18,044 genes, 31,622 4mm$^3$ voxels, 740 NC+MCI+AD subjs.

• Single-locus analysis (vGWAS)
  – Neither attain genome-wide significance
    • Voxel: \( \min_{\text{brain,SNP}} P_{\text{FWE}}^{\text{brain}} = 6.4 \times 10^{-6} \)
    • Cluster: \( \min_{\text{brain,SNP}} P_{\text{FWE}}^{\text{brain}} = 3.5 \times 10^{-6} \)

• Multi-locus analysis (vGeneWAS)
  – Voxel-wise has 12 hits, 2 \( P_{\text{FWE}}^{\text{brain,SNP}} < 0.001 \)
  – Cluster-wise none
## 12 Significant vGeneWAS hits

<table>
<thead>
<tr>
<th>Chr</th>
<th>Gene</th>
<th># of SNP in the gene</th>
<th>Min. $p_{uc}$</th>
<th>Min. peak $p_{FWE}^{Brain}$</th>
<th>Min. peak $p_{FWE}^{Brain, Gene}$</th>
<th>Talairach coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>GRIN2B</td>
<td>131</td>
<td>$1.167 \times 10^{-13}$</td>
<td>$3.213 \times 10^{-9}$</td>
<td>$5.798 \times 10^{-5}$</td>
<td>[20, -84, 44]</td>
</tr>
<tr>
<td>9</td>
<td>X75342</td>
<td>25</td>
<td>$1.963 \times 10^{-13}$</td>
<td>$6.414 \times 10^{-9}$</td>
<td>$1.157 \times 10^{-4}$</td>
<td>[-12, -72, 32]</td>
</tr>
<tr>
<td>7</td>
<td>AK025672</td>
<td>42</td>
<td>$2.214 \times 10^{-12}$</td>
<td>$1.072 \times 10^{-7}$</td>
<td>0.0019</td>
<td>[68, -4, -12]</td>
</tr>
<tr>
<td>1</td>
<td>PGM1</td>
<td>13</td>
<td>$8.274 \times 10^{-12}$</td>
<td>$2.290 \times 10^{-7}$</td>
<td>0.0041</td>
<td>[-48, -76, -16]</td>
</tr>
<tr>
<td>1</td>
<td>BC022483</td>
<td>4</td>
<td>$5.062 \times 10^{-12}$</td>
<td>$3.222 \times 10^{-7}$</td>
<td>0.0058</td>
<td>[12, 40, 4]</td>
</tr>
<tr>
<td>1</td>
<td>AJ249210</td>
<td>22</td>
<td>$8.488 \times 10^{-12}$</td>
<td>$3.441 \times 10^{-7}$</td>
<td>0.0062</td>
<td>[18, 13, 21]</td>
</tr>
<tr>
<td>13</td>
<td>FARP1</td>
<td>101</td>
<td>$1.601 \times 10^{-11}$</td>
<td>$6.010 \times 10^{-7}$</td>
<td>0.0108</td>
<td>[-40, -84, -12]</td>
</tr>
<tr>
<td>18</td>
<td>C18orf58</td>
<td>28</td>
<td>$1.910 \times 10^{-11}$</td>
<td>$8.953 \times 10^{-7}$</td>
<td>0.0162</td>
<td>[-40, -44, 20]</td>
</tr>
<tr>
<td>5</td>
<td>AK092765</td>
<td>7</td>
<td>$8.528 \times 10^{-11}$</td>
<td>$1.505 \times 10^{-6}$</td>
<td>0.0272</td>
<td>[0, 24, 8]</td>
</tr>
<tr>
<td>6</td>
<td>FUT9</td>
<td>38</td>
<td>$7.529 \times 10^{-11}$</td>
<td>$1.770 \times 10^{-6}$</td>
<td>0.0319</td>
<td>[-44, -60, -12]</td>
</tr>
<tr>
<td>10</td>
<td>U69546</td>
<td>52</td>
<td>$9.544 \times 10^{-11}$</td>
<td>$2.271 \times 10^{-6}$</td>
<td>0.0410</td>
<td>[32, -92, -4]</td>
</tr>
<tr>
<td>10</td>
<td>AK131357</td>
<td>22</td>
<td>$9.058 \times 10^{-11}$</td>
<td>$2.422 \times 10^{-6}$</td>
<td>0.0437</td>
<td>[-20, -56, -40]</td>
</tr>
</tbody>
</table>
Top SNPs in GRIN2B

- No individual SNP notable by 1-SNP LSKM analysis
- No rare MAF, less likely that outliers responsible

<table>
<thead>
<tr>
<th>SNP</th>
<th>MAF</th>
<th>in each genotypic group</th>
<th>Min. LSKM $p_{lu}$</th>
<th>Corrected peak $p_{FWE}^B$</th>
<th>Talairach coordinates (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs10845840</td>
<td>0.4412</td>
<td>Maj 226 Hct 375 Min 139</td>
<td>$5.388 \times 10^{-9}$</td>
<td>$1.694 \times 10^{-4}$</td>
<td>[-40, 16, -24]</td>
</tr>
<tr>
<td>rs7301754</td>
<td>0.3101</td>
<td>Maj 356 Hct 299 Min 80</td>
<td>$7.422 \times 10^{-9}$</td>
<td>$2.817 \times 10^{-4}$</td>
<td>[24, -84, 48]</td>
</tr>
<tr>
<td>rs2216344</td>
<td>0.4534</td>
<td>Maj 218 Hct 369 Min 151</td>
<td>$1.976 \times 10^{-8}$</td>
<td>$5.087 \times 10^{-4}$</td>
<td>[16, -88, 40]</td>
</tr>
<tr>
<td>rs220573</td>
<td>0.3480</td>
<td>Maj 316 Hct 333 Min 91</td>
<td>$2.211 \times 10^{-8}$</td>
<td>$8.079 \times 10^{-4}$</td>
<td>[20, -84, 44]</td>
</tr>
<tr>
<td>rs220575</td>
<td>0.3480</td>
<td>Maj 316 Hct 333 Min 91</td>
<td>$2.211 \times 10^{-8}$</td>
<td>$8.079 \times 10^{-4}$</td>
<td>[20, -84, 44]</td>
</tr>
<tr>
<td>rs11055612</td>
<td>0.4953</td>
<td>Maj 179 Hct 381 Min 176</td>
<td>$6.147 \times 10^{-8}$</td>
<td>$1.300 \times 10^{-3}$</td>
<td>[-40, 16, -24]</td>
</tr>
<tr>
<td>rs918168</td>
<td>0.2959</td>
<td>Maj 363 Hct 316 Min 61</td>
<td>$7.654 \times 10^{-8}$</td>
<td>$4.722 \times 10^{-4}$</td>
<td>[-40, 48, 4]</td>
</tr>
<tr>
<td>rs2300267</td>
<td>0.4041</td>
<td>Maj 271 Hct 336 Min 131</td>
<td>$1.730 \times 10^{-7}$</td>
<td>$4.100 \times 10^{-3}$</td>
<td>[20, -84, 44]</td>
</tr>
<tr>
<td>rs2160730</td>
<td>0.3730</td>
<td>Maj 288 Hct 346 Min 103</td>
<td>$2.038 \times 10^{-7}$</td>
<td>$4.900 \times 10^{-3}$</td>
<td>[4, 48, 36]</td>
</tr>
<tr>
<td>rs11055651</td>
<td>0.3378</td>
<td>Maj 295 Hct 328 Min 86</td>
<td>$2.521 \times 10^{-7}$</td>
<td>$2.600 \times 10^{-3}$</td>
<td>[48, -80, -28]</td>
</tr>
</tbody>
</table>
Localization of GRIN2B Effect

- Brain- genome-wide significance (yellow)
- Post-hoc brain-wide (dark blue)
- $P,0.001$ uncorrected (light blue)
Conclusions

• Brain-wide, Genome-wide analyses
  – A huge, but feasible computational challenge
  – Power gains from using standard imaging statistical methods

• Power as serious concern
  – N=700 a tiny genetics sample size!
  – But valid statistical methods invaluable in any such discovery exercise
Mass Univariate Heritability Inference: HCP Motivation

• Human Connectome Project
  – WashU+Minn (see also UCLA/USC-MGH HCP)
  – n = 1,200 subject population sample

• Extended Twins Design
  – Missouri Twins Registry
  – Siblings only
    • One twin pair
    • One or more additional siblings
    • Target family size: 4
Estimating Heritability

**“ACE” Model**

Var(MZ) = Var(DZ) = Var(Sib)

= A + C + E

Cov(MZ₁,MZ₂) = A + C

Cov(DZ₁,DZ₂) = Cov(Sib₁,Sib₂)

= A/2 + C

A = Additive genetic effects

C = Common environment

E = Unique env. & random err.

**HCP Example: h² = 20%, c² = 20%**

- Requires variance components, iterative estimation
- Current methods slooow
  - Mx/OpenMx is current favorite
- New (old) method
  - Linear Regression with Squared Differences (LR-SD)
  - Due to Grimes & Harvey (1980)

Linear Regression on Squared Differences

- Heritability inference without iteration (Grimes & Harvey, 1980)
  - Relate squared differences of data pairs to variance components $A, C, E$:
    
    \[
    E \left[ (MZ_1 - MZ_2)^2 \right] = 2E \\
    E \left[ (DZ_1 - DZ_2)^2 \right] = A + 2E \\
    E \left[ (l_1 - l_2)^2 \right] = 2A + 2C + 2E
    \]
    
    Then estimate $A, C & E$ with linear regression!

- Modification of Grimes and Harvey’s method: $n(n - 1)/2$ obs. $\rightarrow (n_{MZ} + n_{DZ})/2$ obs. (50,721 vs. 141)

- Permutation Inference
  - Under $H0$: $h^2 = 0$, MZ and DZ twin pairs are exchangeable
  - $\binom{n_{MZ} + n_{DZ}}{n_{MZ}/2}$ possible permutations

- Still use parametric likelihood to conduct inference w/ LRT
  
  
  Chen & Nichols, *in preparation*
LR-SD Evaluations

Simulation Setting

- 10,000 simulations
- Sample sizes: 10+10, 50+50
- 15 ACE parameter settings:

<table>
<thead>
<tr>
<th></th>
<th>E</th>
<th>CE</th>
<th></th>
<th></th>
<th>AE</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.2</td>
<td>0.3</td>
<td>0.5</td>
<td>0.7</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>0.2</td>
<td>0.3</td>
<td>0.5</td>
<td>0.7</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>1</td>
<td>0.8</td>
<td>0.7</td>
<td>0.5</td>
<td>0.3</td>
<td>0.8</td>
<td>0.7</td>
<td>0.5</td>
</tr>
</tbody>
</table>

|     |     |     |     |     | ACE |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|
| A   | 0.2 | 0.3 | 0.2 | 0.5 | 0.3 | 0.2 |
| C   | 0.2 | 0.2 | 0.3 | 0.2 | 0.3 | 0.5 |
| E   | 0.6 | 0.5 | 0.5 | 0.3 | 0.3 | 0.3 |
LR-SD Evaluations

Simulations: MSE Comparison

Mean squared error comparison between LR-SD and OpenMx

Chen & Nichols, in preparation
SD-LR Evaluations

Simulations: Power Comparison

Statistical power comparison between LR-SD and OpenMx

n = 20:
\( n_{MZ} = 10 \)
\( n_{DZ} = 10 \)

n = 100:
\( n_{MZ} = 50 \)
\( n_{DZ} = 50 \)

Chen & Nichols, *in preparation*
Simulations: Running Time Comparison

Overall running time comparison between LR-SD and OpenMx
→ On average, our LR-SD is around 300 times faster than OpenMx

<table>
<thead>
<tr>
<th>n</th>
<th>nMZ</th>
<th>nDZ</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>100</td>
<td>50</td>
<td>50</td>
</tr>
</tbody>
</table>
**fMRI Work Memory Example**

- Blokland et al. 2011
- N-back, 2-back vs. 0-back contrast
- n = 319: n\textsubscript{MZ} = 150, n\textsubscript{DZ} = 132, n\textsubscript{UR} = 37
  - 199F/120M, Age 20 – 28 (mean ± SD: 23.6 ± 1.8 yr)
- Age, gender & 2-back performance covariates
- Results
  - Running time:
    - One perm: LR-SD 6 min, Mx 2 days
    - 1000 perms, 10x parallelisation: LR-SD 15.5 hours
fMRI Work Memory Example: Results

- Voxel-wise: Min $P_{FWE}$ 0.006, 3 voxels at 5% FWE
- Clusters-wise:
  Min $P_{FWE}$ 0.003
  3 clusters at 5% FWE (127, 201, & 210 voxels)
Fast Heritability Conclusions

• LR-SD method so fast allows permutation
• Allows standard imaging statistics (on LRT image)
  – FWE corrected voxel/peak, cluster size, cluster mass, etc.
• Limitations
  – No permutation test for $c^2$; never any CI’s
  – Use bootstrap for CI’s
• Future work
  – Genetic correlations
• Software
  – APACE: Accelerated Permutation Inference for ACE model  Jun 2014!
Univariate Imaging Genetics: Heritability Ranking

• Whither heritability
  – Need heritability to call a trait a “phenotype”
  – But, for brains, when is $H_0: h^2 = 0$ ever true!?
  – Still, useful as an measure of biological validity

• Example: Preprocessing comparisons
  – Resting state BOLD fMRI
    • Global time course regression (yes or no)
    • Low vs. high dimensional inter-subject registration
  – Gray Matter structural morphometry
    • Use of auxiliary data: functional/diffusion (yes or no)

Joint Work
Steve Smith, Oxford
Xu Chen, Warwick
HCP Resting State fMRI: ICA-based Functional Connectivity

• For each subject:
  – 200 ICA components (based on 1h of data!)
    • 1 time series, 1 spatial map per IC
  – $200 \times 200$ network matrix
    • Correlation (full or partial) between each IC’s time series
  – Unwrap: $200 \times 199/2 = 19,900$ edges

• Network Matrix
  – $N_{\text{subject}} \times 19,900$
    matrix of resting func. conn. strength
Engineers Gone Wild!

- For each pair of Netmat rows, compute correlation coefficients $r$
  - 209 subjects,
    $209 \times 208/2 = 21,736$ r’s
- Plot by relationship
  - Test for MZ-DZ difference
  - Evidence of “heritability”
- What!?!?
  - Heritability: inter-subject correlations, per phenotype
  - Not: intra-subject, cross-phenotype correlations
What exactly do these r’s mean!?

\[
\begin{align*}
E(\langle r_{MZ} \rangle) & \approx \frac{\text{Var}(\mu)}{\sigma^2} + \frac{h^2 + c^2}{\text{Var}(\mu)} - \frac{\text{ERV}}{\text{Var}(\mu)} + \frac{1 - \rho^P}{\text{Var}(\mu)} \\
E(\langle r_{DZ} \rangle) & \approx \frac{\text{Var}(\mu)}{\sigma^2} + \frac{1}{2}h^2 + c^2 - \frac{1}{2}\text{ERV} - \frac{1 - \rho^P}{\text{Var}(\mu)} \\
E(\langle r_{UR} \rangle) & \approx \frac{\text{Var}(\mu)}{\sigma^2} - \frac{1 - \rho^P}{\text{Var}(\mu)}
\end{align*}
\]

- Huge influence of phenotype mean
  - Variance of mean constant effect
  - Demean, then \( \text{Var}(\mu) = 0 \)

\[\text{Var}(\mu) \quad \text{Variance of mean } \mu_j \text{ of voxel } j\]
\[\mu_j \quad \text{Mean of voxel } j\]
\[\overline{\sigma^2} = \frac{1}{J} \sum_j \sigma^2_j \quad \text{Average of voxel variance}\]

voxel = phenotype element
netmat element
What do these mean, w/out the mean?

\[ E(\langle r_{MZ} \rangle) \approx \frac{\tilde{h}^2 + \tilde{c}^2 - \tilde{ERV}}{1 - \rho^P} \]

\[ E(\langle r_{DZ} \rangle) \approx \frac{\frac{1}{2}\tilde{h}^2 + \tilde{c}^2 - \frac{1}{2}\tilde{ERV}}{1 - \rho^P} \]

\[ E(\langle r_{UR} \rangle) = 0 \]

- So a group comparison gives...

\[ E(\langle r_{MZ} \rangle - \langle r_{DZ} \rangle) \approx \frac{1}{2} \frac{\tilde{h}^2 - \tilde{ERV}}{1 - \rho^P} \]

Variance-weighted average heritability

\[ \tilde{h}^2 = \frac{1}{J} \sum_j \left( \frac{\sigma_j^2}{\sigma^2_j} \right) h_j^2 \]

Variance-weighted avg. common var.

\[ \tilde{c}^2 = \frac{1}{J} \sum_j \left( \frac{\sigma_j^2}{\sigma^2} \right) c_j^2 \]

Variance-weighted ERV

\[ \tilde{ERV} = \frac{2}{J(J - 1)} \sum_{j>j'} \left( \frac{\sigma_j \sigma_{j'}}{\sigma^2} \right) ERV_{jj'} \]

ERV: Heritability \times genetic correlation

\[ ERV_{jj'} = h_j h_{j'} \rho_{jj'}^G \]
What does a difference in means mean?

- What if this effect is significant?

\[ E(\langle r_{MZ} \rangle - \langle r_{DZ} \rangle) \approx \frac{1}{2} \frac{h^2 - \overline{ERV}}{1 - \rho^P} > 0 \]

- Indicates significant heritability (-ish)
  - Reduced by \( \overline{ERV} \) amplified by \( (1 - \rho^P)^{-1} \)

Variance-weighted average inter-voxel correlation

\[ \rho^P = \frac{2}{J(J - 1)} \sum_{j > j'} \left( \frac{\sigma_j \sigma_{j'}}{\sigma^2} \right) \rho_{jj'}^P \]

- But! It’s a valid test for any heritability!

\[ H_0 : h_j^2 = 0 \ \forall j \ \Rightarrow \ \overline{ERV}_{jj'} = 0 \ \forall j, j' \ \Rightarrow \ E(\langle r_{MZ} \rangle - \langle r_{DZ} \rangle) = 0 \]
Applications?

• “Aggregate Heritability” (AgHe)
  \[
  \text{AgHe} = 2(\langle r_{\text{MZ}} \rangle - \langle r_{\text{DZ}} \rangle) \approx \frac{\hat{h}^2 - \overline{\text{ERV}}}{1 - \rho^P} \approx \hat{h}^2
  \]
  
  – Biased estimate of variance-weighted heritability

• High-dim. Phenotype Heritability Ranking
  – Mean \( h^2 \), \( \overline{h}^2 \)
    • \( h^2 \) computed at each element/voxel, then averaged
  – Var-Weighted Mean \( h^2 \), \( \hat{h}^2 \)
    • For BOLD phenotypes, not so crazy!
    • Most active voxels most variable

\[
\hat{h}^2 = \frac{1}{J} \sum_{j} \left( \frac{\sigma_j^2}{\sigma^2} \right) h_j^2
\]
AgHe Properties

• Assess via Monte Carlo simulation
  – 1000-dimensional phenotypes
  – 2 sample size settings
    • $N_{\text{subject}} = 58$ ($N_{MZ} 32$, $N_{DZ} 26$)
    • $N_{\text{subject}} = 580$ ($N_{MZ} 320$, $N_{DZ} 260$)
  – Heterogeneous variance: $\sigma_j^2 = j$
  – Phenotypic correlation: $\rho^P = 0.1665$ ($\rho^P = 0.2143$)
  – Range of heritability & common env.’s (but ERV = 0)
  – Consider raw data, demeaned & standardized
    • Demeaned, AgHe $\approx \hat{h}^2$  Standardized, AgHe $\approx \hat{h}^2$

• Measure
  – Bias, Sd & MSE of AgHe vs true $\overline{h^2}$ & $\hat{h}^2$
AgHe Accuracy vs. $h^2$ (Var-Wt mn. $h^2$)

$n = 58$

Bias comparison for AgHe ~ $h^2$: raw data (blue) vs. demeaning (green) vs. demeaning & variance-normalisation (red)

Relative bias ≈ 20%

Stdev comparison for AgHe ~ $h^2$: raw data (blue) vs. demeaning (green) vs. demeaning & variance-normalisation (red)

MSE comparison for AgHe ~ $h^2$: raw data (blue) vs. demeaning (green) vs. demeaning & variance-normalisation (red)

(a^2, c^2)
AgHe Accuracy vs. $\bar{h^2}$ (mean $h^2$)

n=58

Bias comparison for AgHe ~ mean h2: raw data (blue) vs. demeaning (green) vs. demeaning & variance-normalisation (red)

- **Raw**
- **Demeaned**
- **Standardized**

MSE comparison for AgHe ~ mean h2: raw data (blue) vs. demeaning (green) vs. demeaning & variance-normalisation (red)

- **Raw**
- **Demeaned**
- **Standardized**

Rel Bias ≈ 28%
HCP Phenotype Ranking

• 22 HCP Phenotypes...
  – nElm = 3k-60k

• For each
  – Compute AgHe, $\bar{h}^2$ & $\tilde{h}^2$
    • APACE used to find P-values & CI’s

• Hypothesis:
  – Ranking will be similar between the 3 methods
  – AgHe most similar to $\tilde{h}^2$ (Var-Wt mean $h^2$)
HCP Phenotype Ranking: Estimates

- Good monotonic relationship
  - Tighter for $\bar{h}^2$ (variance-weighted mean $h^2$)

$\text{AgHe vs. } \bar{h}^2$

$\text{AgHe vs. } \bar{h}^2$

$r = 0.81$

$r = 0.89$
HCP Phenotype Ranking: P-values

- Good agreement for strong significance
  - AgHe more optimistic... possibly due to $(1 - \tilde{\rho}^P)^{-1}$

P-values: AgHe vs. $\overline{h}^2$

P-values: AgHe vs. $\tilde{h}^2$
Non-High-Dimensional Phenotypes: Not so good

- Ranking of group 200-dimensional ICA
  - 200 phenotypes, each with 199 elements: Connection strength to each other node
- AgHe not so biased, but huge variance

\[ \text{AgHe vs. } \overline{h^2} \]

\[ \text{AgHe vs. } \tilde{h^2} \]
AgHe: Conclusions

• AgHe
  – Trivially fast
  – Easy to explain to non-imagers
  – Biased estimate of $\tilde{h}^2$ but potentially useful for ranking
  – Best for high-dimensional phenotypes ($n_{Elm} \geq 1k$)
    • High variance reduces utility otherwise
    • But don’t need its speed for low-dim phenotypes
  – If nothing else, suggested $\tilde{h}^2$ (i.e. get it exactly w/ APACE)

• Great utility for HCP
  – Quick vetting of 100’s of imaging phenotypes, preprocessing options
  – Especially useful for sorting multitudes of rs-fMRI options
FINAL Conclusions

• Imaging Genetics needs
  – Powerful, computational tractable methods...
    • To find weak association
    • To measure heritability in high-dimension phenotypes

• Everyone’s toolbox should include both
  – Univariate tools
    • “Default first pass” if nothing else
  – Multivariate tools
    • If distributed effects can be captured well by the model