UC San Diego
UC San Diego Previously Published Works

Title

Generalization of cortical MOSTest genome-wide associations within and across samples

Permalink

https://escholarship.org/uc/item/893342rx

Authors

Loughnan, Robert J

Shadrin, Alexey A

Frei, Oleksandr

et al.

Publication Date

DOI

Copyright Information

Peer reviewed

Generalization of Cortical MOSTest Genome-Wide Associations Within and Across Samples

Robert J. Loughnan[1,2], **Alexey A. Shadrin**[4], **Oleksandr Frei**[3,4], **Dennis van der Meer**[3,5], **Weiqi Zhao**[1], **Clare E. Palmer**[6], **Wesley K. Thompson**[2,7], **Carolina Makowski**[8,9], **Terry L. Jernigan**[1,6,9,10], **Ole A. Andreassen**[4], **Chun Chieh Fan**[2,8], **Anders M. Dale**[1,8,9,11]

[1]Department of Cognitive Science, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA

[2]Population Neuroscience and Genetics, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92161, USA

[3]NORMENT Centre, Division of Mental Health and Addiction, Oslo University Hospital & Institute of Clinical Medicine, University of Oslo, Oslo, Norway

[4]Centre for Bioinformatics, Department of Informatics, University of Oslo, Oslo, Norway

[5]School of Mental Health and Neuroscience, Faculty of Health, Medicine and Life Sciences, Maastricht University, The Netherlands

[6]Center for Human Development, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92161, USA

[7]Herbert Wertheim School of Public Health, University of California, San Diego, La Jolla, CA, USA

[8]Center for Multimodal Imaging and Genetics, University of California, San Diego School of Medicine, 9444 Medical Center Dr, La Jolla, CA 92037, USA

[9]Department of Radiology, University of California, San Diego School of Medicine, 9500 Gilman Drive, La Jolla, CA 92037, USA

[10]Department of Psychiatry, University of California, San Diego School of Medicine, 9500 Gilman Drive, La Jolla, CA 92037, USA

[11]Department of Neuroscience, University of California, San Diego School of Medicine, 9500 Gilman Drive, La Jolla, CA 92037, USA

Abstract

Genome-Wide Association studies have typically been limited to single phenotypes, given that high dimensional phenotypes incur a large multiple comparisons burden: ~1 million tests across the genome multiplied by the number of phenotypes. Recent work demonstrates that a Multivariate

Omnibus Statistic Test (MOSTest) is well powered to discover genomic effects distributed across multiple phenotypes. Applied to cortical brain MRI morphology measures, MOSTest has resulted in a drastic improvement in power to discover loci when compared to established approaches (min-P). One question that arises is how well these discovered loci replicate in independent data. Here we perform 10 times cross validation within 34,973 individuals from UK Biobank for imaging measures of cortical area, thickness and sulcal depth (>1,000 dimensionality for each). By deploying a replication method that aggregates discovered effects distributed across multiple phenotypes, termed PolyVertex Score (PVS), we demonstrate a higher replication yield and comparable replication rate of discovered loci for MOSTest (# replicated loci: 309–747, replication rate: 93–94%) in independent data when compared with the established min-P approach (# replicated loci: 28–59, replication rate: 67–81%). An out-of-sample replication of discovered loci was conducted with a sample of 4,069 individuals from the Adolescent Brain Cognitive Development® (ABCD) study, who are on average 50 years younger than UK Biobank individuals. We observe a higher replication yield and comparable replication rate of MOSTest compared to min-P. This finding underscores the importance of using well-powered multivariate techniques for both discovery and replication of high dimensional phenotypes in Genome-Wide Association studies.

Introduction

Performing Genome Wide Association Studies (GWAS) on high dimensional phenotypes incurs a large multiple comparisons burden (number of independent genetic tests by number of phenotypes) using traditional approaches, which can result in low power to detect associations. Vertex-wise measures of cortical morphology (area, thickness and sulcal depth) represent high dimensional phenotypes (>1000 dimensions) and, from twin studies, are known to have high heritabilities of up to 90% and 50% for total and regional area respectively, and 80% and 60% for mean and regional thickness respectively[1,2]. Our group has previously developed a novel Multivariate Omnibus Test (MOSTest) [3–5], which aggregates the effect of a genomic variant across the cortex. This method significantly boosts discovery of genetic loci linked to cortical morphology, with an up to 10x increase in number of loci discovered – when compared to an established approach (min-P) deployed for the same phenotypes[5]. Additionally, discovered loci show strong enrichment with pathways involved in neurogenesis and cell differentiation. Two main benefits of MOSTest over established techniques, like min-P, are: 1) its ability to aggregate pleiotropic effects into a single statistical test and 2) it drastically reduces the multiple comparison burden across the dimensionality of phenotypes, while still accounting for genome-wide multiple comparisons correction. Given such a dramatic increase in discovery of genomic loci, it is of interest to understand how well these discoveries replicate in independent data.

Here we perform 10-times cross validation with simulated and real brain imaging data taken from the UK Biobank, and randomly split the sample into $\frac{2}{3}$ training and $\frac{1}{3}$ replication splits. For the training samples we perform discovery of vertex-wise measures of area, thickness and sulcal depth as in [4]. Having discovered genomic loci in training folds, we perform replication of these loci in the test sets. To perform replication for each SNP we calculate a PolyVertex Score (PVS) (similar to [6,7]) from imaging data in the test set for each

MOSTest discovered locus. This PVS aggregates the distributed effects across the cortex by taking a weighted sum across all vertices using mass univariate z statistics as weights from the training set. This approach is similar to the widely used method of Polygenic Risk Scores (PRS) in genetics[8], where instead of predicting a phenotype we are predicting a single genomic variant and instead of using distributed effects across the genome as predictors we use the distributed effects across the cortex, estimated in the training set. For each discovered loci in the training set we generate a PVS for each individual, which represents a continuous prediction of the genotype in the test set. We then correlate each PVS with its corresponding measured genomic variant in the test set to test how well these discovered loci replicate (one tailed t test, p<0.05). We test this MOSTest discovery and PVS replication, against an established GWAS approach (min-P)[9]. Figure 1 displays a schematic of how replication of how min-P and MOSTest differs for a single discovered variant. Firstly, we perform this process on simulated data under a range of different genetic architectures to demonstrate higher power to detect associations for MOSTest vs min-P while appropriately controlling type-I errors. Next, repeating this for real imaging data from the UK Biobank, we confirm a higher replication yield and comparable replication rate MOSTest versus min-P. Finally, we test the generalization of loci discovered in UK Biobank to a developmental cohort of 9–10 year old children from the Adolescent Brain Cognitive Development® (ABCD; https://abcdstudy.org) Study, where we see a higher yield of replicated loci for MOSTest versus min-P.

Results

Simulations

Across a range of simulated conditions MOSTest exhibited greater power than min-P to discover and replicate casual variants – see Figure 2. At highly heritable, low polygenic ($h^2 = 0.04$, nc=10) and low heritable, highly polygenic ($h^2 = 0.004$, nc=10,000) phenotypes the two methods showed comparable power. Both min-P and MOSTest demonstrated a well-controlled type 1 error rate, with MOSTest having zero and min-P having one false positive replication across all simulations and folds. For min-P across 8 sets of simulated phenotypes, 10 folds and 7.2 million variants this single false positive represents an error rate of less than 10^{-8}. This low type 1 error rate for both methods is likely due to the stringent criteria required for variants to be both discovered (p<10^{-8}) in training and replicated (0.05) on test data.

Real Data

Across training folds, the UK Biobank sample, we confirm that MOSTest confers up to a 10-fold increase in discovered loci over min-P – see solid bars Figure 3. When replication of loci is defined at the nominal level (p<0.05, see methods) we see a higher number of replicated loci for MOSTest (area: 396, thickness: 309, sulcal depth: 747) vs min-P (area: 32, thickness:28, sulcal depth: 59), as well as comparable replication rate for MOSTest (area:93%, thickness: 94%, sulcal depth: 93%) vs min-P (area:67%, thickness: 81%, sulcal depth: 79%) – see hatched bars Figure 3. Averaged across cross-validation folds, we found that the lead SNP of the top locus accounted for more variance in the replication set with MOSTest (R^2 (95%CI)= area:0.029 (0.024–0.034), thickness: 0.058 (0.040–0.076), sulcal

depth: 0.043 (0.037–0.049)) compared to min-P (R^2 (95%CI)= area: 0.0092 (0.0062–0.012), thickness: 0.012 (0.010–0.14), sulcal depth: 0.014 (0.0095–0.019)). If replication is defined more conservatively with significance corrected for the number of discovered loci (p<0.05/# of discovered loci), we again find that MOSTest confers a comparable replication rate (area: 65%, thickness: 70%, sulcal depth: 61%) to min-P (area: 42%, thickness: 62%, sulcal depth: 53%).

Next, we tested the generalization performance of loci discovered in each training fold of UK Biobank to a developmental cohort of adolescents from the Adolescent Brain Cognitive Development study. Here we once again see a higher absolute number of replicated loci (nominal p<0.05 level), as well as a comparable replication rate for MOSTest (area: 63%, thickness: 60%, sulcal depth: 63%) to min-P (area: 44%, thickness: 45%, sulcal depth: 47%) - see Figure 4. Again, the variance explained by the lead SNP of the top locus (averaged across cross-validation folds) accounted for more variance for MOSTest (R^2 = area: 0.019 (0.016–0.022), thickness: 0.042 (0.025–0.059), sulcal depth: 0.033 (0.030–0.036)) than for min-P (R^2 = area: 0.0083 (0.0073–0.0093), thickness: 0.0066 (0.0046–0.0086), sulcal depth: 0.013 (NA-NA*)). * Confidence interval was not estimable due to zero variance across folds.

Discussion

In simulations we demonstrated the increased power of MOSTest vs min-P to replicate through the deployment of a PVS method, whilst exhibiting appropriate control of type-I errors. On real imaging data from UK Biobank we show a higher replication yield and comparable replication rate of MOSTest compared to a conventional min-P approach. The comparable replication rate for MOSTest loci (93–94% vs 67–81% for min-P) indicates that the difference in absolute number of replicated loci for MOSTest vs min-P is not merely a result of MOSTest discovering a higher number of loci. Furthermore, we still see a comparable replication rate when we penalize the replication significance threshold by the number of loci discovered by each method (i.e. p<0.05/ # of discovered loci). This underscores the distributed effects of the genome across the cortex, which multivariate methods are better powered to capture and in turn, will display stronger generalization to independent data.

Additionally, we have shown that genetic-cortical morphology associations learned within an adult population (mean age 64 years) of individuals from the UK generalize out of sample to adolescents aged 9–10 years old in the United States of America taken from the ABCD study. There are marked differences between the training sample of UK Biobank and validation sample of ABCD including: large age differences, different scanners used, imaging protocols and the number of individuals in validation sets. In spite of these differences we observe a high replication rate in ABCD of discoveries found within UK Biobank via MOSTest. We see higher replication for cortical area and sulcal depth in ABCD than for cortical thickness. Cortical thickness changes more dynamically over the lifespan[10], therefore, given the large age disparity between the two samples, perhaps it is not a surprise to see that cortical thickness is the measure that exhibits the largest reduction

in replication rates in ABCD when compared across cross-validation folds of UK Biobank for MOSTest (60% vs 94%). We may expect that the replication rate of discovered cortical thickness loci to increase as the children develop, a hypothesis that can be tested as more longitudinal ABCD data is collected. Despite differences across these datasets we observe greater replication of UK Biobank discovered loci in ABCD when taking into account the multivariate nature of associations across the cortex (i.e. MOSTest and PVS).

Furthermore, we demonstrated that lead MOSTest discoveries explained a notable amount of variance out of sample, by GWAS standards: 3–6% in UK Biobank and 2–4% in ABCD. Methods, such as MOSTest and PVS, that result in high replication yield and out of sample variance explained may support precision medicine efforts[11]. In particular if these methods are deployed on disorders of the brain they may provide complimentary predictive power to well established models such as Polygenic Risk Scores.

The training data used here to detect loci and train PVS projections weights were taken from individuals of European ancestry from the UK Biobank. We may expect that the genetic architecture of cortical morphology to differ between ancestry groups[12]. We also acknowledge that our use of PVS to predict genotypes out of sample is just one possible projection weighting scheme, which may not provide optimal out of sample prediction. Here we have demonstrated the high generalization performance of cortical morphology discoveries using MOSTest to independent data. This was shown both within study (UK Biobank) and across studies (UK Biobank to ABCD) despite substantial age differences of participants. This work underscores the importance of deploying well powered multivariate methods when performing GWAS on high dimensional phenotypes, both for discovery and replication.

Methods

The UK Biobank sample used and methods described for min-P and MOSTest discovery overlap with previous work[4,5].

UK Biobank Sample

Genotypes, MRI scans, demographic and clinical data were obtained from the UK Biobank under accession number 27412, excluding participants who withdrew their consent. For this study we selected white British individuals (derived from both self-declared ethnicity and principal component analysis) who were unrelated (no relation greater than 3rd degree relatives using PLINK[13] –king-cutoff 0.0625) and who had undergone the neuroimaging protocol. This was done to obtain a homogenous group of unrelated individuals of a single ancestry as is considered best practice to control for confounds in genetic analyses[14]. The resulting sample contained 34,973 individuals with a mean age of 64.4 years (standard deviation 7.5 years), 18,031 female. T_1-weighted MRI scans were collected from three scanning sites throughout the United Kingdom, all on identically configured Siemens Skyra 3T scanners, with 32-channel receive head coils. We used UK Biobank v3 imputed genotype data[15].

Adolescent Brain Cognitive Development® (ABCD) Sample

The ABCD study is a longitudinal study across 21 data acquisition site following 11,878 children starting at 9 and 10 years old. This paper analyzed the full baseline sample from data release 3.0 (NDA DOI:10.151.54/1519007). The ABCD study used school-based recruitment strategies to create a population-based, demographically diverse sample with heterogeneous ancestry. T_1-weighted MRI scans were collected using Siemens Prisma, GE 750 and Phillips 3T scanners. Scanning protocols were harmonized across 21 acquisition sites. Genetic ancestry factors were estimated using fastStructure[16] with four ancestry groups. Genotype data was imputed at the Michigan Imputation Server[17], using the HRC reference panel as described in[18,19]. We selected individuals who had passed neuroimaging and genetic quality control checks. Additionally, we restricted to individuals who were unrelated (no relation greater than 3rd degree relatives) and of European ancestry (>90% estimated European from fastStructure) as training and testing in the same ancestry groups is considered best practice for genetic analyses[20]. This resulted in 4,069 individuals with a mean age of 9.9 years (standard deviation 0.62 years), 1,889 female.

Data processing

T1-weighted structural MRI scans were processed with the FreeSurfer v5.3 standard "recon-all" processing pipeline[21] to generate 1,284 non-smoothed vertex-wise measures (ico3 downsampling with the medial wall removed) summarizing cortical surface area, thickness and sulcal depth. Vertices with no variation across subjects were removed. All measures were pre-residualized for age, sex, scanner site, Euler number and the first ten genetic principal components. In contrast to other MOSTest work[3,5] we did not pre-residualize for global measures specific to each set of variables (total cortical surface area or mean cortical thickness) as there is no clear analogous global measure for sulcal depth. Subsequently, a rank-based inverse normal transformation was applied to the residualized measures. For genomic data we carried out standard quality-checks as described previously[3], setting a minor allele frequency threshold of 0.5% and finding the intersecting variants between UK Biobank and ABCD, leaving 7,340,618 variants. Variants were tested for association with cortical surface area, cortical thickness and sulcal depth at each vertex using the standard univariate GWAS procedure. Resulting univariate p-values and effect sizes were further combined in the MOSTest and min-P analyses to identify area, thickness and sulcal depth associated loci.

Cross validation

We performed 10 times cross validation within UK Biobank with random $\frac{2}{3}$ training and $\frac{1}{3}$ testing splits. Validation in ABCD was performed across each of these $\frac{2}{3}$ UKB Training samples.

Simulations

Within the UK Biobank sample we generated simulated phenotypes which exhibited effects distributed across a large number of phenotypes following a similar process to previously published work[3]. For this we included the full sample of $N = 34,973$ individuals, across $M=7,340,618$ variants. Using an additive genetic model, we drew genetic effects

β from a gaussian distribution, and then calculated quantitative phenotypes y_i of the ith sample as $y_i = \sum_{j=1}^{M} x_{ij}\beta_j + \varepsilon_i$, where $X = (x_{ij})$ is an N by M genotype matrix, containing the number of reference alleles for the ith sample and j[th] variant β_j is the causal effect size, and ε is a normally distributed residual that yields a pre-defined level of heritability, $h^2 = Var(X\beta)/Var(y)$, for our simulations $h^2 = 0.004$ or 0.04. To simulate a realistic multivariate scenario of genetically and phenotypically correlated phenotypes, we introduced a covariance structure in both β and ε across a total of T = 1,175 phenotypes set to the correlation structure across vertex wise measures of sulcal depth. Each set of T = 1,175 simulated phenotypes had a pre-defined number of causal variants, nc, where nc=10, 100, 10000 and 100,000. These casual variants were restricted to 102,484 markers of chromosome 21 to facilitate the calculation of type-I error rate, as null markers (i.e. those not on chromosome 21) would not be in LD (linkage disequilibrium) with causal markers. Type-I errors were calculated as the number of null variants discovered (genome-wide significance level $p<5\times10^{-8}$) and replicated ($p<0.05$) across training and testing folds, by min-P and MOSTest methods respectively – described below (without pruning described in "Locus Definition" section). Replication power was analogously calculated as the proportion of causal variants discovered and replicated by each method across training and testing folds. In total we generated 8 sets of simulated phenotypes (each of dimensionality T = 1,175) under different genetic architectures at 2 h^2 values and 4 nc values. These simulated phenotypes were then each discovered and replicated using min-P and MOSTest for each of the 10 training and testing folds. Code for generating these simulations can be found at: https://github.com/precimed/mostest/tree/master/simu.

MOSTest Discovery

Consider M variants and P (pre-residualized) phenotypes. Let $z_{j,k}$ be a z-score from the univariate association test between j[th] variant and k[th] (residualized) phenotype, and z_j be the vector of z-scores of the jth variant across phenotypes. Let Y be a matrix of (pre-residualized) phenotypes with N (individuals) rows and P (phenotypes) columns, and R be its correlation matrix. R can be decomposed using singular valued decomposition as $R = USU^T$ (U – unitary matrix, S – diagonal matrix with singular values on its diagonal). Consider the regularized version of the correlation matrix $R = US_rU^T$, where S_r is obtained from S by keeping r largest singular values and replacing the remaining with r^{th} largest. The MOSTest statistics for j[th] variant (scalar) is then estimated as $\chi_j = z_j R_r^{-1} z_j^T$, where regularization parameter is selected separately for cortical area, thickness and sulcal depth to maximize the yield of genome-wide significant loci. As established in previous work[3–5] the largest yield for cortical surface area is obtained with $r = 10$; the optimal choice for cortical thickness and sulcal depth was $r = 20$. For simulations we set $r = 20$ as simulated phenotypes were based on correlation structure of sulcal depth. The distribution of the test statistics under null (CDF_{null}^{most}) is approximated from the observed distribution of the test statistics with permuted genotypes, using the empirical distribution in the 99.99 percentile and Gamma distribution in the upper tail, where shape and scale parameters of Gamma distribution are fit to the observed data. The p-value of the MOSTest test statistic for the jth variant is then obtained as $p_{most} = CDF_{null}^{most}(\chi_j)$.

min-P Discovery

Similar to the MOSTest analysis, consider M variants P and pre-residualized phenotypes. Let $z_{i,j}$ be a z-score from the univariate association test between jth variant and kth (residualized) phenotype and z_j be the vector of z-scores of the jth variant across phenotypes. The min-P statistics for the jth variant is then estimated as $v_j = 2\Phi\left(-\max_{k=1\ldots P}(|z_{j,k}|)\right)$, where Φ is a cumulative distribution function of the standard normal distribution. The distribution of the min-P test statistics under null (CDF_{null}^{minP}) is approximated from the observed distribution of the test statistics with permuted genotypes, using the empirical distribution in the 99.99th percentile and Beta distribution in the upper tail, where shape parameters of Beta distribution (α and β) are fit to the observed data. The p-value of the min-P test statistic for the jth variant is then obtained as $p_{minP} = CDF_{null}^{minP}(v_j)$.

Locus definition

Independent significant SNPs and genomic loci were identified in accordance with the PGC locus definition, as also used in FUMA SNP2GENE[22]. First, we select a subset of SNPs that pass genome-wide significance threshold 5×10^{-8}, and use PLINK[13] to perform a clumping procedure at LD r2=0.6, to identify the list of independent significant SNPs. Second, we clump the list of independent significant SNPs at LD r2=0.1 threshold to identify lead SNPs. Third, we query the reference panel for all candidate SNPs in LD r2 of 0.1 or higher with any lead SNPs. Further, for each lead SNP, it's corresponding genomic loci is defined as a contiguous region of the lead SNPs' chromosome, containing all candidate SNPs in r2=0.1 or higher LD with the lead SNP. Finally, adjacent genomic loci are merged if they are separated by less than 250 KB. Allele LD correlations are computed from EUR population of the 1000 genomes Phase 3 data. Obtained clumps of variants were considered as independent genome-wide significant genetic loci. This process was only performed for min-P and MOSTest discovery of loci for real data, and not on simulated data.

Replication of Discovered Variants

A schematic displaying the difference between min-P and MOSTest replication is displayed in Figure 1. For genome-wide significant loci defined in the training folds, we performed replication in test folds of UK Biobank, as well as the whole sample of ABCD. Let X^{test} represent the genotype matrix of individuals in the test set of N individuals and M variants and Y^{test} represent the phenotype matrix of N individuals and P (pre-residualized) phenotypes. Replication was performed in one of two ways, depending on whether the genetic variant was discovered using min-P or MOSTest. Firstly, for a min-P discovery, implicated by the association statistic $z_{j,k}$, the jth variant, x_j^{test}, is associated with the kth (residualized) phenotype y_k^{test}, in the test set. Secondly, for MOSTest validation the jth discovered loci corresponds to a vector of mass univariate association statistics across all vertices z_j - these are used to generate projection weights to create a PolyVertex Score (PVS) [7], $x_{PVS,j}^{test}$. This approach largely mirrors the use of polygenic scores used in genetics, where here we are aggregating effects of vertices across the cortex. For polygenic scores, it is well known that the correlation structure (i.e. linkage disequilibrium) across

the genome can result in suboptimal out of sample performance. This has motivated techniques like LD-Pred[23] and PRSice[24] to first account for this genomic correlation before generating scores. Similarly, we decorrelate the association statistics, z_j, as $\boldsymbol{w}_j = \widetilde{\mathbf{R}}_r \boldsymbol{z}_j$ using the regularized correlation matrix $\widetilde{\mathbf{R}}_r$ that was learned in the training fold. We then generate the polyvertex score for the jth genomic variant as the dot product of $\boldsymbol{w}_j$ with the (pre-residualized) phenotype matrix, $\boldsymbol{Y}^{test}$, in the test set: $\widehat{\boldsymbol{x}}^{test}_{PVS,j} = \boldsymbol{w}_j \boldsymbol{Y}^{test}$. This PVS, $\widehat{\boldsymbol{x}}^{test}_{PVS,j}$, was then associated with its corresponding genetic variant, x_j, in the test set.

As the phenotype matrix, $\boldsymbol{Y}^{test}$, was pre-residualized for covariates before taking the most significant vertex (min-P) or computing the PVS (MOSTest) we did not need to further control for covariates. For both min-P and MOSTest validation, we calculated one-tailed p values from computed replication t statistics as we assume the effect to be in the same direction for training folds and test sets. To define replicated loci we use a nominal p value threshold of 0.05 for associations. Due to the higher number of discovered loci for MOSTest vs min-P, we additionally report the number of loci validated at a Bonferroni corrected threshold, where this number of independent tests is taken to be the number of discovered loci in the training set. This corrected threshold penalizes MOSTest to a greater extent than min-P for discovering a larger number of loci. We calculate the variance explained by the single lead j[th] variant in the replication sample from t statistics of x_j and degrees of freedom (df) as: $R^2 = \dfrac{t^2}{\left(t^2 + df\right)}$. We report the mean of R^2 across cross validation folds, as well as an estimate of its 95% confidence interval as 1.96 x σ, where σ is the standard deviation of R^2 across cross validation folds. Code for discovery and replication is available at (https://github.com/robloughnan/MOSTest_generalization).

ABCD Acknowledgement

Bibliography

1. Panizzon MS et al. Distinct Genetic Influences on Cortical Surface Area and Cortical Thickness. Cereb. Cortex 19, 2728–2735 (2009). [PubMed: 19299253]

2. Eyler LT et al. A Comparison of Heritability Maps of Cortical Surface Area and Thickness and the Influence of Adjustment for Whole Brain Measures: A Magnetic Resonance Imaging Twin Study. Twin Res. Hum. Genet. 15, 304–314 (2012). [PubMed: 22856366]

3. van der Meer D et al. Understanding the genetic determinants of the brain with MOSTest. Nat. Commun. 11, 3512 (2020). [PubMed: 32665545]

4. van der Meer D et al. The genetic architecture of human cortical folding. bioRxiv 2021.01.13.426555 (2021). doi:10.1101/2021.01.13.426555

5. Shadrin AA et al. Multivariate genome-wide association study identifies 780 unique genetic loci associated with cortical morphology. bioRxiv 2020.10.22.350298 (2021). doi:10.1101/2020.10.22.350298

6. Zhao W et al. The Bayesian polyvertex score (PVS-B): a whole-brain phenotypic prediction framework for neuroimaging studies. bioRxiv 813915 (2019). doi:10.1101/813915

7. Zhao W et al. Individual Differences in Cognitive Performance Are Better Predicted by Global Rather Than Localized BOLD Activity Patterns Across the Cortex. Cereb. Cortex 31, 1478–1488 (2021). [PubMed: 33145600]

8. Lewis CM & Vassos E Polygenic risk scores: from research tools to clinical instruments. Genome Med. 12, 44 (2020). [PubMed: 32423490]

9. Grasby KL et al. The genetic architecture of the human cerebral cortex. (2018).

10. Walhovd KB, Fjell AM, Giedd J, Dale AM & Brown TT Through Thick and Thin: a Need to Reconcile Contradictory Results on Trajectories in Human Cortical Development. Cereb. Cortex 27, 1472–1481 (2017). [PubMed: 28365755]

11. Denny JC & Collins FS Precision medicine in 2030—seven ways to transform healthcare. Cell 184, 1415–1419 (2021). [PubMed: 33740447]

12. Fan CC et al. Modeling the 3D Geometry of the Cortical Surface with Genetic Ancestry. Curr. Biol. 25, 1988–1992 (2015). [PubMed: 26166778]

13. Chang CC et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 4, (2015).

14. Marees AT et al. A tutorial on conducting genome-wide association studies: Quality control and statistical analysis. Int. J. Methods Psychiatr. Res. 27, 1–10 (2018).

15. Bycroft C et al. The UK Biobank resource with deep phenotyping and genomic data. Nature 562, 203–209 (2018). [PubMed: 30305743]

16. Raj A, Stephens M & Pritchard JK FastSTRUCTURE: Variational inference of population structure in large SNP data sets. Genetics 197, 573–589 (2014). [PubMed: 24700103]

17. Das S et al. Next-generation genotype imputation service and methods. Nat. Genet. 48, 1284 (2016). [PubMed: 27571263]

18. Loughnan RJ et al. Gene-experience correlation during cognitive development: Evidence from the Adolescent Brain Cognitive Development (ABCD) Study. bioRxiv 637512 (2021). doi:10.1101/637512

19. Palmer CE et al. Delineating genetic and familial risk for psychopathology in the ABCD study. medRxiv 2020.09.08.20186908 (2020). doi:10.1101/2020.09.08.20186908

20. Duncan L et al. Analysis of polygenic risk score usage and performance in diverse human populations. Nat. Commun. 10, (2019).

21. Dale AM et al. Automatically Parcellating the Human Cerebral Cortex. Cereb. Cortex 14, 11–22 (2004). [PubMed: 14654453]

22. Watanabe K, Taskesen E, van Bochoven A & Posthuma D Functional mapping and annotation of genetic associations with FUMA. Nat. Commun. 8, 1826 (2017). [PubMed: 29184056]

23. Vilhjálmsson BJ et al. Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. Am. J. Hum. Genet. 97, 576–592 (2015). [PubMed: 26430803]

24. Choi SW, Mak TS-H & O'Reilly PF Tutorial: a guide to performing polygenic risk score analyses. Nat. Protoc. 15, 2759–2772 (2020). [PubMed: 32709988]
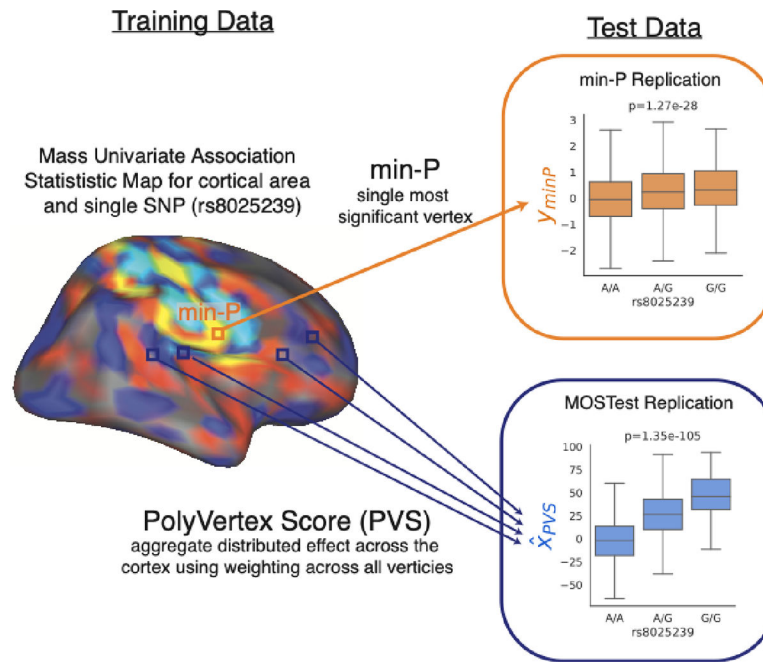
Figure 1.
Schematic of replication process for a single SNP. Variant rs8025239 is discovered in training fold and has mass univariate map of association statistics with cortical area. Min-P replication (indicated by orange box and arrow) takes most significant vertex and associates that vertex with variant rs8025239 in test data. MOSTest replication (indicated by blue box and arrow) computes a PolyVertex Score (PVS) in test data which aggregates all effects across the cortex by taking a weighted sum (using association statistics from training set) across all vertices – the PVS is then correlated with the variant rs8025239. This process is repeated for all discovered variants in training set with a separate PVS being generated for each MOSTest discovery. Replication of a variant is defined as $p<0.05$ in one tailed t-test.
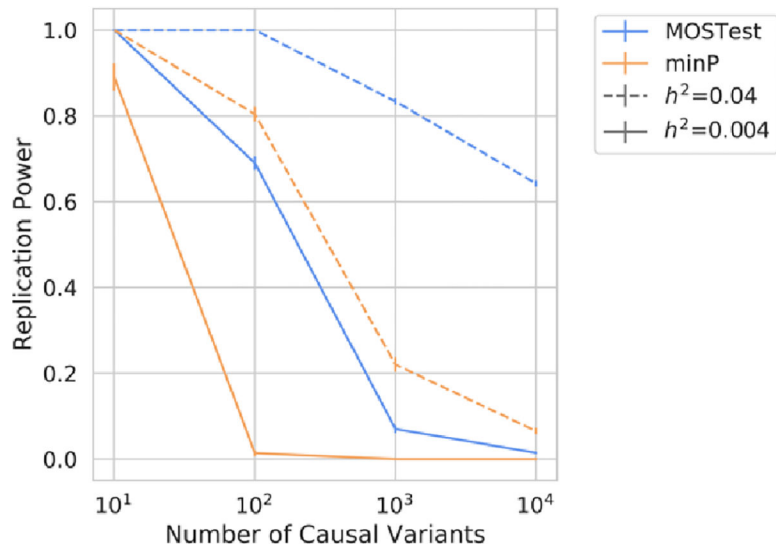
Figure 2.
Power to detect and replicate causal associations from simulated high dimensional phenotypes (T=1,175) for min-P and MOSTest under different genetic architectures. Error bars represent standard deviations across cross-validation folds.
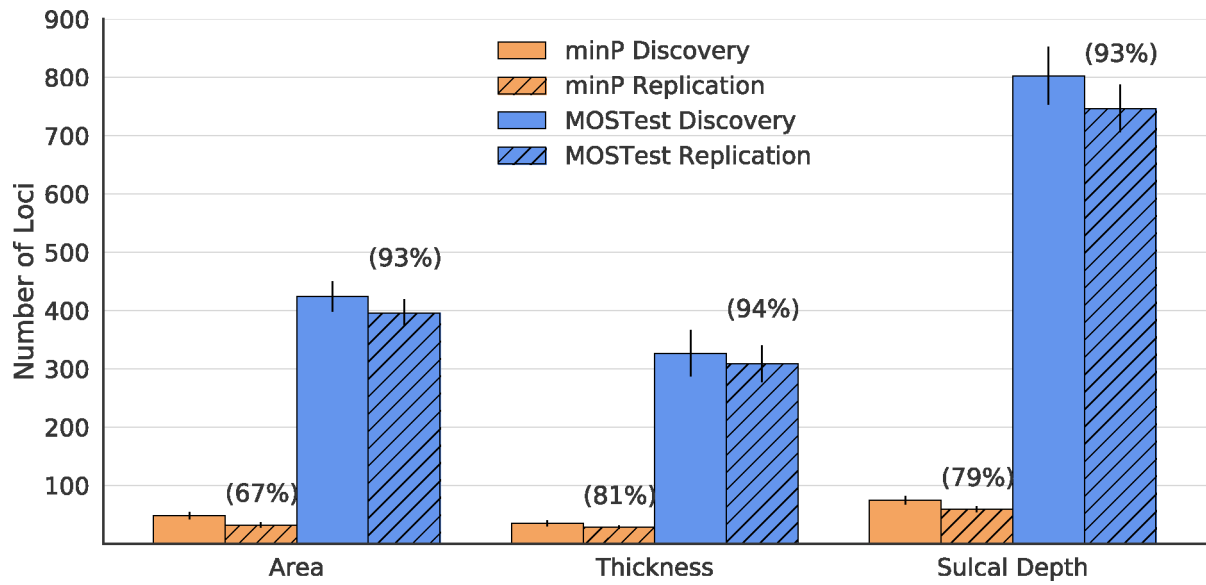
Figure 3.
Cross-validation discovery and replication yield within 10-times cross validation within UK Biobank for cortical morphometry measures. Solid bars represent the number of genome wide significant loci associated with each measure. Hashed bars represent the number of loci that replicate in test folds at a nominal significance level (p<0.05). Error bars are standard deviations across 10 cross-validation repetitions. Numbers in parentheses represent replication rate (# of discovered loci / # replicated loci) for each method-phenotype pair.
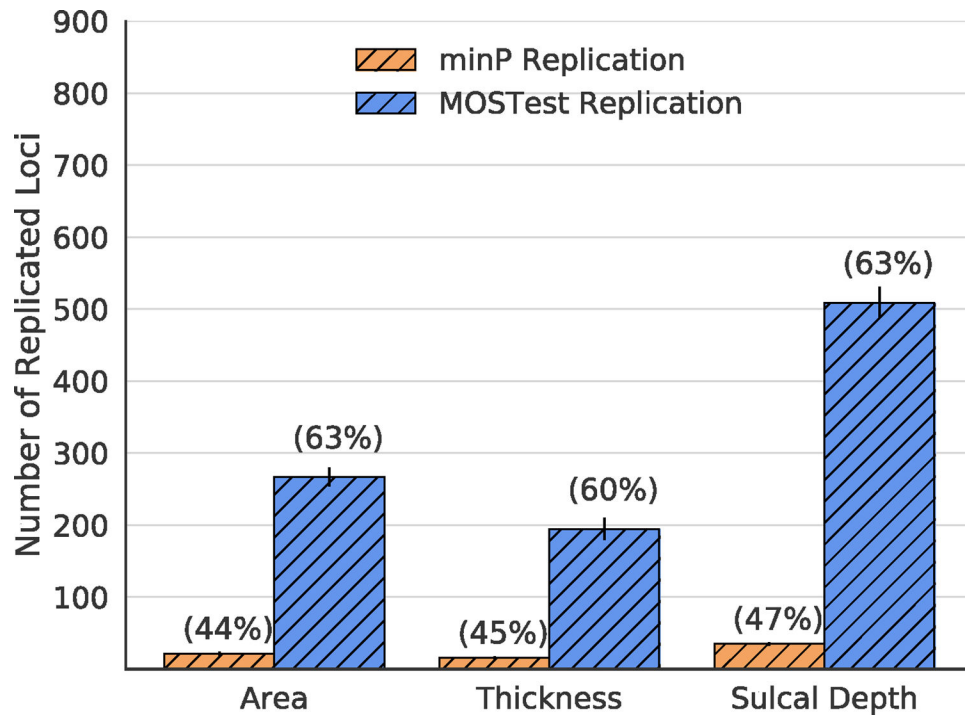
Figure 4.
Replication yield within the ABCD dataset across 10 training folds of UK Biobank for cortical morphometry measures. Bars represent the number of loci that replicate in ABCD at a nominal significance level (p<0.05). Error bars are standard deviations across 10 training sets of UK Biobank. Numbers in parentheses represent replication rate (# of discovered loci / # replicated loci). ABCD: Adolescent Brain Cognitive Development Study.