Application number/Title: 18177 - Multi-trait GWAS analyses in the UK Biobank

Applicant PI: Dr Paul O'Reilly

Application Institution: King's College London, MRC SGDP Centre, Psychiatry, Psychology and Neuroscience, DeCrespigny Park, Denmark Hill, London SE5 8AF, United Kingdom

Funding body: MRC

Keywords provided by the Applicant PI to describe the research project:

Multivariate, GWAS, pleiotropy, correlated, discovery, methods

Application Lay Summary:

1a: GWAS methods have been developed to analyse associations between SNPs and multiple phenotypes jointly. We have produced one such method (MultiPhen) and performed a simulation study finding that multivariate analyses can double the discovery of trait associated genetic variation compared with univariate analyses. Multi-trait analyses, such as polygenic risk scores, offer insights into shared and distinct aetiology among different phenotypes, such as ADHD, autism, schizophrenia, eating disorders and obesity. We will perform single and multi-trait analyses on the UK Biobank to boost discovery power of causal genetic variants, identify shared aetiology among phenotypes, and evaluate method performance on real data.

1b: Identifying shared genetic risk between physical and psychiatric phenotypes, in particular, could shed light on the aetiology of psychiatric disorders (which our group focuses on). Joint analysis of multiple traits may lead to substantially greater identification of novel genotype-phenotype associations and provide insights into the biological network underlying correlated phenotypes. Exposing the genetics responsible for comorbidity between phenotypes could also uncover possibilities for drug repositioning. By evaluating method performance we may find whether power is optimised by performing such analyses on clinically related phenotypes, those that are most highly correlated, or on sets of

phenotypes with heterogeneous correlation structure.

1c: Genome-wide association study (GWAS) analyses will be performed on subsets of the UK Biobank phenotypes, allowing some traits to be analysed jointly for the first time. We will use several statistical approaches to investigate patterns of shared and distinct genetic risk between and within psychiatric and physical traits. We will use 'polygenic risk scoring' and 'linkage disequilibrium score regression' analyses software to infer genetic overlap and estimate genetic correlations between different phenotypes and outcomes, using our local computing facilities.

1d: We wish to include the full UK Biobank cohort in our analyses.

In response to feedback from our approved preliminary application: we intend to use data from self-reported baseline information and health records to derive phenotypes. We acknowledge the complexity of looking across outcomes from different medical record sources and the difficulty in reliably ascertaining case status across multiple sources. We have undertaken preliminary analyses aimed at identifying case status in the UK Biobank using multivariate measures (self-report, hospital data, treatment status) as part of our collaboration on application 16577, and will extend this work with approval of this application.

Project extension: GWAS methods have been developed to analyse associations between SNPs and multiple phenotypes jointly. We have produced one such method (MultiPhen) and performed a simulation study finding that multivariate analyses can double the discovery of trait associated genetic variation compared with univariate analyses. Multi-trait analyses, such as polygenic risk scores, offer insights into shared and distinct aetiology among different phenotypes, such as ADHD, autism, schizophrenia, eating disorders and obesity. We will perform single and multi-trait analyses on the UK Biobank to boost discovery power of causal genetic variants, identify shared aetiology among phenotypes, and evaluate method performance on real data.

Our expertise in multivariate methodology will enable powerful investigations of the shared aetiology of psychiatric and physical traits, including brain-related phenotypes. The release of imaging data in the UKB provides an opportunity to further integrate neuroimaging into our psychiatric research. We request the T1-MRI data to enable deeper investigation of brain-related phenotypes and their involvement in psychiatric disorders.

For example, we will generate a novel neuroimaging phenotype - 'brain-age', representing an age-adjusted index of brain health. This biomarker has been used to explore trajectories of general health in ageing as well as in psychiatric disorders. It is heritable, and therefore informative to consider as a physical trait (with psychiatric relevance) alongside other physical phenotypes in our multi-trait GWAS analysis. Thus, the shared and unique genetic risks with poorer brain health and poorer physical health can be established.

The brain-age phenotype is a single value per individual, and amenable to polygenic score analysis, to determine what genetic variation is important for brain health, and how much variability in brain-age can be explained by composite genetic measures. Brain-age is calculated using custom image analysis software that takes T1-MRI data and generates a brain-age value. These values will be returned to the UKB.