bioRxiv preprint doi: https://doi.org/10.1101/048991; this version posted April 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Title: Analysis of shared heritability in common disorders of the brain

Authors: V Anttila*[1,1,2,3], B Bulik-Sullivan[1,3], H Finucane[4,5], J Bras[6], L Duncan[1,3], V Escott-Price[7], G Falcone[8], P Gormley[9], R Malik[10], N Patsopoulos[3,11], S Ripke[1,2,3,12], R Walters[1,2,3], Z Wei[13], D Yu[2,9], PH Lee[2,9], *IGAP consortium, IHGC consortium, ILAE Consortium on Complex Epilepsies, IMSGC consortium, IPDGC consortium, METASTROKE and Intracerebral Hemorrhage Studies of the International Stroke Genetics Consortium, Attention-Deficit Hyperactivity Disorder Working Group of the Psychiatric Genomics Consortium, Anorexia Nervosa Working Group of the Psychiatric Genomics Consortium, Autism Spectrum Disorders Working Group of The Psychiatric Genomics Consortium, Bipolar Disorders Working Group of the Psychiatric Genomics Consortium, Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium, Obsessive Compulsive Disorder and Tourette Syndrome Working Group of the Psychiatric Genomics Consortium, Schizophrenia Working Group of the Psychiatric Genomics Consortium*, G Breen[14,15], C Bulik[16,17], M Daly[1,2,3], M Dichgans[10,18], S Faraone[19], R Guerreiro[20], P Holmans[7], K Kendler[21], B Koeleman[22], CA Mathews[23], JM Scharf[2,3,8,9,24,25], P Sklar[26], J Williams[7], N Wood[20,27,28], C Cotsapas[3,29], A Palotie[1,2,3,9,30,31], JW Smoller[2,9], P Sullivan[16,32], J Rosand[3,8,25], A Corvin†*[2,33], BM Neale †*[3,1,2,3], on behalf of the Brainstorm consortium

Affiliations:

1) Analytic and Translational Genetics Unit, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts, USA

2) Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA

3) Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA

4) Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

5) Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA

6) Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK

7) Cardiff University, Medical Research Council Center for Neuropsychiatric Genetics & Genomics, Institute of Psychology, Medicine & Clinical Neuroscience, Cardiff, Wales, UK

8) Center for Human Genetic Research, Department of Neurology, Massachusetts General Hospital, Boston, MA, USA

9) Psychiatric and Neurodevelopmental Genetics Unit (PNGU), Center for Human Genetics Research, Massachusetts General Hospital, Boston, MA, USA

10) Institute for Stroke and Dementia Research, Klinikum der Universität München, Ludwig-Maximilians-University, Munich, Germany

11) Department of Neurology, Brigham & Women's Hospital, Harvard Medical School, Boston, MA

12) Charite Universitatsmedizin Berlin, Berlin, Germany

13) Department of Computer Science, New Jersey Institute of Technology, New Jersey, USA

14) Kings College London, Institute of Psychiatry, Psychology & Neuroscience, Social Genetics & Developmental Psychiatry Center, MRC, London, England

15) NIHR, Biomed Research Center for Mental Health, South London & Maudsley NHS Trust & Institute Psychiatry, London, England

16) Departments of Psychiatry and Nutrition, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA.

17) Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden.

18) Munich Cluster for Systems Neurology (SyNergy), Munich, Germany

19 Departments of Psychiatry and of Neuroscience and Physiology, SUNY Upstate Medical University, Syracuse, NY, USA

20) Department of Molecular Neuroscience, Institute of Neurology, University College London, London, UK

21) Department of Psychiatry, Virginia Commonwealth University, Richmond, VA, USA

22) Division Biomedical Genetics, University Medical Center Utrecht, Utrecht, the Netherlands

23) Department of Psychiatry and UF Genetics Institute, University of Florida: Gainesville, Florida, USA

24) Division of Cognitive and Behavioral Neurology, Brigham and Women's Hospital, Boston, MA, USA

25) Department of Neurology, Massachusetts General Hospital, Boston, MA, USA

26) Icahn School of Medicine at Mount Sinai, New York, New York, USA

27) Sobell Department of Motor Neuroscience and Movement Disorders, Institute of Neurology, University College London, London, UK

28) Institute of Genetics, University College London, London, UK

29) Department of Neurology, Yale School of Medicine, New Haven, CT, USA

30) Institute for Molecular Medicine Finland (FIMM), Helsinki, Finland

31) Department of Neurology, Massachusetts General Hospital, Boston, MA, USA.

bioRxiv preprint doi: https://doi.org/10.1101/048991; this version posted April 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

32) Department of Genetics and Psychiatry, University of North Carolina School of Medicine, Chapel Hill, North Carolina, USA
33) Neuropsychiatric Genetics Research Group, Department of Psychiatry, Trinity College Dublin, Ireland

*[1] anttila@broadinstitute.org
*[2] acorvin@tcd.ie
*[3] bneale@broadinstitute.org

One Sentence Summary: Comprehensive heritability analysis of brain phenotypes demonstrates a clear role for common genetic variation across neurological and psychiatric disorders, with substantial overlap within the latter.

Abstract: Disorders of the brain exhibit considerable epidemiological comorbidity and frequently share symptoms, provoking debate about the extent of their etiologic overlap. Here we apply linkage disequilibrium score regression (LDSC) to quantify the extent of shared genetic contributions across 23 brain disorders (n=842,820), 11 quantitative and four dichotomous traits of interest (n=722,125) based on genome-wide association meta-analyses. Psychiatric disorders show substantial sharing of common variant risk, while many neurological disorders appear more distinct from one another, suggesting substantive differences in the specificity of the genetic etiology of these disorders. Further, we observe little evidence of widespread sharing of the common genetic risk between neurological and psychiatric disorders studied. In addition, we identify significant sharing of genetic influences between the certain quantitative measures and brain disorders, including major depressive disorder and neuroticism personality score. These results highlight the importance of common genetic variation as a source of risk for brain disorders and the potential of using heritability methods to obtain a more comprehensive view of the genetic architecture of brain phenotypes.

Main Text:

The classification of brain disorders has evolved over the past century, reflecting the medical and scientific communities' best assessments of the presumed root causes of clinical phenomena such as behavioral change, loss of motor function, spontaneous movements or alterations of consciousness. A division between neurology and psychiatry developed as advances in histopathology and neuroimaging informed the biological understanding of the disorders, with the more directly observable phenomena (such as the presence of emboli, protein tangles, or unusual electrical activity patterns) generally defining the neurological disorders(*1*). A reassessment of long-held categorical distinctions between brain disorders may be helpful in informing the next steps in our search for the biological pathways that underlie their pathophysiology(*2, 3*).

Epidemiological and twin studies have explored these phenotypic overlaps generally (*4-6*), and substantial epidemiological comorbidity has been reported for many pairs of disorders, including bipolar disorder-migraine(*7*), stroke-major depressive disorder(*8*), epilepsy-autism spectrum disorders (ASD) and epilepsy-attention deficit hyperactivity disorder (ADHD)(*9*). Furthermore, genetic studies in neurological and psychiatric conditions have shown that mutations in the same ion channel genes confer pleiotropic risk for multiple distinct brain phenotypes. Examples include *de novo* coding mutations in *SCN2A* for epilepsy and ASD (*10*),

bioRxiv preprint doi: https://doi.org/10.1101/048991; this version posted April 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

coding variants in *CACNA1A* for migraine, epilepsy and episodic ataxia(*11, 12*) and recurrent microdeletions for autism, epilepsy and schizophrenia(*13*). Recently, genome-wide association studies (GWAS) have convincingly demonstrated that individual common risk variants show overlap across traditional diagnostic boundaries (e.g. migraine-stroke(*14*) and schizophrenia-bipolar disorder(*15*). Recent work in psychiatric genetics has also demonstrated strong genetic correlations across schizophrenia, major depressive disorder and bipolar disorder(*16*).

In general, brain disorders (excepting those caused by trauma, infection or cancer) show substantial heritability from twin and family studies (*17*). Subsequent genome-wide association studies (GWAS) have demonstrated that common genetic variation contributes to this heritability, in most cases implicating a large number of common variants with small risk effects. Notable recent successes have identified risk loci for schizophrenia(*18*), Alzheimer's disease(*19*), Parkinson's disease(*20*) and migraine(*21*). In addition to locus discovery, these larger sample sizes enable informative analyses of shared genetic influences, in order to improve our understanding of the degree of distinctiveness of brain disorders(*22*). Using recently developed heritability-based methods(*23*) we can now extend our evaluation of the nature of these diagnostic boundaries and explore the extent of shared common variant genetic influences across a wide set of neurological and psychiatric phenotypes.

Study design

The goal of this study was to systematically evaluate the evidence for sharing of risk alleles across a wide spectrum of brain disorders. We formed the Brainstorm consortium, a collaboration among GWAS meta-analysis consortia from 23 disorders (see Data sources), to analyze the most recent available summary statistics. In total, the study sample consists of 206,606 cases and 636,214 controls (Table 1; for sample overlap, see Fig. S1), as well as 722,125 additional samples for traits of interest (578,829 samples with quantitative phenotypes and 302,830 with dichotomous traits, with some overlap; Table 2). Data were centralized and underwent uniform quality control and processing. In order to avoid bias from population differences, we generated European-ancestry versions of those studies with a major proportion of non-European ancestry contributions to the meta-analysis (e.g. epilepsy, schizophrenia).

We have recently developed a novel heritability estimation method, linkage disequilibrium score regression (LDSC)(*24*), which allows for rapid estimation of heritability and genetic correlations from summary statistics. For a given trait, the total additive common SNP heritability in a set of GWAS summary statistics (h^2g) is estimated by regressing the association χ^2 statistic of a SNP on the total amount of genetic variation tagged by that SNP (i.e., the sum of r^2 between that SNP and the surrounding SNPs, termed the LD Score). We then extend this to estimate the genetic correlation, r_g, (i.e., the genome-wide average shared genetic risk) for a pair of phenotypes by regressing the product of Z-score for each phenotype for each

bioRxiv preprint doi: https://doi.org/10.1101/048991; this version posted April 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

SNP, instead of the χ^2 statistic. The LD Score is estimated from a common reference panel (for this work, the European subset of the 1000 Genomes Project reference) (*25*). In this framework, including LD in the regression allows us to distinguish and account for LD-independent error sources (such as sample sharing and population stratification) from LD-dependent sources (like polygenic signal)(*23*). It is essential to use an approach which is not biased by sample overlaps when analyzing summary statistics, given the large amount of control sharing between the GWAS meta-analyses in the study (Fig. S1).

Heritability analysis

GWAS summary statistics of the 23 disorders and 15 traits of interest underwent uniform quality control (Methods, Table S1 and S2), and were subsequently used to estimate their liability-scale heritability (Fig. S2A and S2B). Three ischemic stroke subtypes (cardioembolic, large-vessel disease and small-vessel disease) as well as the agreeableness personality measure from NEO Five-Factor inventory(*26*) had insufficiently powered heritability estimates for robust analysis and were excluded (see Methods). We performed a weighted-least squares regression analysis to evaluate whether differences relating to study makeup (effective sample size [Fig. S2C] or case/control ratio[Fig. S2D]) or disorder-specific measures (disorder prevalence [Fig. S2E] or age of onset [Fig. S2F]) were correlated with the magnitude of the estimates (Table S3). No measure was significantly associated with the liability-scale heritability.

The heritability estimates for the study disorders were generally somewhat lower than previously reported estimates from common variants (Table S4), with a few more pronounced cases, including ADHD, which showed considerably lower heritability in this study compared with the most recently reported estimate (*16*) (estimated h^2g of 0.10 [SE 0.01] and 0.28 [SE 0.02], respectively), Parkinson's disease(*27*) (estimated h^2g of 0.11 [SE 0.04] and 0.27 [SE 0.05], respectively) and Alzheimer's disease(*28*) (estimated h^2g of 0.13 [SE 0.01] and 0.24 [SE 0.03], respectively). These differences may reflect methodological variability due to heterogeneity in the case collections across the larger meta-analyses. Other potential causes include the impact of ancestry differences on the estimates, differences in phenotyping stringency (leading to a lower average polygenic load) or a relatively higher impact of rare variants (with a concomitant decrease in the proportion of disease due to common variants). For example, applying our analysis to the summary statistics of the GERAD cohort (3,941 cases and 7,848 controls) from the Alzheimer's disease meta-analysis, where the previous heritability estimate was calculated, we obtained an estimated h^2g of 0.25 [SE 0.04], which agrees closely with the published estimate of 0.24 [SE 0.03].

Heritability estimates for psychiatric phenotypes tended to be slightly higher than neurological disorders (with the exception of intracerebral hemorrhage and generalized epilepsy). These heritability estimates should be interpreted somewhat cautiously, as they may

bioRxiv preprint doi: https://doi.org/10.1101/048991; this version posted April 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

reflect potential biases present in the original case ascertainment, and will tend to be deflated by diagnostic heterogeneity and ascertainment errors or unusual contributions of high-impact rare variants. However, in LDSC the magnitude of the heritability estimate does not bias the estimates of genetic correlation.

We observed a similar range of heritability estimates among the traits of interest as among the disorders (Table S5). Three of the "Big Five" personality measures (agreeableness, neuroticism and extraversion) yielded very low heritability (estimated h^2g for all < 0.04), while conscientiousness (estimated h^2g 0.08 [SE 0.03]) and openness to experience (estimated h^2g 0.12 [SE 0.03]) showed more robust estimates, though similar to the disorder phenotypes, somewhat lower than previously estimated in smaller sample sizes(*29*). Measures related to cognitive ability, such as childhood cognitive performance (estimated h^2g 0.19, [SE 0.03]) and college attainment (estimated h^2g 0.15 [SE 0.01]), yielded estimates that were more consistent with previous estimates of the heritability of intelligence(*30, 31*), suggesting that the cognitive measures may be less prone to phenotypic measurement error and/or have a higher heritability overall than the personality measures.

We also performed a functional partitioning analysis using stratified LD score regression to examine whether the observed heritability was enriched in any tissue-specific regulatory partitions of the genome, using the ten top-level partitions from Finucane et al(*32*) (Figure S3). In the functional enrichment analysis, we replicated the previously reported significant central nervous system (CNS) enrichment for schizophrenia and bipolar disorder(*32*), and demonstrated novel heritability enrichments for generalized epilepsy and major depressive disorder (to CNS) and multiple sclerosis (to immune system cells and tissues). Phenotype-tissue pairs with an FDR-corrected q-value < 0.05 for enrichment are listed in Table S5.

Correlations among brain disorders

Several distinct patterns emerged in the shared heritability analysis. In expanding on the number of psychiatric disorders analyzed from previous work(*16*), we observed broad sharing across psychiatric disorders (Figure 1; average r_g=0.21). Notably, schizophrenia showed significant genetic correlation with most of the other studied psychiatric disorders. Schizophrenia, bipolar disorder, major depressive disorder and ADHD each showed a high degree of correlation to the others (average r_g = 0.41). Anorexia nervosa, OCD and schizophrenia also demonstrated significant sharing among each other. Tourette Syndrome was only significantly correlated with OCD and migraine and not with any of the other disorders. From this analysis, the common variant genetic architectures of ASD and Tourette Syndrome appear to be somewhat distinct from other psychiatric disorders (with the exception of the OCD-Tourette Syndrome pair) in that they do not show significant correlations with other phenotypes, at least at current sample sizes.

bioRxiv preprint doi: https://doi.org/10.1101/048991; this version posted April 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

In contrast to psychiatric disorders, the neurological disorders included revealed greater specificity, and the extent of genetic correlation was generally more limited (Figure 2; after excluding subtype-specific analyses, average r_g = 0.06). Several disorders — Parkinson's disease, Alzheimer's disease, generalized epilepsy and multiple sclerosis — showed little or no overlap with any other brain phenotypes. Focal epilepsy showed the highest degree of genetic correlation among the neurological disorders, but none were significant, possibly reflecting the relatively modest power of the current focal epilepsy meta-analysis. However, the modest heritability and the broad pattern of sharing may be consistent with low levels of heterogeneity and potentially misclassification across a wide range of neurological conditions. Disorder-specific summaries of the results are shown in Figure S4 and Table S6.

Cross-category analysis and traits of interest

In the cross-category correlation analysis, the overall pattern is consistent with limited sharing across neurological and psychiatric traits (Figure 3; average r_g=0.03). The only significant cross-category correlations were migraine-ADHD, migraine-Tourette Syndrome and migraine-major depressive disorder and these were relatively modest (average r_g=0.22). The strongest positive correlations were all within-category (highest was focal epilepsy and intracranial hemorrhage; r_g = 0.83, p=0.03) while the negative correlations were all across category (lowest was focal epilepsy and OCD; r_g = -0.42, p=0.04).

We observed a number of significant genetic correlations between the traits of interest and the brain disorders. Among anthropometric traits, BMI was the only trait with significant correlations, with a positive correlation with ADHD and negative correlations with anorexia nervosa (previously reported (*23*)), OCD and schizophrenia. In contrast, Crohn's disease was included as a proxy for immunological effects, but no correlations with any of the study phenotypes were observed. However, the phenotype chosen to represent vascular pathophysiology, coronary artery disease, did show significant correlation to stroke-related phenotypes (r_g = 0.69, p=2.47 x 10^{-6} to ischemic stroke and r_g = 0.86, p=2.26 x 10^{-5} for early-onset stroke), suggesting a clear role for shared genetic effects across these phenotypes.

In measures related to cognitive function, three psychiatric phenotypes (anorexia nervosa, ASD and bipolar disorder) showed significant positive correlation to these measures, while in contrast two neurological phenotypes (Alzheimer's disease and intracerebral hemorrhage) showed a negative correlation. A more detailed comparison of all phenotypes with a correlation Z-score>3 in this category are shown in Fig. S6. Among the personality measures, the only significant result was for the major depressive disorder-neuroticism pair, which showed a significantly positive genetic correlation (r_g = 0.83, p= 3.60 x 10^{-9}). For smoking-related measures, the only significant genetic correlation was with ADHD-never/ever smoked (r_g=0.38,

bioRxiv preprint doi: https://doi.org/10.1101/048991; this version posted April 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

$p=3.17 \times 10^{-7}$), potentially supporting findings from twin studies(*33*) that suggest impulsivity contributes to smoking initiation.

Discussion

By integrating and analyzing the current genome-wide association summary statistic data from consortia of 23 brain disorders, we find that psychiatric disorders broadly share a considerable portion of their common variant genetic risk, especially across schizophrenia, major depressive disorder, bipolar disorder and ADHD, while neurological disorders generally do not. Across categories, psychiatric and neurologic disorders share relatively little of their common genetic risk, suggesting that multiple different and largely independently regulated etiological paths may give rise to similar clinical manifestations (e.g., psychosis, which manifests in both schizophrenia(*34*) and Alzheimer's disease(*35*)). The clinical delineation between neurology and psychiatry is thus recapitulated at the level of common variant risk for the studied disorders.

The high degree of genetic correlation among the psychiatric traits suggests that genetic risk factors for psychiatric disorders do not respect clinical diagnostic boundaries, congruent with the clinical controversies in classification. The broad and continuous nature of psychiatric disorder spectra have been clinically recognized for a long time(*36-38*), and these results suggest that shared biological mechanisms substantially contribute across psychiatric diagnoses. The observed positive genetic correlations are consistent with a few different scenarios. For example, the observed r_g may reflect the existence of some considerable portion of common genetic risk factors conferring equal risks to multiple disorders and that other additional factors which contribute to the eventual clinical representation are mostly distinct. Alternatively, all common genetic effects could be shared between a pair of traits, but each individual effect may confer different degrees of risk and lead to different aggregate genetic risk profiles. It is likely that the truth lies somewhere in between, and it will become increasingly feasible to evaluate these overlaps at the locus level as more genome-wide significant loci are identified.

The presence of significant genetic correlation may also reflect the phenotypic overlap between any two disorders. For example, the sharing between schizophrenia and ADHD might reflect underlying difficulties in executive functioning, which are well-established in both disorders(*39*). Similarly, the sharing between anorexia nervosa, OCD and schizophrenia may reflect a shared cognitive effect spectrum from overvalued ideas to psychotic thinking. Another possibility is that a heritable intermediate trait confers risk to both outcomes, thereby giving rise to the genetic correlation, as the genetic influences on this trait will be shared for both outcomes (e.g., obesity as a risk factor for both type 2 diabetes and myocardial infarction). Alternatively, misclassification of cases (or comorbidity) across each of the studies may also introduce genetic correlation, though previous work suggests that even substantial misclassification across

bioRxiv preprint doi: https://doi.org/10.1101/048991; this version posted April 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

schizophrenia and bipolar disorder is insufficient to explain the observed genetic correlation(*40*), though it will likely decrease the power to estimate the extent of genetic overlap.

The low correlations across neurological disorders, suggest that the current classification reflects specific genetic etiology, although the limited sample size for some of these disorders and lack of inclusion of disorders conceived as "circuit-based", such as restless legs syndrome, sleep disorders and possibly essential tremor, constrains the generalizability of this conclusion. This recapitulates the current understanding that the disorder etiologies are largely non-overlapping; degenerative disorders (such as Alzheimer's and Parkinson's diseases) would not be expected *a priori* to share their polygenic risk profiles with a neuro-immunological disorder (like multiple sclerosis) or neurovascular disorder (like ischemic stroke). Considering neurological disorder subtypes, migraine with and without aura (r_g = 0.45, p=3.86 x 10^{-5}) show substantial genetic correlation; whereas focal and generalized epilepsy (r_g = 0.15, p= 0.407) show limited genetic correlation.

We did observe a few significant correlations across neurology and psychiatry, namely between migraine and ADHD, major depressive disorder and Tourette Syndrome, suggesting modest shared etiological overlap across the neurology/psychiatry distinction. The epidemiological co-morbidity of migraine with major depressive disorder and ADHD has been previously reported in epidemiological studies (*41-43*), while in contrast, the previously reported epidemiological co-morbidity between migraine and bipolar disorder (*44*) is not reflected at all in our estimate of genetic correlation (r_g = -0.02, p= 0.414). Similarly, we see limited evidence for the epidemiological co-morbidity between migraine with aura and ischemic stroke(*45*) (r_g = 0.08, p=0.299); however, the standard errors of this comparison are too high to draw strong conclusions. Further, as the datasets are cross-sectional instead of prospective, survival effects may contribute to the differences.

Several phenotypes show only very low-level correlations with any of the other studied disorders and traits, despite large sample size and high heritability, suggesting their common variant genetic risk may largely be unique. In particular, Alzheimer's disease, Parkinson's disease and multiple sclerosis show extremely limited sharing with the other phenotypes and with each other. Neuroinflammation has been implicated in the pathophysiology of each of these conditions(*46-48*), as it has for migraine(*49*) and schizophrenia(*50*), but no considerable shared heritability was observed with either of them nor with Crohn's disease, which was included among the traits of interest due to its inflammatory component. While this observation does not preclude shared neuroinflammatory mechanisms in these disorders, it does suggest that on a large scale, common variant genetic influences on these inflammatory mechanisms are not shared between the disorders, but are instead likely to be disorder-specific. Further, we only observed enrichment of heritability to immunological cells and tissues in multiple sclerosis, showing that inflammation-specific regulatory marks in the genome do not show overall enrichment for common variant risk for Alzheimer's and Parkinson's diseases (though this does not preclude the effects of specific, non-polygenic neuroinflammatory mechanisms(*51*)). Among

bioRxiv preprint doi: https://doi.org/10.1101/048991; this version posted April 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

psychiatric disorders, ASD and Tourette Syndrome showed a similar absence of correlation with other disorders, although this could reflect small sample sizes.

Traits of interest

We observed a number of significant correlations between brain disorders and the traits of interest. First, the observed correlation between ADHD and smoking initiation (r_g = 0.38, p=1.09 x 10^{-6}) is consistent with the epidemiological evidence of overlap(*52*), and with the existing hypothesis that impulsivity inherent in ADHD may drive smoking initiation and potentially dependence.

BMI shows significant positive genetic correlation to ADHD, consistent with a meta-analysis linking ADHD to obesity(*53*), and negative genetic correlation with anorexia nervosa, OCD and schizophrenia. These results connect well with the evidence for enrichment of BMI heritability in CNS tissues(*32*) and that many reported signals suggest neuronal involvement(*54*); this may also provide a partial genetic explanation for lower BMI in anorexia nervosa patients after recovery(*55*). Given that no strong correlations were observed between BMI and any of the neurological phenotypes, it is possible to hypothesize that at least part of BMI's brain-specific genetic architecture may be more closely related to behavioral phenotypes. However, ischemic stroke and BMI show surprisingly little genetic correlation in this analysis (r_g = 0.06, p=0.296), suggesting that although BMI is a strong risk factor for stroke(*56*), there is limited evidence for shared common genetic effects. These analyses also provide an opportunity to explore whether the genetic influences on the reported reduced rate of cardiovascular disease in individuals with anorexia nervosa are due to BMI-related effects; with the limited evidence of overlap of anorexia nervosa with intracerebral hemorrhage, ischemic stroke, early-onset stroke and myocardial infarction (*57, 58*), these results suggest that any lower cardiovascular mortality is more likely due to direct BMI-related effects rather than any shared common genetic risk variants.

Analysis of the Big Five personality measures suggests that the current sample sizes for personality data are now starting to be sufficiently large for correlation testing. The significant positive genetic correlation between major depressive disorder and neuroticism (r_g = 0.83, p= 3.60 x 10^{-9}) provides further evidence for the link between these phenotypes, which has been reported previously with polygenic risk scores(*53*). Further, there are considerable overlap with several other (primarily psychiatric) disorders; among the positive, ADHD-extraversion (r_g = 0.34, p=0.001), migraine without aura-neuroticism(r_g = 0.36, p=0.003) and anorexia nervosa-neuroticism (r_g = 0.40, p=0.005), and among the negative, major depressive disorder-conscientiousness (r_g = -0.42, p=0.02) and ADHD-conscientiousness (r_g = -0.47, p= 0.02).

bioRxiv preprint doi: https://doi.org/10.1101/048991; this version posted April 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

The observed pattern of positive genetic correlation to cognitive measures (which are all from early life; the childhood cognitive performance is measured between ages 6-18, which roughly covers the range thought to be contributing to educational attainment as well) in several psychiatric phenotypes concurrent with negative genetic correlation to neurological phenotypes is potentially interesting for follow-up studies (Fig. S6). While ADHD was the only psychiatric trait to show a significant negative correlation to cognitive measures, three (anorexia nervosa, ASD and bipolar disorder) of the eight phenotypes showed significant positive genetic correlation with one or more cognitive measures and a further two (OCD and schizophrenia) show moderate positive genetic correlation (Z-score > 2). For the neurological phenotypes, two (Alzheimer's disease and intracerebral hemorrhage) showed significant negative genetic correlation to the cognitive measures, while a further four (epilepsy, ischemic stroke, early-onset stroke and migraine) showed moderate negative genetic correlation (Z-score < -2) and one (Parkinson's disease) moderate positive correlation (Z-score >2). For Alzheimer's disease, poor cognitive performance in youth has been linked with increased risk for developing the disorder in later life(*59*), but to our knowledge no such connection has been reported for the other phenotypes. These results suggest the existence of a link between genetic risk for several brain disorders and cognitive performance in early life.

Heritability methods

The application of heritability-based methodology promises to be a good alternative approach for estimating shared genetic correlation between phenotypes with differing prevalence, age of onset and diagnostic challenges. Being able to estimate the genetic correlational structure directly from the summary statistics of an entire meta-analysis for a condition, as opposed to a subset of study participants who have been measured for multiple phenotypes, is a major benefit for using LD score. Accordingly, sample collection and phenotyping of each disorder can occur independently and following to the appropriate disorder-specific gold standard, rather than needing to capture each phenotype (potentially less comprehensively) in every cohort. Heritability-based methods, such as the one used here, may further be a useful tool in the future for evaluating the validity of those diagnostic standards in capturing maximally informative patient samples for genetic studies, both for the formal diagnosis itself and in relation to their overlap with secondary and tertiary phenotypes.

Although fairly large sample sizes are required for obtaining robust heritability estimates, we did not observe systematic biases resulting from different case/control ratios, disorder prevalence or sample size. Selection and survival biases in the underlying data, though reflecting the best efforts of the GWAS community and participating consortia, may attenuate the heritability estimates and correlations, as may within-disorder heterogeneity in the larger meta-analyses. Even given the advantages of our approach, some of the individual studies are still modestly powered for robust estimation of genetic correlations. Moreover, our analyses only

bioRxiv preprint doi: https://doi.org/10.1101/048991; this version posted April 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

examine the properties of common variant contributions, and extending these analyses to rare variants may better inform the extent of the overlaps. The restriction to common variants could explain the limited overlap observed in this study between schizophrenia and ASD, which has been previously reported to be present for rare, novel loss-of-function variants(*60*). This may suggest that the rare and common variant contributions to genetic correlation may behave differently, and incorporating the two variant classes into a single analysis remains a challenge. In particular, more targeted studies are required to ascertain the pathway- and other biological implications of the results.

The broader implication of our findings is that the current clinical boundaries for the studied psychiatric phenotypes do not reflect distinct underlying pathogenic processes based on the genetic evidence. In contrast, the included neurological disorders show considerably greater genetic specificity. Although it is important to emphasize that while some disorders are under-presented here (e.g anxiety and personality disorders in psychiatry and circuit-based disorders such as restless leg syndrome in neurology), these results clearly demonstrate limited evidence for widespread common genetic risk sharing between psychiatry and neurology. Genetically informed analyses may provide the basis for some degree of restructuring of psychiatric nosology (consistent with the historical impact of twin- and family-based results). Further elucidation of the genetic overlap, especially as distinct loci map onto a subset of disorders, may form the basis for either defining new clinical phenotypes or support a move to a more continuous view of psychiatric phenotypes. Ultimately, such developments give hope for reducing diagnostic heterogeneity and eventually improving the diagnosis and treatment of psychiatric disorders.

bioRxiv preprint doi: https://doi.org/10.1101/048991; this version posted April 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Author Information Correspondence and requests for materials should be addressed to V.A. (anttila@broadinstitute.org), A.C. (acorvin@tcd.ie) or B.M.N. (bneale@broadinstitute.org).

Data sources

Disorder or trait – Consortium or dataset identifier – web address:

Psychiatric disorders
ADHD – PGC-ADD2 - www.med.unc.edu/pgc
Anorexia nervosa – PGC-AN - www.med.unc.edu/pgc
Autism spectrum disorders(*16*) – PGC-AUT - www.med.unc.edu/pgc
Bipolar disorder – PGC-BIP2 - www.med.unc.edu/pgc
Major depressive disorder – PGC-MDD2 - www.med.unc.edu/pgc
OCD – IOCDFGC - https://iocdf.org/
Schizophrenia(*18*) – PGC-SCZ2 - www.med.unc.edu/pgc
Tourette Syndrome – TSAIGC – N/A

Neurological disorders
Alzheimer's disease(*19*) – IGAP - http://www.pasteur-lille.fr/en/recherche/u744/igap
Epilepsy and subtypes, focal and generalized(*61*) – ILAE – http://www.epigad.org/page/show/gwas_index
Intracerebral hemorrhage(*62*) – ISGC - http://www.strokegenetics.com/
Ischemic stroke and subtypes (cardioembolic, early-onset, small-vessel and large-vessel)(*63*) – METASTROKE dataset of the ISGC – http://www.strokegenetics.com/
Migraine and subtypes, migraine with and without aura – IHGC – www.headachegenetics.org
Multiple sclerosis(*64*) – IMSGC - http://eaglep.case.edu/imsgc_web
Parkinson's disease(*20*) – IPDGC – www.pdgene.org

Traits of interest
BMI(*54*) – GIANT – https://www.broadinstitute.org/collaboration/giant
Height(*65*) – GIANT – https://www.broadinstitute.org/collaboration/giant
Infant head circumference at birth(*66*) – EGG – http://egg-consortium.org/
College attainment, years of education(*67*) – SSGAC – http://ssgac.org/Data.php
Childhood cognitive performance(*68*) – SSGAC – http://ssgac.org/Data.php
Neuroticism, extraversion, agreeableness, conscientiousness and openness (*26*) – GPC – http://www.tweelingenregister.org/GPC/
Crohn's disease(*69*) – IIBDGC - http://www.ibdgenetics.org/downloads.html
Coronary artery disease(*70*) – Cardiogram – http://www.cardiogramplusc4d.org/downloads/
Never/ever smoked, cigarettes per day(*71*) - TAG - https://www.med.unc.edu/pgc/downloads

Other
matSpD - neurogenetics.qimrberghofer.edu.au/matSpD/
Circos – circos.ca

bioRxiv preprint doi: https://doi.org/10.1101/048991; this version posted April 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

References:

1. J. B. Martin, The integration of neurology, psychiatry, and neuroscience in the 21st century. *Am J Psychiatry* **159**, 695 (May, 2002).
2. J. W. Smoller, Disorders and borders: psychiatric genetics and nosology. *Am J Med Genet B Neuropsychiatr Genet* **162B**, 559 (Oct, 2013).
3. T. R. Insel, P. S. Wang, Rethinking mental illness. *JAMA* **303**, 1970 (May 19, 2010).
4. K. S. Kendler, C. A. Prescott, J. Myers, M. C. Neale, The structure of genetic and environmental risk factors for common psychiatric and substance use disorders in men and women. *Arch Gen Psychiatry* **60**, 929 (Sep, 2003).
5. R. Jensen, L. J. Stovner, Epidemiology and comorbidity of headache. *Lancet neurology* **7**, 354 (Apr, 2008).
6. J. Nuyen *et al.*, Comorbidity was associated with neurologic and psychiatric diseases: a general practice-based controlled study. *J Clin Epidemiol* **59**, 1274 (Dec, 2006).
7. R. M. Hirschfeld *et al.*, Screening for bipolar disorder in the community. *The Journal of clinical psychiatry* **64**, 53 (Jan, 2003).
8. A. Pan, Q. Sun, O. I. Okereke, K. M. Rexrode, F. B. Hu, Depression and risk of stroke morbidity and mortality: a meta-analysis and systematic review. *JAMA* **306**, 1241 (Sep 21, 2011).
9. A. Lo-Castro, P. Curatolo, Epilepsy associated with autism and attention deficit hyperactivity disorder: is there a genetic link? *Brain & development* **36**, 185 (Mar, 2014).
10. T. D. Graves, M. G. Hanna, Neurological channelopathies. *Postgraduate medical journal* **81**, 20 (Jan, 2005).
11. J. Haan, G. M. Terwindt, A. M. van den Maagdenberg, A. H. Stam, M. D. Ferrari, A review of the genetic relation between migraine and epilepsy. *Cephalalgia* **28**, 105 (Feb, 2008).
12. B. de Vries, R. R. Frants, M. D. Ferrari, A. M. van den Maagdenberg, Molecular genetics of migraine. *Hum Genet* **126**, 115 (Jul, 2009).
13. C. G. de Kovel *et al.*, Recurrent microdeletions at 15q11.2 and 16p13.11 predispose to idiopathic generalized epilepsies. *Brain* **133**, 23 (Jan, 2010).
14. S. Debette *et al.*, Common variation in PHACTR1 is associated with susceptibility to cervical artery dissection. *Nat Genet* **47**, 78 (Jan, 2015).
15. S. M. Purcell *et al.*, Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460**, 748 (Aug 6, 2009).
16. C. Cross-Disorder Group of the Psychiatric Genomics *et al.*, Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet* **45**, 984 (Sep, 2013).
17. T. J. Polderman *et al.*, Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat Genet* **47**, 702 (Jul, 2015).
18. C. Schizophrenia Working Group of the Psychiatric Genomics, Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421 (Jul 24, 2014).
19. J. C. Lambert *et al.*, Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet* **45**, 1452 (Dec, 2013).
20. M. A. Nalls *et al.*, Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. *Nat Genet* **46**, 989 (Sep, 2014).
21. V. Anttila *et al.*, Genome-wide meta-analysis identifies new susceptibility loci for migraine. *Nat Genet* **45**, 912 (Aug, 2013).
22. N. Solovieff, C. Cotsapas, P. H. Lee, S. M. Purcell, J. W. Smoller, Pleiotropy in complex traits: challenges and strategies. *Nat Rev Genet* **14**, 483 (Jul, 2013).
23. B. Bulik-Sullivan *et al.*, An atlas of genetic correlations across human diseases and traits. *Nat Genet* **47**, 1236 (Nov, 2015).

bioRxiv preprint doi: https://doi.org/10.1101/048991; this version posted April 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

24. B. K. Bulik-Sullivan *et al.*, LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* **47**, 291 (Mar, 2015).
25. C. Genomes Project *et al.*, A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061 (Oct 28, 2010).
26. M. H. de Moor *et al.*, Meta-analysis of genome-wide association studies for personality. *Mol Psychiatry* **17**, 337 (Mar, 2012).
27. M. F. Keller *et al.*, Using genome-wide complex trait analysis to quantify 'missing heritability' in Parkinson's disease. *Hum Mol Genet* **21**, 4996 (Nov 15, 2012).
28. S. H. Lee *et al.*, Estimation and partitioning of polygenic variation captured by common SNPs for Alzheimer's disease, multiple sclerosis and endometriosis. *Hum Mol Genet* **22**, 832 (Feb 15, 2013).
29. R. A. Power, M. Pluess, Heritability estimates of the Big Five personality traits based on common genetic variants. *Translational psychiatry* **5**, e604 (2015).
30. C. M. Haworth *et al.*, The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Mol Psychiatry* **15**, 1112 (Nov, 2010).
31. I. J. Deary *et al.*, Genetic contributions to stability and change in intelligence from childhood to old age. *Nature* **482**, 212 (Feb 9, 2012).
32. H. K. Finucane *et al.*, Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat Genet* **47**, 1228 (Nov, 2015).
33. T. Korhonen *et al.*, Externalizing behaviors and cigarette smoking as predictors for use of illicit drugs: a longitudinal study among Finnish adolescent twins. *Twin Res Hum Genet* **13**, 550 (Dec, 2010).
34. P. F. Buckley, B. J. Miller, D. S. Lehrer, D. J. Castle, Psychiatric comorbidities and schizophrenia. *Schizophrenia bulletin* **35**, 383 (Mar, 2009).
35. C. G. Lyketsos *et al.*, Mental and behavioral disturbances in dementia: findings from the Cache County Study on Memory in Aging. *Am J Psychiatry* **157**, 708 (May, 2000).
36. R. Kendell, A. Jablensky, Distinguishing between the validity and utility of psychiatric diagnoses. *Am J Psychiatry* **160**, 4 (Jan, 2003).
37. A. S. Cristino *et al.*, Neurodevelopmental and neuropsychiatric disorders represent an interconnected molecular system. *Mol Psychiatry* **19**, 294 (Mar, 2014).
38. D. A. Regier *et al.*, Limitations of diagnostic criteria and assessment instruments for mental disorders. Implications for research and policy. *Arch Gen Psychiatry* **55**, 109 (Feb, 1998).
39. E. G. Willcutt, A. E. Doyle, J. T. Nigg, S. V. Faraone, B. F. Pennington, Validity of the executive function theory of attention-deficit/hyperactivity disorder: a meta-analytic review. *Biol Psychiatry* **57**, 1336 (Jun 1, 2005).
40. N. R. Wray, S. H. Lee, K. S. Kendler, Impact of diagnostic misclassification on estimation of genetic correlations using genome-wide genotypes. *Eur J Hum Genet* **20**, 668 (Jun, 2012).
41. O. B. Fasmer, A. Halmoy, K. J. Oedegaard, J. Haavik, Adult attention deficit hyperactivity disorder is associated with migraine headaches. *European archives of psychiatry and clinical neuroscience* **261**, 595 (Dec, 2011).
42. N. Breslau, R. B. Lipton, W. F. Stewart, L. R. Schultz, K. M. Welch, Comorbidity of migraine and depression: investigating potential etiology and prognosis. *Neurology* **60**, 1308 (Apr 22, 2003).
43. K. R. Merikangas, J. Angst, H. Isler, Migraine and psychopathology. Results of the Zurich cohort study of young adults. *Arch Gen Psychiatry* **47**, 849 (1990).
44. R. S. McIntyre *et al.*, The prevalence and impact of migraine headache in bipolar disorder: results from the Canadian Community Health Survey. *Headache* **46**, 973 (Jun, 2006).
45. J. T. Spector *et al.*, Migraine headache and ischemic stroke risk: an updated meta-analysis. *The American journal of medicine* **123**, 612 (Jul, 2010).

bioRxiv preprint doi: https://doi.org/10.1101/048991; this version posted April 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

46. M. T. Heneka *et al.*, Neuroinflammation in Alzheimer's disease. *Lancet neurology* **14**, 388 (Apr, 2015).
47. E. C. Hirsch, S. Hunot, Neuroinflammation in Parkinson's disease: a target for neuroprotection? *Lancet neurology* **8**, 382 (Apr, 2009).
48. E. M. Frohman, M. K. Racke, C. S. Raine, Multiple sclerosis--the plaque and its pathogenesis. *N Engl J Med* **354**, 942 (Mar 2, 2006).
49. C. Waeber, M. A. Moskowitz, Migraine as an inflammatory disorder. *Neurology* **64**, S9 (May 24, 2005).
50. J. Steiner *et al.*, Increased prevalence of diverse N-methyl-D-aspartate glutamate receptor antibodies in patients with an initial diagnosis of schizophrenia: specific relevance of IgG NR1a antibodies for distinction from N-methyl-D-aspartate glutamate receptor encephalitis. *JAMA psychiatry* **70**, 271 (Mar, 2013).
51. C. International Genomics of Alzheimer's Disease, Convergent genetic and expression data implicate immunity in Alzheimer's disease. *Alzheimer's & dementia : the journal of the Alzheimer's Association* **11**, 658 (Jun, 2015).
52. F. J. McClernon, S. H. Kollins, ADHD and smoking: from genes to brain to behavior. *Ann N Y Acad Sci* **1141**, 131 (Oct, 2008).
53. S. Cortese *et al.*, Association Between ADHD and Obesity: A Systematic Review and Meta-Analysis. *Am J Psychiatry* **173**, 34 (Jan 1, 2016).
54. D. Shungin *et al.*, New genetic loci link adipose and insulin biology to body fat distribution. *Nature* **518**, 187 (Feb 12, 2015).
55. L. Mustelin *et al.*, Long-term outcome in anorexia nervosa in the community. *The International journal of eating disorders* **48**, 851 (Nov, 2015).
56. T. Kurth *et al.*, Prospective study of body mass index and risk of stroke in apparently healthy women. *Circulation* **111**, 1992 (Apr 19, 2005).
57. S. R. Korndorfer *et al.*, Long-term survival of patients with anorexia nervosa: a population-based study in Rochester, Minn. *Mayo Clinic proceedings* **78**, 278 (Mar, 2003).
58. P. F. Sullivan, Discrepant results regarding long-term survival of patients with anorexia nervosa? *Mayo Clinic proceedings* **78**, 273 (Mar, 2003).
59. D. A. Snowdon *et al.*, Linguistic ability in early life and cognitive function and Alzheimer's disease in late life. Findings from the Nun Study. *JAMA* **275**, 528 (Feb 21, 1996).
60. E. M. Kenny *et al.*, Excess of rare novel loss-of-function variants in synaptic genes in schizophrenia and autism spectrum disorders. *Mol Psychiatry* **19**, 872 (Aug, 2014).
61. e.-a. u. e. a. International League Against Epilepsy Consortium on Complex Epilepsies. Electronic address, Genetic determinants of common epilepsies: a meta-analysis of genome-wide association studies. *Lancet neurology* **13**, 893 (Sep, 2014).
62. D. Woo *et al.*, Meta-analysis of genome-wide association studies identifies 1q22 as a susceptibility locus for intracerebral hemorrhage. *Am J Hum Genet* **94**, 511 (Apr 3, 2014).
63. M. Traylor *et al.*, Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE collaboration): a meta-analysis of genome-wide association studies. *Lancet neurology* **11**, 951 (Nov, 2012).
64. N. A. Patsopoulos *et al.*, Genome-wide meta-analysis identifies novel multiple sclerosis susceptibility loci. *Ann Neurol* **70**, 897 (Dec, 2011).
65. A. R. Wood *et al.*, Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet* **46**, 1173 (Nov, 2014).
66. H. R. Taal *et al.*, Common variants at 12q15 and 12q24 are associated with infant head circumference. *Nat Genet* **44**, 532 (May, 2012).

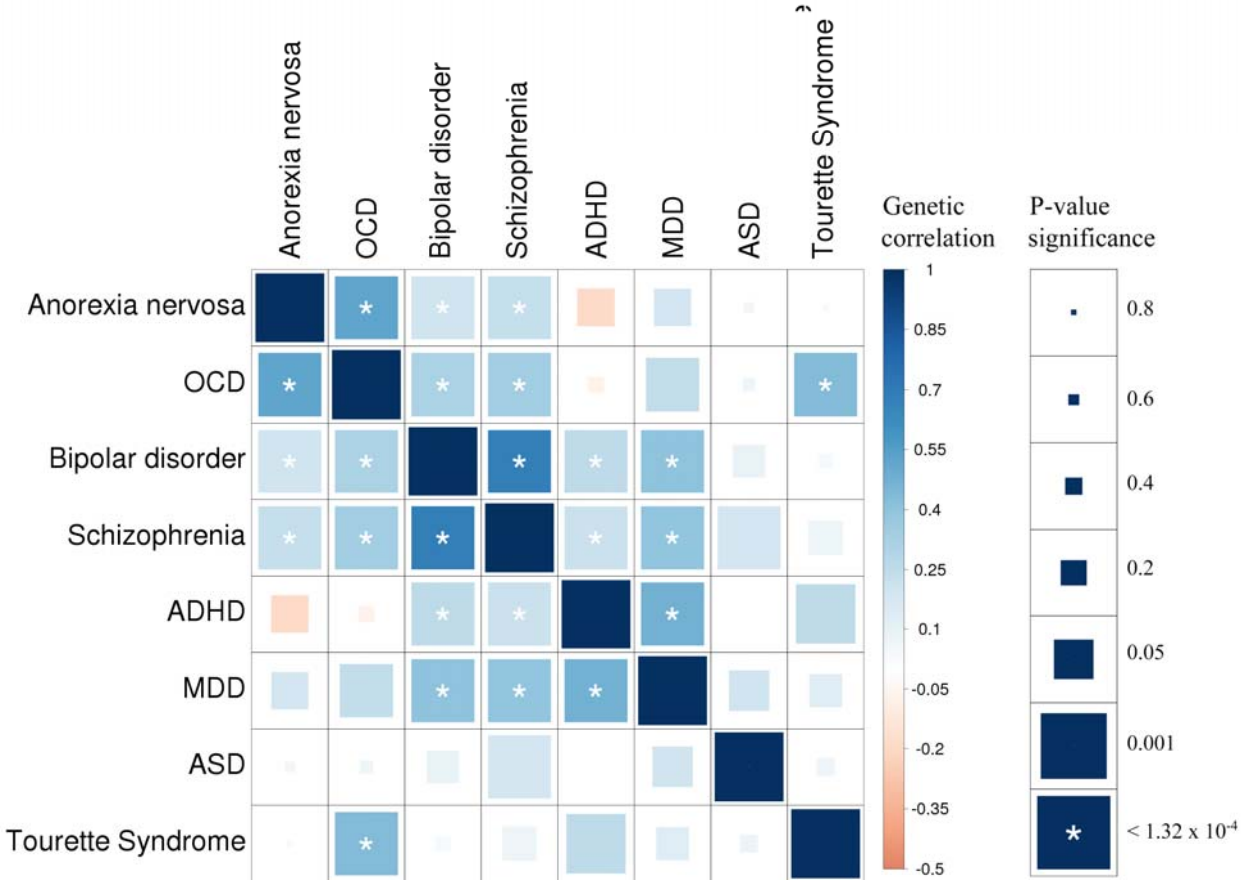
bioRxiv preprint doi: https://doi.org/10.1101/048991; this version posted April 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

67. C. A. Rietveld *et al.*, GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science* **340**, 1467 (Jun 21, 2013).
68. C. A. Rietveld *et al.*, Common genetic variants associated with cognitive performance identified using the proxy-phenotype method. *Proc Natl Acad Sci U S A* **111**, 13790 (Sep 23, 2014).
69. L. Jostins *et al.*, Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* **491**, 119 (Nov 1, 2012).
70. H. Schunkert *et al.*, Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet* **43**, 333 (Apr, 2011).
71. Tobacco, C. Genetics, Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nat Genet* **42**, 441 (May, 2010).

bioRxiv preprint doi: https://doi.org/10.1101/048991; this version posted April 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Acknowledgments The authors wish to acknowledge Claire Churchhouse, Rosy Hoskins and Joanna Martin for their helpful comments on the manuscript, Matthew Whittall for inspiration, and the patients and participants of the respective consortia. For study-specific acknowledgments, see Supplementary Materials. Data on coronary artery disease / myocardial infarction have been contributed by CARDIoGRAMplusC4D investigators and have been downloaded from www.CARDIOGRAMPLUSC4D.ORG. Data on head circumference trait has been contributed by EGG Consortium and has been downloaded from www.egg-consortium.org.

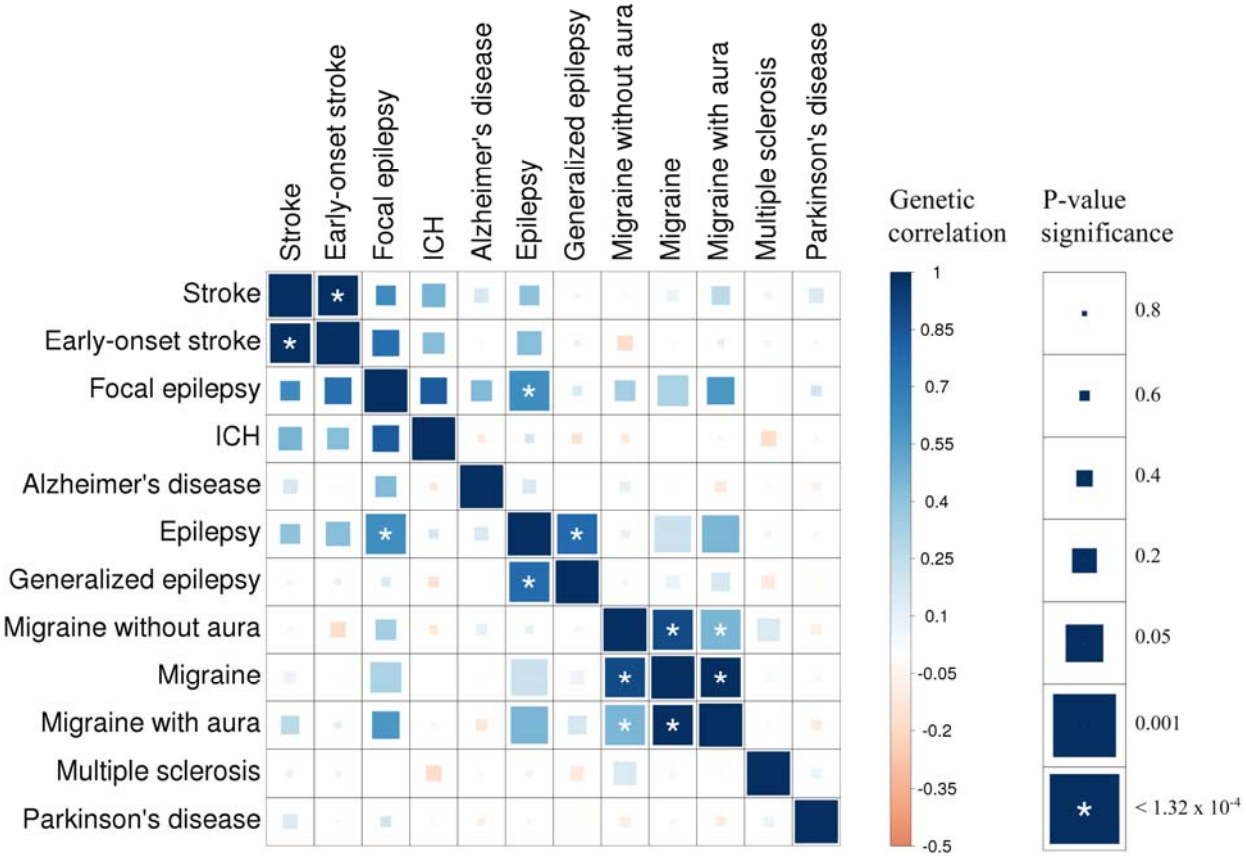
Figure 1. *Genetic correlation matrix across psychiatric phenotypes, after hierarchical clustering.*



Color of each box indicates the magnitude of the correlation, while size of the boxes indicates its significance, with significant correlations filling each box completely. Asterisks indicate genetic correlations which are significant after Bonferroni correction. ADHD – attention deficit hyperactivity disorder; ASD – autism spectrum disorder; MDD – major depressive disorder; OCD – obsessive-compulsive disorder.

bioRxiv preprint doi: https://doi.org/10.1101/048991; this version posted April 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

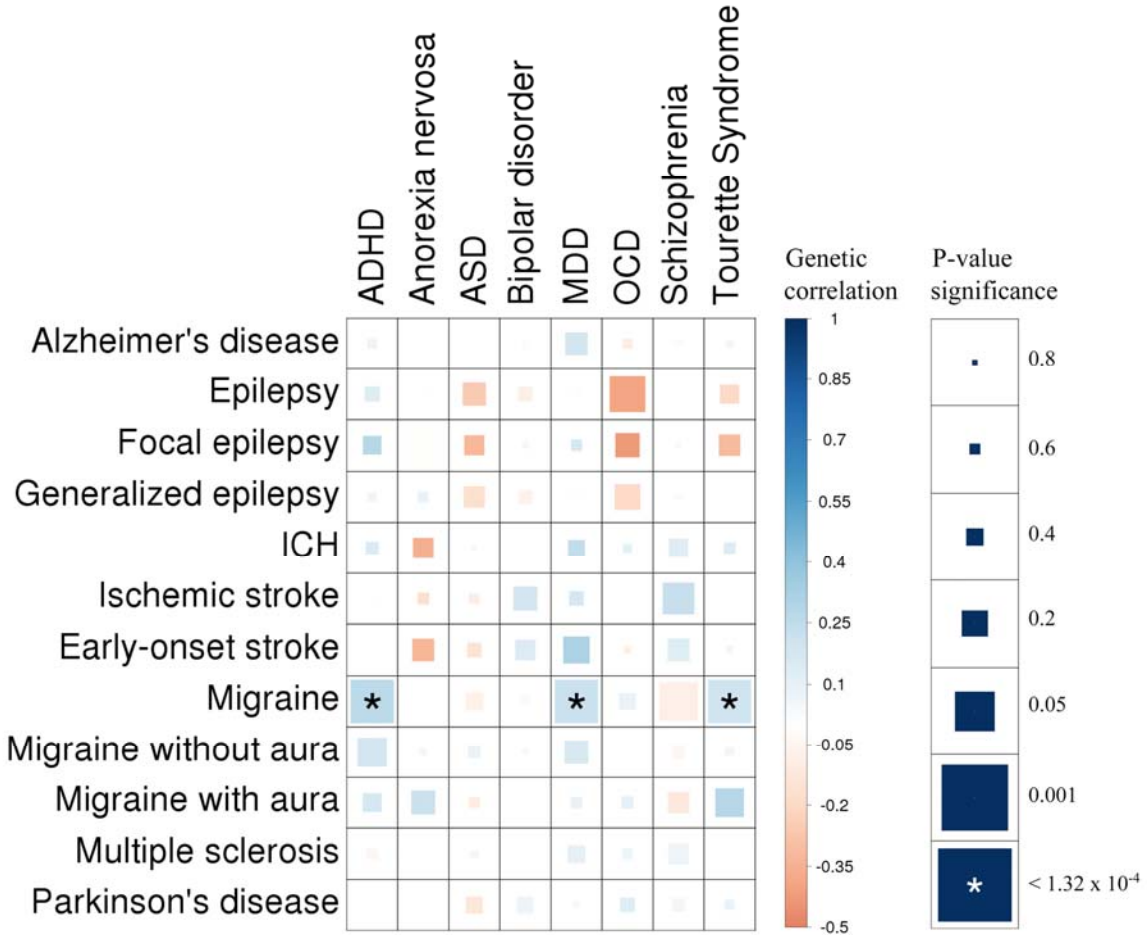
Figure 2. *Genetic correlation matrix across neurological phenotypes, after hierarchical clustering.*



Color of each box indicates the magnitude of the correlation, while size of the boxes indicates its significance, with significant correlations filling each box completely. Asterisks indicate genetic correlations which are significant after Bonferroni correction. Small black boxes indicate tests where studies have complete overlap, e.g. all cases of generalized epilepsy are also cases of epilepsy. Asterisks indicate significant genetic correlation after multiple testing correction.ICH – intracerebral hemorrhage.

bioRxiv preprint doi: https://doi.org/10.1101/048991; this version posted April 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

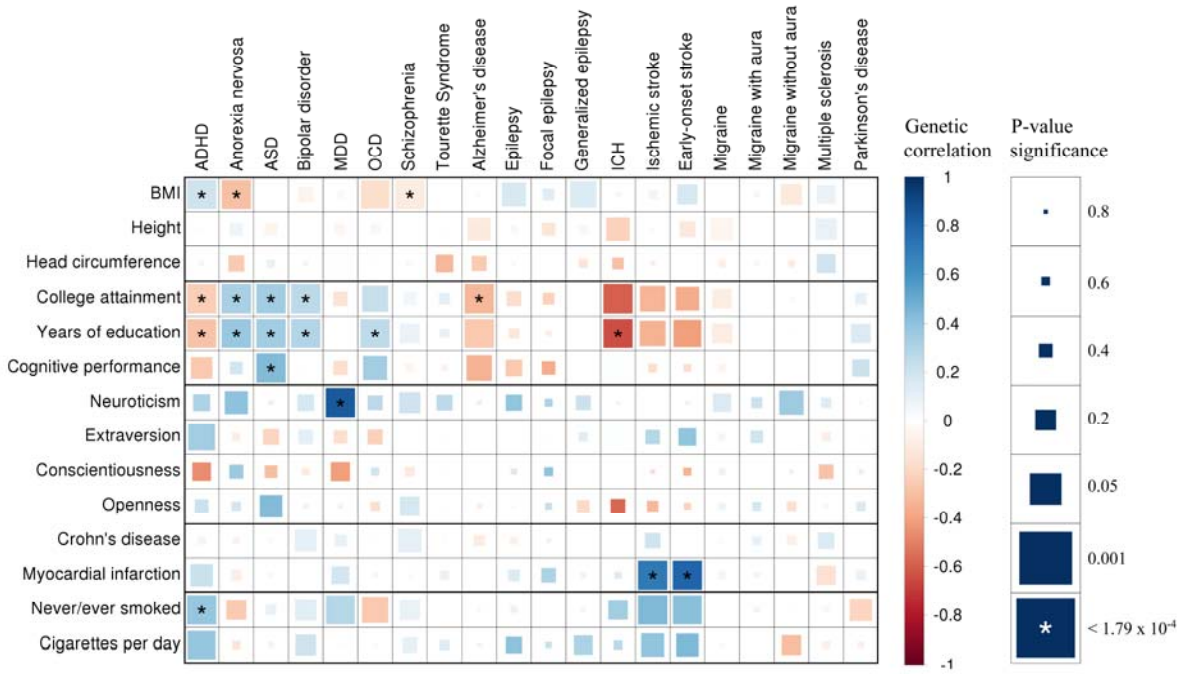
Figure 3. *Genetic correlation matrix across neurological and psychiatric phenotypes.*



Color of each box indicates the magnitude of the correlation, while size of the boxes indicates its significance, with significant correlations filling each box completely. Asterisks indicate genetic correlations which are significant after Bonferroni correction. ADHD – attention deficit hyperactivity disorder; ASD – autism spectrum disorder; ICH – intracerebral hemorrhage; MDD – major depressive disorder; OCD – obsessive-compulsive disorder.

bioRxiv preprint doi: https://doi.org/10.1101/048991; this version posted April 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 4. *Genetic correlations across brain disorders and traits of interest.*



Color of each box indicates the magnitude of the correlation, while size of the boxes indicates its significance, with significant correlations filling each box completely. Asterisks indicate genetic correlations which are significant after Bonferroni correction. ADHD – attention deficit hyperactivity disorder; ASD – autism spectrum disorder; ICH – intracerebral hemorrhage; MDD – major depressive disorder; OCD – obsessive-compulsive disorder; BMI – body-mass index.

bioRxiv preprint doi: https://doi.org/10.1101/048991; this version posted April 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Table 1. *Brain disorder phenotypes in the Brainstorm project. Indented phenotypes are part of a larger whole, e.g. the epilepsy study consists of the joint analysis of focal epilepsy and generalized epilepsy. Numbers in gray denote a control set which is non-unique, e.g. all cardioembolic stroke controls are a subset of ischemic stroke controls. ADHD – attention deficit hyperactivity disorder; OCD – obsessive-compulsive disorder. Source details are listed under Notes.*

Psychiatric phenotypes				**Neurologic phenotypes**			
Phenotype	Source	Cases	Controls	Phenotype	Source	Cases	Controls
ADHD	PGC-ADD2	12,645	84,435	Alzheimer's disease	IGAP	17,008	37,154
Anorexia nervosa	PGC-AN	3,495	11,105	Epilepsy	ILAE	7,779	20,439
Autism spectrum disorder	PGC-AUT	5,305	5,305	Focal epilepsy	"	4,601	17,985
Bipolar disorder	PGC-BIP2	20,352	31,358	Generalized epilepsy	"	2,525	16,244
Major depressive disorder	PGC-MDD2	16,823	25,632	Intracerebral hemorrhage	ISGC	1,545	1,481
OCD	PGC-OCDTS	2,936	7,279	Ischemic stroke	METASTROKE	10,307	19,326
Schizophrenia	PGC-SCZ2	33,640	43,456	Cardioembolic stroke	"	1,859	17,708
Tourette Syndrome	PGC-OCDTS	4,220	8,994	Early-onset stroke	"	3,274	11,012
				Large-vessel disease	"	1,817	17,708
				Small-vessel disease	"	1,349	17,708
				Migraine	IHGC	59,673	316,078
				Migraine with aura	"	6,332	142,817
				Migraine without aura	"	8,348	136,758
				Multiple sclerosis	IMSGC	5,545	12,153
				Parkinson's disease	IPDGC	5,333	12,019
Total psychiatric		*99,416*	*217,564*	*Total neurologic*		*107,190*	*418,650*

bioRxiv preprint doi: https://doi.org/10.1101/048991; this version posted April 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Table 2. *Traits of interest in the study. Numbers in gray denote overlapping study sets, e.g. samples in the college attainment analysis are a subset of those in the analysis for years of education. (d) – dichotomous trait, (q) – quantitative trait. BMI – body-mass index. Source details are listed under Notes.*

Phenotype	Source	Samples
Anthropometric		
BMI (q)	GIANT	339,224
Height (q)	"	253,288
Head circumference (q)	EGG	10,768
Cognitive		
Years of education (q)	SSGAC	126,559
College attainment (d)	"	120,917
Cognitive performance (q)	"	17,989
Personality measures		
Neuroticism (q)	GPC	63,661
Extraversion (q)	"	63,030
Agreeableness (q)	"	17,375
Conscientiousness (q)	"	17,375
Openness (q)	"	17,375
Disease phenotypes		
Crohn's disease (d)	IIBDGC	20,883
Myocardial infarction (d)	Cardiogram	86,995
Smoking-related		
Never/ever smoked (d)	TAG	74,035
Cigarettes per day (q)	TAG	38,617
Total		**722,125**

Supplementary Materials

Figures S1-6

Tables S1-8

Supplementary Notes

Materials and methods

Study-specific acknowledgements

Additional author notes

References S1-S25