Widespread associations between grey matter structure and the human phenome

Baptiste Couvy-Duchesne, PhD[1], Lachlan T. Strike, PhD[2], Futao Zhang, PhD[1], Yan Holtz, MS,[1,2], Zhili Zheng, MD[1,3], Kathryn E. Kemper, PhD[1], Loic Yengo, PhD[1], Olivier Colliot[4,5,6,7,8], Margaret J. Wright, PhD[2,9], Naomi R. Wray, PhD[1,2,*], Jian Yang, PhD[1,2,3,*], Peter M. Visscher, PhD[1,2,*]

1: Institute for Molecular Bioscience, the University of Queensland, 4072 St Lucia, QLD, Australia

2: Queensland Brain Institute, the University of Queensland, 4072 St Lucia, QLD, Australia

3: Institute for Advanced Research, Wenzhou Medical University, Wenzhou, Zhejiang 325027, China

4 : Inria, ARAMIS Project-team, 75013, Paris, France

5 : Institut du Cerveau et de la Moelle épinière, 75013, Paris, France

6 Inserm, U1127, 75013, Paris, France

7 CNRS, UMR 7225, 75013, Paris, France

8 Sorbonne Universitee, 75013, Paris, France

9: Centre for Advanced Imaging, the University of Queensland, 4072 St Lucia, QLD, Australia

*: These authors jointly supervised this work.

Correspondence: BCD (b.couvyduchesne@uq.edu.au), JY (jian.yang@uq.edu.au) or PMV (peter.visscher@uq.edu.au). Institute for Molecular Bioscience, 306 Carmody Rd, St Lucia QLD 4072. +61 7 3346 6348.

Abstract

The recent availability of large-scale neuroimaging cohorts (here the UK Biobank [UKB] and the Human Connectome Project [HCP]) facilitates deeper characterisation of the relationship between phenotypic and brain architecture variation in humans. We tested the association between 654,386 vertex-wise measures of cortical and subcortical morphology (from T1w and T2w MRI images) and behavioural, cognitive, psychiatric and lifestyle data. We found a significant association of grey-matter structure with 58 out of 167 UKB phenotypes spanning substance use, blood assay results, education or income level, diet, depression, being a twin as well as cognition domains (UKB discovery sample: N=9,888). Twenty-three of the 58 associations replicated (UKB replication sample: N=4,561; HCP, N=1,110). In addition, differences in body size (height, weight, BMI, waist and hip circumference, body fat percentage) could account for a substantial proportion of the association, providing possible insight into previous MRI case-control studies for psychiatric disorders where case status is associated with body mass index. Using the same linear mixed model, we showed that most of the associated characteristics (e.g. age, sex, body size, diabetes, being a twin, maternal smoking, body size) could be significantly predicted using all the brain measurements in out-of-sample prediction. Finally, we demonstrated other applications of our approach including a Region Of Interest (ROI) analysis that retain the vertex-wise complexity and ranking of the information contained across MRI processing options.

Highlights

- Our linear mixed model approach unifies association and prediction analyses for highly dimensional vertex-wise MRI data

- Grey-matter structure is associated with measures of substance use, blood assay results, education or income level, diet, depression, being a twin as well as cognition domains
- Body size (height, weight, BMI, waist and hip circumference) is an important source of covariation between the phenome and grey-matter structure
- Grey-matter scores quantify grey-matter based risk for the associated traits and allow to study phenotypes not collected
- The most general cortical processing ("fsaverage" mesh with no smoothing) maximises the brain-morphometricity for all UKB phenotypes

1. Introduction

The field of MRI studies is at a turning point owing to the recent availability of large data sets to researchers, including the UKB (Miller et al., 2016) and HCP (Van Essen et al., 2013; Van Essen et al., 2012b) samples. These datasets promote not only the replication of previous findings, but also expand the range of phenotypes available for study (e.g. psychiatric symptoms and lifestyle factors). In addition, such data sets can offer insights into the brain markers that may be shared between phenotypes, helping to draw new links between brain and behaviour. Finally, these community samples can complement the typical case-control paradigm by identifying confounders of MRI analyses or by studying related traits (e.g. cognition domains relevant in Alzheimer's disease).

Here, we introduce a set of analyses that leverages large sample sizes to fully exploit the spatial resolution of MRI images using linear mixed models (LMM) implemented in the OSCA software tool (Zhang et al., 2019). Our high-resolution approach (i.e. vertex-wise morphological measures) has the advantage of retaining all the brain complexity data of current MRI acquisitions rather than relying on prior-based data reduction techniques (e.g. the region-of-interest [ROI] approach), and allows for the elucidation of precise brain-phenotype associations.

Specifically, we used an efficient implementation of LMMs to estimate the multivariate correlation of 600,000+ cortical and subcortical measurement at vertices extracted from T1 weighted (T1w) and T2 weighted (T2w) MRI images with a phenotype of interest (previously coined morphometricity (Sabuncu et al., 2016), here we prefer the more specific brain-morphometricity). We extended this framework to also estimate the proportion of variance in a trait associated with the vertex-wise data from specific brain features, hemispheres and regions of interest. We further introduce multi-trait LMMs that can further quantify shared

84 brain-morphometricity (grey-matter correlation) between traits, reflecting causal, bi-
85 directional or confounded relationships. In addition, we show how LMMs can estimate the
86 joint effects of all brain features on a trait to construct a trait predictor from brain features
87 (grey-matter score) that can be applied and tested in an independent sample. As such, our
88 approach unifies association studies and prediction analyses, in order to unravel the brain-
89 phenome relationships (Rosenberg et al., 2018).

90 We analysed two of the largest MRI datasets available (UKB [split into discovery N=9,888
91 and replication N=4,561] and HCP [N=1,110]) and considered a wide range of phenotypes
92 spanning demographics, blood cell composition, diet, psychiatric and traumatic history,
93 physical capacities or substance use. We have released our image processing and analysis
94 software/scripts as well as all summary statistics to facilitate replication and re-use of the
95 results.

96

97 **2. Materials and Methods**

98 *2.1.UK Biobank (UKB) sample*

99 2.1.1. *Participants recruitment, inclusion and exclusion criteria*

100 The UKB participants were unselected volunteers from the United Kingdom (Sudlow
101 et al., 2015). Participants who had participated in the baseline UKB data collection were
102 invited to undergo the imaging study if they lived within travelling distance of the imaging
103 centre. Exclusion criteria were limited to: presence of metal implant, recent surgery and
104 health conditions problematic for MRI imaging (e.g. hearing, breathing problems or extreme
105 claustrophobia) (Miller et al., 2016).

106

107 2.1.2. *T1 and T2 FLAIR image collection*

108 MRI images were collected in Cheadle (greater Manchester) using a 3T Siemens
109 Skyra machine (software platform VD13) and a 32-channel head coil (Miller et al., 2016).

The T1 weighted (T1w) images were acquired over 4:54 minutes, voxel size 1.0x1.0x1.0mm, matrix of 208x256x256mm, using a 3D MPRAGE sequence (Mugler and Brookeman, 1990), sagittal orientation of slice acquisition, R=2 (in plane acceleration factor), TI/TR=880/2000 ms (Miller et al., 2016). The T2 FLAIR acquisition lasted 5:52 minutes, voxel size 1.05x1.0x1.0 mm, matrix of 192x256x256 voxels, 3D SPACE sequence (Mugler et al., 2000), sagittal orientation, R=2, partial Fourier 7/8, fat saturated, TI/TR=1800/5000ms, elliptical (Miller et al., 2016).

2.1.3. *Image processing*

We processed the T1w and T2 FLAIR images using the ENIGMA (Thompson et al., 2014) protocols for cortical surface and thickness (Stein et al., 2012) as well as subcortical structure (Gutman et al., 2013; Gutman et al., 2012). When both T1w and T2 FLAIR were available for a participant, we processed them together to enhance the tissue segmentation in FreeSurfer 6.0 (Fischl, 2012), hence a more precise skull stripping and pial surfaces definition. When the T2 FLAIR was not acquired or not usable, we processed the T1w image by itself. We retained the full image information by using the (fsaverage) vertex-wise level data in the cortical surface and thickness analyses. This corresponded to 149,960 cortical vertices in the left hemisphere and 149,933 in the right hemisphere, for each modality. In addition, we extracted subcortical radial thickness and log Jacobian determinant (that measures surface deformation from a template, somewhat analogous to a relative surface area (Roshchupkin et al., 2016)) for 27,300 vertices per hemisphere mapping 7 subcortical volumes (hippocampus, putamen, amygdala, thalamus, caudate, pallidum and accumbens) (Gutman et al., 2013). Overall, the imaging data used in the analyses comprised 654,386 vertex measurements per individual: 299,893 describing cortical thickness, another 299,893

for cortical surface area, 27,300 for subcortical thickness and 27,300 for subcortical curvature.

For comparison with previous ENIGMA publications, we also extracted cortical thickness and surface area of 34 cortical regions delimited by the Desikan atlas (Desikan et al., 2006; Fischl et al., 2004), as described on the ENIGMA website. To further the comparison of processing options, we extracted cortical measurements from smoothed fsaverage meshes (fwhm 5, 10, 15, 20 and 25mm) as well as (unsmoothed) coarser meshes provided by FreeSurfer: fsaverage6 (149,091 vertices for both hemispheres and modalities), fsaverage5 (37,455 vertices), fsaverage4 (9,457 vertices) and fsaverage3 (2,414).

2.1.4. *Discovery Sample description*

At the time of download (July 2017), T1w images were available for 10,102 participants of the UK Biobank (UKB) project. None of the participants had withdrawn consent after the data was collected. We excluded 175 participants with T1w images labelled as unusable by the UKB, leaving 9,928 MRI scans to process. T2 FLAIR images were available for 9,755 of those. The FreeSurfer processing failed or did not complete within 48 hours for a handful of participants: 37 for cortical processing, 19 for subcortical, including 17 for whom both processing failed. For simplicity, we chose not to re-run image processing on these participants as their exclusion should have a minimal impact on the results obtained from the full sample. Excluded individuals are described in **Dataset S1**. Our final sample comprised 9,890 participants with usable cortical data, 9,908 with subcortical data and 9,888 with both cortical and subcortical data. This sample consisted of 9,888 adults aged 62.5 on average (SD=7.5, range 44.6–79.6) and comprised 52.4% of female participants.

We excluded 391 participants with extreme brains (outliers) or likely to have a large effect on the analyses (see **Appendix S1** for details of QC and **Dataset S1** for description of the excluded participants).

2.1.5. *Variables used*

We included 168 variables grouped in several categories: demographics, cognition, physical test, psychiatry, recent feelings, stress and traumas, substance use, miscellaneous, brain measurements, blood assay and diet (see **Dataset S2** for details). When longitudinal observations were available for a participant, we used the one collected as part of the imaging assessment (when available) or the closest in time.

2.1.6. *Replication Sample description*

Replication data set was downloaded in May 2018 and consisted in an additional 4,942 participants with a T1w image. Image processing and phenotype selection were identical to that of the discovery sample. This led to the exclusion of 381 participants whose processing failed and 238 excluded from QC. The final sample (N=4,323) included in the replication analysis was on average 63.1 years old (SD=7.46, range 46.1-80.3) with 52.1% of females. The age difference between discovery and replication sample was small but significant (p=9.02e-7). See **Dataset S1** for a full description of replication participants (final, QCed and failed processing) in addition to a comparison of the discovery and replication samples.

2.2.Human Connectome Project (HCP) sample

2.2.1. Participants recruitment, inclusion and exclusion criteria

HCP participants were recruited from ongoing longitudinal studies of the Missouri Family Study (Edens et al., 2010; Sartor et al., 2011) and had to be between 22 and 35 years

of age. Inclusion and exclusion criteria have been described previously (Van Essen et al., 2012b).

2.2.2. T1 and T2 weighted image collection

T1w and T2 weighted (T2w) images were collected at the Washington University (St Louis, Missouri) on a 3T Siemens Skyra scanner using a standard 32-channel head coil (Van Essen et al., 2013; Van Essen et al., 2012b). Two T1w images were acquired, each over 7 minutes and 40 seconds with a voxel size of 0.7x0.7x0.7mm, matrix/FOV of 224x224x224mm using a 3D MPRAGE sequence (Mugler and Brookeman, 1990), TR/TE/TI=2400/2.14/1000ms, flip angle 8degrees, R=2, sagittal orientation of slice acquisition (Glasser et al., 2013). Similarly, two T2w images were acquired over 8:24 min each, voxel size 0.7x0.7x0.7mm, matrix of 224x224x224mm, 3DSPACE sequence (Mugler et al., 2000), sagittal orientation, R=2, TR/TE=3200/565, no fat suppression pulse.

2.2.3. *Image processing*

The HCP team (Glasser et al., 2013; Marcus et al., 2013; Van Essen et al., 2012a) pre-processed the structural scans to facilitate scan comparison across individuals, removing spatial artefacts and improve T1w and T2w alignment using FSL (Jenkinson et al., 2002; Jenkinson et al., 2012) and FreeSurfer (Fischl, 2012). When both passed HCP quality control (QC), T1w and T2w images they processed them together in FreeSurfer 6.0 (Fischl, 2012), otherwise data extraction relied on a single scan (Glasser et al., 2013). Participants with poor quality T1w and T2w scans were re-imaged (Glasser et al., 2013). Cortical processing (recon-all procedure in FreeSurfer) was also performed by the HCP team and included down sampling to 1mm size voxels and 256x256x256 matrix, aided registration using customised brain mask, and two manual steps performed outside of the recon-all procedure to enhance

white matter and pial reconstruction (Glasser et al., 2013). We downloaded the processed images (Marcus et al., 2011) and performed ENIGMA shape analysis (Gutman et al., 2013; Gutman et al., 2012) to extract vertex-wise measurements of the subcortical thickness and curvature. As for the UKB sample, a total of 654,386 vertex measurements were extracted for each individual. We excluded 24 outliers with extreme brains or likely to bias the analyses (see **Appendix S1** and **Dataset S2** for description of excluded participants).

2.2.4. *Sample description*

As per the HCP "1200 Subjects data release" (1st of March 2017), 1,113 participants were scanned on the 3T MRI and underwent extensive behavioural testing. Participants were mostly (54.4%) females and were 28.8 years old on average (SD=3.7, range 22–37). The sample comprised 455 twins (41.0%), 286 monozygotic twins (138 complete pairs) and 169 dizygotic twins (78 complete pairs). In addition, siblings and half siblings of twins were also recruited which resulted in only 445 distinct families in the sample.

2.2.5. *Variables used*

For the HCP sample, we included 161 variables, some of which were also available in the UKB (e.g. demographics, cognition, physical assessment, blood assay or psychiatry). We also included interesting variables only present in the HCP sample: personality, emotion, in depth mental health assessment (Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA) and Adult Self Report (ASR) (Achenbach, 2009; Achenbach et al., 2003)), detailed cognition, Pittsburgh sleep index (PSQI) (Buysse et al., 1989), or results from the urine drug tests (see **Dataset S2**).

2.3.Variance component analyses and brain relatedness matrix calculation

2.3.1. *The linear mixed model*

We aimed to estimate the variance of a trait accounted for by brain features, which Sabuncu et al., called "morphometricity" (Sabuncu et al., 2016). To do so we consider the following linear mixed model that allows estimating the association between a phenotype and M vertices even when M is greater than the sample size (N):

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{b} + \mathbf{e} \quad (1)$$

where $\mathbf{Y}_{N,1}$ is the phenotype considered with N the number of observations, $\mathbf{X}_{N,c}$ is a matrix of c covariates (as such does not include any vertex variable), $\boldsymbol{\beta}_{c,1}$ is a vector of fixed effects, $\mathbf{b}$ is a random effect with $\mathbf{b} \sim \mathcal{N}\left(0, \mathbf{B}\sigma_b^2\right)$ and $\mathbf{e}$ is the error term with $\mathbf{e} \sim \mathcal{N}(0, \mathbf{I}\sigma_e^2)$. In this formulation $\mathbf{I}_{N,N}$ is the identity matrix as we assume the error terms to be independent and identically distributed. $\mathbf{B}_{N,N}$ is a matrix of variance-covariance between individuals calculated from all vertex measurements, which we will refer to as the brain relatedness matrix (BRM). Off diagonal elements of the BRM can be interpreted as a measure of brain similarity between two individuals (see **S2 Appendix**). Finally, σ_e^2 and σ_b^2 are the variance components for the random effects $\mathbf{e}$ and $\mathbf{b}$. For context, this model is analogous to that used in complex trait genetics to estimate SNP-based heritability, where a Genetic Relatedness Matrix (GRM) replaces the BRM (Yang et al., 2010; Yang et al., 2011). The element i,j of the BRM can be calculated as the inner product of brain measurements of individuals i and j: $b_{i,j} = \frac{\sum_1^M z_{i,m} z_{j,m}}{M}$. Here, $z_{i,m}$ represents the value of vertex m for individual i centred and standardised by its standard deviation over all individuals, $z_{j,m}$ the value of vertex m for individual j centred and standardised over all individuals, M is the total number of vertices or brain features included. We can equivalently use matrix notation, then: $\mathbf{B} = \frac{\mathbf{Z}\mathbf{Z}'}{M}$, with $\mathbf{Z}_{N,M}$ a matrix of the centred and standardised brain observations, for N

individuals and M brain features. We are interested in estimating the parameters σ^2_e and σ^2_b so we can derive the proportion of the trait variance captured by the brain similarities: $R^2 = \frac{\sigma^2_b}{\sigma^2_b + \sigma^2_e}$. To do so we used the REstricted Maximum Likelihood (REML) method (Patterson and Thompson, 1971) implemented in OSCA.

2.3.2. *Mixed model with several random effects*

Here, we are dealing with several types of brain measurements: cortical vs. subcortical or thickness vs. surface area for instance. To accommodate the different modalities, we can extend the LMM presented above to jointly estimate the variance accounted for by the different types of measurements: $\mathbf{Y} = \mathbf{X\beta} + \mathbf{b_1} + \mathbf{b_2} + \mathbf{b_3} + \mathbf{b_4} + \mathbf{e}$ now, with $\mathbf{b}_i \sim \mathcal{N}\left(0, \mathbf{B}_i\sigma^2_{bi}\right)$, $i \in [\![1;4]\!]$, and all other parameters left unchanged. Note that since all $\mathbf{b}_i$ are estimated jointly, each estimate is conditional on the other three parameters fitted in the model. We constructed the BRM $\mathbf{B_1}$ from the cortical thickness measurements, $\mathbf{B_2}$ from the cortical surface area, $\mathbf{B_3}$ from the subcortical radial thickness and $\mathbf{B_4}$ from the subcortical curvature. The variance components σ^2_{bi} quantify the specific variance attributed to each type of measurement and the quantity $\frac{\sigma^2_{b1}+\sigma^2_{b2}+\sigma^2_{b3}+\sigma^2_{b4}}{\sigma^2_{b1}+\sigma^2_{b2}+\sigma^2_{b3}+\sigma^2_{b4}+\sigma^2_e}$ represents the proportion of the trait variance captured by all our brain measurements not biased towards the cortical measurements.

2.3.3. *Bivariate models to estimate grey-matter correlation*

Finally, we are interested in estimating the correlation (or covariance) between two traits that is attributable to the same brain similarities, which we call grey-matter correlation r_{GM}. This can be achieved by fitting a bivariate LMM, a direct extension of the models presented above (Thompson, 1973). We used the AI-REML algorithm in GCTA (Lee

et al., 2012) as the multivariate option is not yet available in OSCA. We restricted our bivariate analysis to variables that were significantly associated with grey-matter structure. We derived the residual correlations (r_E) from the phenotypic (r) and grey-matter correlations estimated by GCTA: $r_E = \frac{r - r_{GM} * \sqrt{R_1^2 * R_2^2}}{\sqrt{(1-R_1^2) * (1-R_2^2)}}$ with R_1^2 and R_2^2 the brain-morphometricity of the two traits included in the bivariate model. For significance testing, we derived SE of r_E) from a first order Taylor series approximation (delta method, see **Appendix S3** and (Bijma and Bastiaansen, 2014; Lee et al., 2012; Visscher, 1998)).

2.4. *Covariates used*

Our baseline model included commonly used covariates in MRI analyses: acquisition variables (UKB imaging wave, processing with T1w or with combined T1w+T2w), age, sex, and head size (intra-cranial volume (ICV) as well as left and right total cortical surface area and cortical thickness that correspond to the measurements used here). In a follow-up analysis, we included other covariates such as height, weight and BMI to evaluate their confounding effect on the reported associations. We reported the associations between phenotypes and covariates using the adjusted R-squared calculated from linear models estimated in R3.3.3 (R Development Core Team, 2012). As some of the covariates are correlated we report the R^2 calculated by adding progressively the covariates (same order as above). Thus, the fixed effect R^2 should not be compared between covariates, but can be contrasted between phenotypes or with the random effect R^2. We compared the covariates' associations with our phenotypes of interest in the UKB discovery and replication samples and found highly concordant results between the two samples (**Figure S1**). Thus, any brain-

morphometricity difference found between UKB discovery and replication sample should reflect a true difference in the phenotype grey-matter structure association.

2.5.*Test statistics in mixed linear models*

We tested whether the variance accounted for by the brain similarities was significantly different from 0 using a likelihood ratio test on nested models (with and without the random effect). The test statistic follows a chi-square distribution with x degree of freedom (x being the number of variance components tested) for a σ_0^2 value inside the parameter space. However, when testing $H0$: $\sigma_0^2 = 0$ $vs.$ $H1$: $\sigma_1^2 > 0$, the p-value should be interpreted with caution as the estimator may not be asymptotically normally distributed because 0 is a boundary of the parameter space (Self and Liang, 1987; Stram and Lee, 1994). Some have suggested that the p-value could be better approximated using a mixture of chi-square distributions in the test of significance (Self and Liang, 1987; Stram and Lee, 1994). However, a 50:50 mixture has been shown to be sometimes inappropriate (Crainiceanu and Ruppert, 2004; Pinheiro and Bates, 2000) as the test relies on assumptions often not met in LMM (such as i.i.d. observations) (Crainiceanu and Ruppert, 2004). Thus, we preferred using a $\chi 2$ (x df.), the only consequence being a less powerful hence conservative test (Bates et al., 2015; Crainiceanu and Ruppert, 2004; Pinheiro and Bates, 2000). Such test is implemented in OSCA (Zhang et al., 2019), as well as in GCTA (Yang et al., 2011).

2.6.Statistical power of the current analyses

In the UKB discovery sample (assuming N=9,500), we have 80% power to detect an effect >2.2% of variance accounted for by the combined BRM (gathering all features), while taking into account multiple testing (pvalue significance threshold p<0.05/175, to ensure a type I error<5%). In the HCP sample (assuming N=1,000), considering the number of tests performed (p<0.05/160), we would need an effect of 20% of variance accounted for to yield

the same power (**Appendix S4 and** (Visscher et al., 2014)). For brain correlations, the calculation of statistical power depends on the sample size (set to 9,500), the variance accounted for in each phenotype (we chose 5%), the phenotypic correlation (set to r=0.2), the significance threshold (p<4.2e-5, based on our number of tests) as well as the variance of off-diagonal elements of the BRM $var(B_{ij})$ (0.00096, for the BRM of all brain features) (Visscher et al., 2014). In this example, we had 80% power to detect a brain correlation greater than 0.35, but only a 7% power for a brain correlation of 0.2. Using a sample of N=1,000, as per the HCP, and selecting phenotypes with >20% variance accounted for (everything else being equal), we have a 1% power to detect a brain correlation of 0.35, and we would need a brain correlation greater than 0.99 to achieve 80% power.

2.7.Vertex level associations of specific brain features and regions

We conducted post-hoc analyses to identify associations with each type of brain measurement (i.e. left or right measurements of cortical thickness, cortical surface area, subcortical curvature and subcortical thickness) in each cortical (Desikan-Killiany atlas (Desikan et al., 2006)) or subcortical region. For this, we used BRMs specific to each region and brain measurement. Brain regions of interest (ROI) contained between 272 and 12,179 vertices in the left cortex, and between 369 and 11,878 for the right hemisphere. The smallest ROI was the frontal pole and the largest the superior frontal gyrus. Subcortical structures ranged from 930 vertices (Accumbens) to 2502 (Caudate, Hippocampus and Putamen) (Gutman et al., 2013; Gutman et al., 2012). We used the same covariates as in previous LMMs.

2.8.In sample prediction (10-fold cross-validation)

352 We derived brain prediction scores using the Best Linear Unbiased Predictors (BLUP)
353 (Henderson, 1950, 1975; Robinson, 1991) and evaluated them in the UKB discovery sample
354 using a 10-fold cross-validation design. Note that the BLUP predictor was derived from the
355 LMM (REML) analysis described above. When measuring the correlation between grey-
356 matter scores and observed value, we controlled for the same covariates used in the LMMs
357 and included dummy variables to account for hypothetical differences between the groups
358 selected in the cross-validation design. BLUP estimates the predicted values of the random
359 effects (here, **b**, see equation 1) instead of relying on the estimates of fixed effects for all
360 brain features (Goddard et al., 2009; Robinson, 1991). In short, BLUP scores integrate the
361 correlations between vertices to derive weights that correspond to the joint effects of all
362 the vertices. BLUP have desirable statistical properties: they are unbiased and are best
363 predictors in the sense that they minimise the mean square error in the class of linear
364 unbiased predictors (Henderson, 1975; Robinson, 1991), leading to more accurate
365 prediction than other linear predictors (Robinson et al., 2017; Vilhjalmsson et al., 2015).
366 Among others, BLUP scores are routinely used in animal breeding (Robinson, 1991),
367 prediction of individual genetic risk (Robinson et al., 2017) as well as to calculate
368 transcriptomic or methylation age (Peters et al., 2015). BLUP predictors can be calculated in
369 OSCA (Zhang et al., 2019) from summary statistics (analogous to GCTA-SBLUP (Robinson et
370 al., 2017)) and known correlation between vertex measurements.
371
372 *2.9.Out of sample prediction*
373 Finally, we derived BLUP brain prediction scores constructed from the UKB discovery
374 sample, and applied them to the UKB replication and HCP participants. We evaluated the

375 predictive performance using the correlation between grey-matter score and corresponding
376 observed phenotype, controlling for covariates used in the LMMs.
377
378 *2.10. Application of LMMs to identify "best" cortical processing*
379 Here, we defined as "best" processing the MRI cortical processing that maximises
380 the association with a trait of interest, from the minimal number of features (vertices).
381 Thus, we evaluated which of our 10 FreeSurfer processing (fsaverage – no smoothing;
382 fsaverage – smoothing fwhm5, 10, 15, 20, 25; fsaverage6, 5, 4, 3 – no smoothing; ENIGMA
383 ROI processing; **see 2.1.3**) maximised the brain-morphometricity, for all UKB traits
384 considered.
385 As the ENIGMA processing only consists of 150 measurements (14 subcortical volumes
386 measurements, cortical surface or thickness averaged over 78 ROI defined by the Desikan-
387 Killiany atlas (Desikan et al., 2006)) we used generalised linear models (GLMs – multiple
388 regression) to estimate the brain-morphomometricity. For context, the LMM approach used
389 in the vertex level analyses is a direct extension of GLMs that allows the number of features
390 to exceed the number of participants (p>N).
391
392 *2.11. Data and code availability statement*
393 Data used in this manuscript is held and distributed by the HCP and UKB teams. We
394 have released the scripts used in image processing and LMM analyses to facilitate
395 replication and dissemination of the results (see **URLs**). We have also released BLUP weights
396 to allow meta-analyses or application of the grey-matter scores in independent cohorts.
397
398 **3. Results**

3.1.Associations between phenotypes and all grey-matter structure vertices

For the phenotypes of interest, we summarised in circular barplots (**Figure 1**) the association (R^2) with all 654,386 vertex-wise grey-matter measures extracted during image processing, as well as with covariates (acquisition, age, sex and brain/head size variables - see **Methods**). The R^2 may be interpreted as the proportion of variance in a phenotype captured by all grey-matter morphology. **Figure 1** shows only the significant results (Bonferroni significance threshold; $p_{UKB\_discovery}$=0.05/175=2.8e-4, p_{HCP}=0.05/160=2.9e-4) with the full results available in **Dataset S3, S4** (see **Figure S2** for positive control associations)**.**

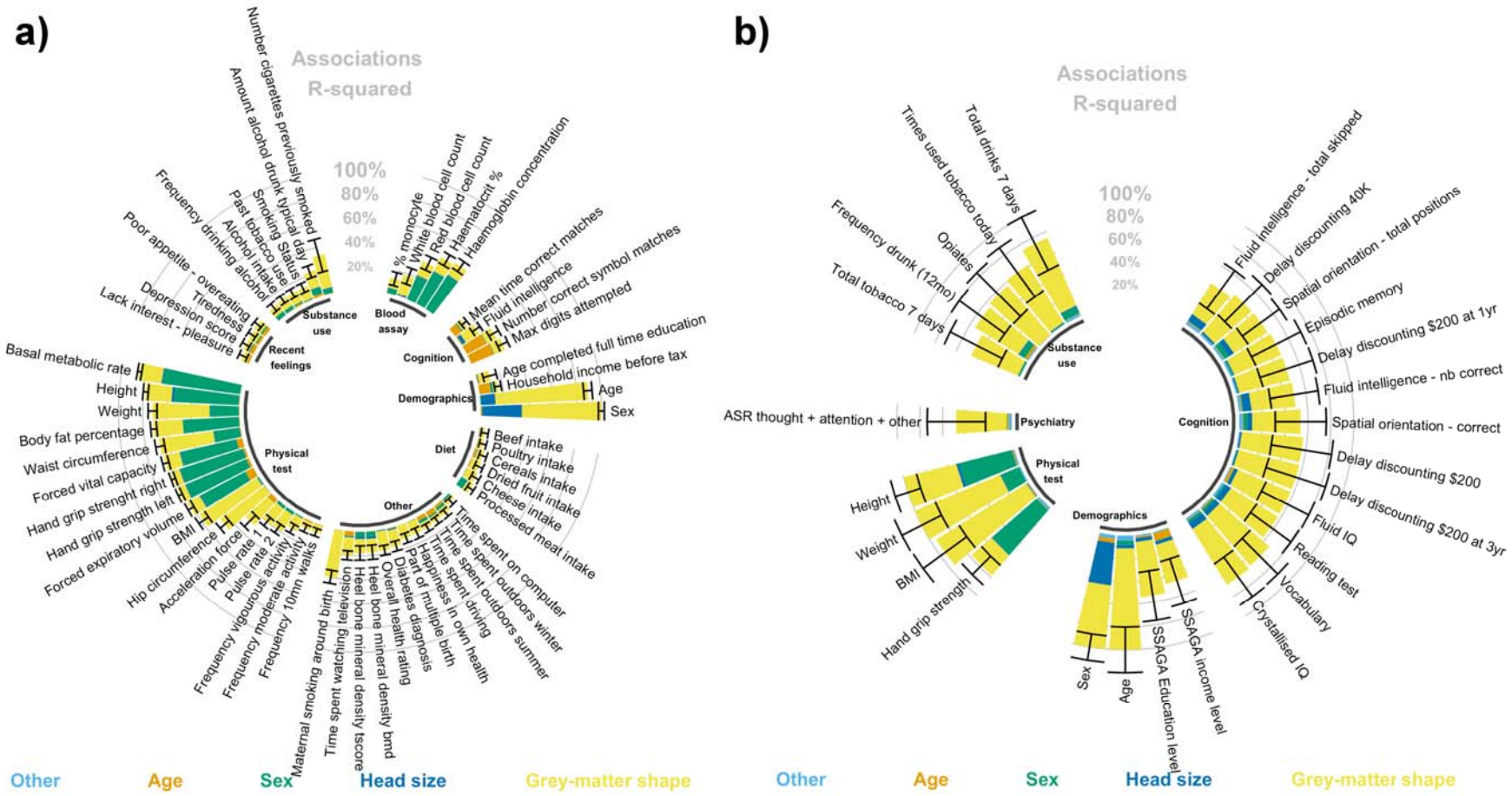
Grey-matter structure was strongly associated with age (R^2_{UKB}=0.77, SE=0.018; R^2_{HCP}=0.88, SE=0.10), sex (R^2_{UKB}=0.66, SE=0.012; R^2_{HCP}=0.56, SE=0.059), as well as height (R^2_{UKB}=0.22, SE=0.011; R^2_{HCP}=0.47, SE=0.060) weight (R^2_{UKB}=0.47, SE=0.019; R^2_{HCP}=0.81, SE=0.099) and BMI (R^2_{UKB}=0.57, SE=0.024; R^2_{HCP}=0.92, SE=0.12). Measures of build, body fat and metabolism were also associated with grey-matter structure (R^2_{UKB}=0.45, SE=0.019 with waist circumference, R^2_{UKB}=0.24, SE=0.013 with body fat percentage, R^2_{UKB}=0.19, SE=0.009 with basal metabolic rate; corresponding measures not available in the HCP dataset). In addition, grey-matter structure was associated with measures of strength in both samples (e.g. hand grip: R^2_{UKB}=0.074, SE=0.009; R^2_{HCP}=0.23, SE=0.58) and levels of physical activity (R^2_{UKB} ranging between 0.059-0.25, not-significant in the HCP).

Grey-matter structure was further associated with cognitive domains (R^2_{UKB} ranging in 0.048-0.13, R^2_{HCP} in 0.34-0.57), smoking (R^2_{UKB} ranging in 0.11-0.28, R^2_{HCP} in 0.45-0.65), alcohol consumption (R^2_{UKB} ranging between 0.071-0.14, R^2_{HCP}=0.63, SE=0.13), educational attainment (R^2_{UKB}=0.097, SE=0.029; R^2_{HCP}=0.39, SE=0.11) and income level (R^2_{UKB}=0.042, SE=0.014; R^2_{HCP}=0.32, SE=0.10). Associations with diet, blood assay results, depression score

and symptoms, diabetes, bone density, lifestyle and maternal smoking around birth were only observed in the UKB, the phenotypes not being available in the HCP (**Figure 1**).

We replicated 23 of the 58 associations listed above in the UKB replication sample ($p<0.05/58$; **Figure S3a**). Replication of blood assay phenotypes was limited due to the small sample sizes (N~300), being only collected for the first imaging waves. Beyond statistical significance that depends on sample and effect sizes, the brain-morphometricity estimates were highly similar between the discovery and replication UKB samples (cor=0.95, excluding blood assay, **Figure S3b**). Full replication results have been added to **Dataset S4**.

In the UKB (discovery), results and conclusions did not change regardless of fitting a single random effect or several random effects each corresponding to one of the grey-matter modalities (i.e. cortical thickness, cortical surface, subcortical thickness, and subcortical area) (**Figure S4**). In the HCP, we observed 3 extra significant associations between grey-matter structure and cocaine (urine test), self-reported number of times used cocaine or hallucinogens. Similar to the association found with opiate (urine test), the small number of positive participants warrants replication. We chose not to include these 3 variables in the subsequent analyses.



439

440 **Figure 1: Circular barplot of the associations (R^2) between phenotypes and grey-matter structure vertices (morphometricity)**

441 For clarity, we only plotted the significant associations in the UKB discovery (panel **a**) and HCP sample (panel **b**). We applied Bonferroni

correcting to account for multiple testing in each sample. The black bars represent the 95% confidence intervals of the morphometricity estimates. For context, we also present the association R^2 between phenotypes and covariates of the baseline model, as per the legend under the barplot. As some covariates may be correlated, the R^2 was calculated by adding progressively the covariates in that order: acquisition and processing variables (labelled "other"), age, sex and head size (ICV, total cortical thickness and surface area). Age and sex were not included as covariates when studying them as phenotypes. See **Dataset S3-4** for full results. See **Figure S1** for positive control associations.

3.2.Controlling for body size

The large associations between grey-matter structure and height, weight, BMI, waist and hip circumference (even after controlling for acquisition, age, sex and head size differences**, Figure 1**) led us to perform a sensitivity analysis to evaluate their contribution to the brain-morphometricity of the traits studied. We repeated the analysis further controlling for height, weight and BMI, which yielded lower R^2 estimates (**Figure S5**) and fewer significant associations with grey-matter structure. Thus, when correcting for height in the UKB, 4 of the 58 associations with grey-matter structure did not remain significant: household income, monocyte percentage, beef intake, and time spent using computer, see **Dataset S3**). Such finding is consistent with the reported association between body size and income or socio-economic status in the UKB (Tyrrell et al., 2016). When further correcting for weight and BMI another 14 associations did not remain significant including educational attainment, frequency drinking alcohol, most diet items (cereal, dried fruits, poultry, processed meat), time spent driving, red blood cell count, frequency of walks and small exercise. Notably, the brain-morphometricity of the depression score could be completely explained by differences in weight and BMI ($R^2_{baseline}$=0.050, SE=0.018; $R^2_{baseline+height}$=0.048, SE=0.017, $R^2_{baseline+height+BMI+weight}$<0.001, SE=0.007) and none of the associations between grey-matter structure and depression symptoms remained significant (Tiredness, Anhedonia, Poor appetite- overeating, $R^2_{baseline+height+BMI+weight}$<0.014). The brain-morphometricity estimates (correcting for body size) in the UKB replication sample aligned with results from the discovery sample (cor=0.90), except for age and sex showing larger associations with grey-matter structure in the replication analysis (**Figure S6**).

Similarly, 4 of the 27 associations did not remain significant in the HCP dataset: with fluid intelligence (total skipped), hand grip strength, ASR thought and attention problems

and frequency of being drunk in the past year. Though we had limited power to detect associations smaller than R^2 of 0.2 in this sample (see **2.6**).

In light of these results, we chose a conservative approach to present in the main text results that include body size variables as covariates, though the analyses using baseline covariates can be found in the supplementary. We acknowledge (see discussion) that this may be overly conservative, by implicitly making strong assumptions about the direction of causation between body shape, grey-matter morphology and the rest of the phenome. On the other hand, it avoids reporting associations that may be fully or in part caused by differences in body shape.

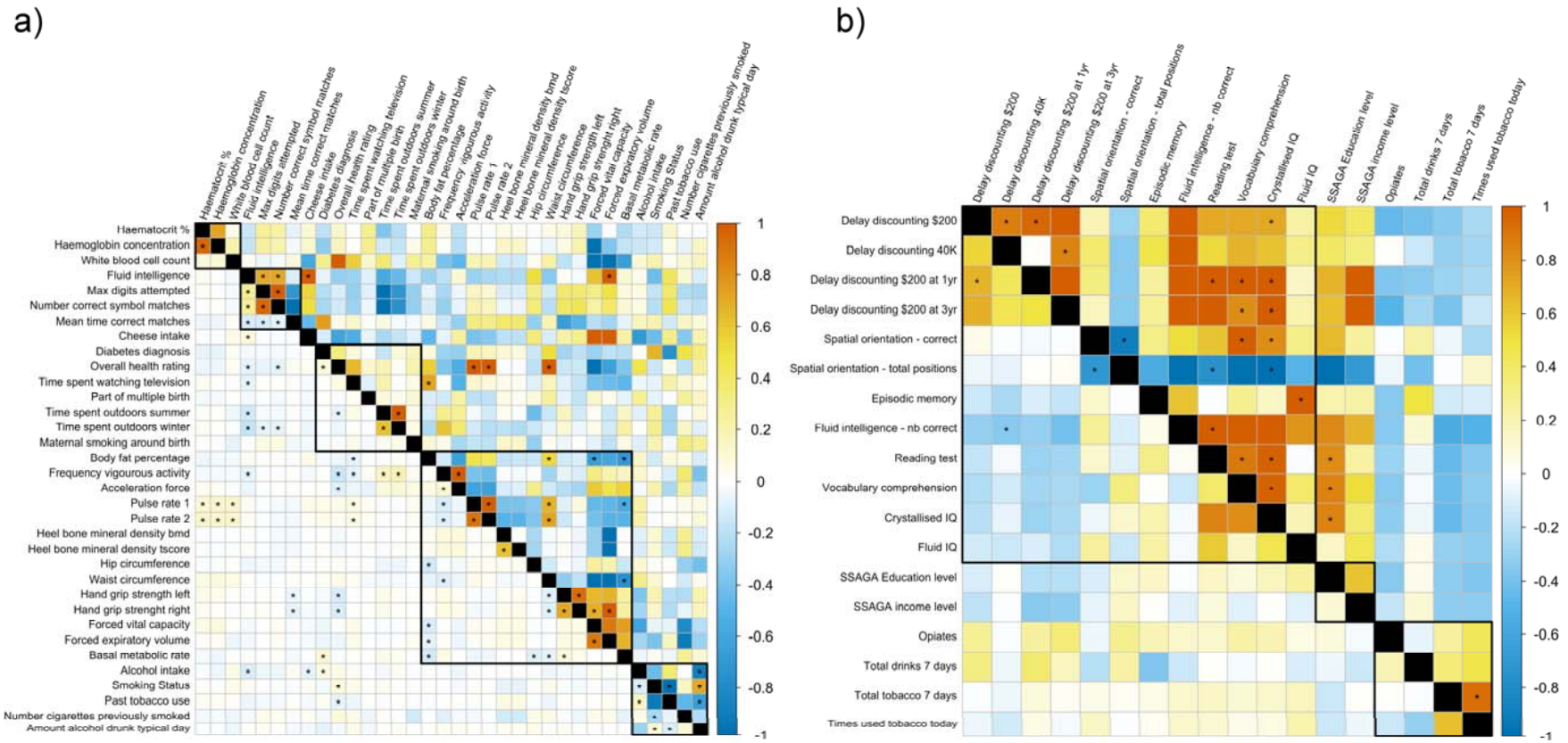


Figure 2: **Matrices of grey-matter correlations (upper diagonals) and residual correlations (lower diagonals) between all the variables showing significant morphometricity after controlling for baseline covariates, as well as height, weight and BMI.**

Panel (a) shows the results for the UKB and panel (b) the HCP results. Correlations significant after multiple testing correction (Bonferroni) are indicated by a star. Blocks circled in black indicate the different phenotype categories used previously (see **Figure 1**). Most grey-matter

correlations are observed within categories (e.g. cognition or substance use) but they can also help identifying shared brain-morphometricity between different types of variables (e.g. cheese intake and pulse rate). r_{GM} is a measure of the shared brain-morphometricity between 2 traits and can arise from causal, bi-directional or confounded relationships between phenotypes. Contrasting r_{GM} and residual correlation (r_E) can indicate how much of the phenotypic correlation is attributable to individual's resemblance in term of grey-matter structure, compared to other factors (brain or non-brain resemblances).

3.3.Grey-matter correlations

We estimated grey-matter and residual correlations (r_{GM} and r_E) between the phenotypes that showed significant brain-morphometricity in the univariate analyses. r_{GM} can be interpreted as the proportion of grey-matter vertices similarly associated with both traits, while r_E offers insight into factors, shared between the traits, but that do not relate to grey-matter structure (e.g. other brain modalities, non-brain contribution). A weighted sum of r_{GM} and r_E make up the phenotypic correlation (see **2.3.3**). In this section, we controlled for height, weight and BMI on top of the baseline covariates, which yields a conservative set of 39 phenotypes and prevents results from being confounded by body size (**Figure 2; Datasets S5** (UKB), **S6** (HCP) for point estimates**)**. We excluded phenotypes used as covariates (age, sex, head and body size) as regressing them out makes them orthogonal (i.e. not associated) with the remaining traits. We used conservative significance thresholds of 0.05/(35*34)=4.2e-5 for UKB and 0.05/(18*17)=1.6e-4 for HCP that account for the total number of correlations performed in each sample. We highlighted below which grey-matter correlations were also significant in the UKB replication sample (significance threshold p<0.05/ntest i.e. p<1.9e-3).

In the UKB, we observed significant positive grey-matter (and residual) correlations between cognition domains (r_{GM} ranging between 0.71, SE=0.12 and 1.0, SE=0.007; corresponding r_E ranging between 0.26, SE=0.014 and 0.94, SE=0.005; **Figure 2**). In addition, we identified grey-matter correlations between measures of physical activity. For example, body fat percentage correlated with waist circumference (r_{GM}=0.52, SE=0.12), forced vital capacity (r_{GM}=-0.66, SE=0.14), basal metabolic rate (r_{GM}=-0.69, SE=0.094, $r_{GM\text{-}replication}$=-0.75, SE=0.13) and time spent watching TV (r_{GM}=0.73, SE=0.13, $r_{GM\text{-}replication}$=0.81, SE=0.16). Pulse rate correlated with waist circumference (r_{GM}=0.67, SE=0.13) and basal metabolic rate

(r_{GM}=-0.55, SE=0.11), acceleration force correlated with frequency of vigorous activity (r_{GM}=-0.64, SE=0.17) while hand grip strength (left and right) was associated with forced vital capacity (replicated) and forced expiratory volume (replicated). In addition, we found significant grey matter correlations between substance use phenotypes such as amount of usual alcohol intake and alcohol intake (r_{GM}=-0.89, SE=0.086; $r_{GM\text{-}replication}$=-1.0, SE=0.12; sign due to coding of the variable, **see Dataset S1**), smoking status (r_{GM}=0.71, SE=0.13) and past tobacco use (r_{GM}=-0.64, SE=0.14). Finally, we identified unexpected large grey-matter correlations. For example, cheese intake and forced expiratory volume were both correlated (r_{GM}=1.0, SE=0.11) with fluid intelligence, and waist circumference correlated with overall health rating and pulse rate (r_{GM}>0.67). Overall, 9 out of the 26 significant correlations replicated in the UKB replication sample; sign of the grey-matter correlation was always consistent between discovery and replication analyses (**Table S1**).

In the HCP, we observed positive grey-matter correlations between cognition domains (**Figure 2** and **Dataset S6**) and between IQ dimensions and education level. In addition, the two tobacco related phenotypes were associated with most of the same grey-matter vertices (r_{GM}=0.92, SE=0.045). To note, residual correlations and grey-matter correlations were of opposite signs between IQ domains and delay discounting variables, and between cognition and substance use phenotypes. These observations remained after rank-inverse transformation of the variable, suggesting it is not an artefact of the trait distribution. More work is needed to confirm these results in larger samples.

For completeness, grey-matter (and residual) correlations under the baseline model are reported in **Figure S7**, which reveals many large grey-matter correlations between measures of body size and diet, blood assay, activity levels and depression symptoms and score. This further highlights that in the phenome, the brain-morphometricity of some traits

may be accounted for by the covariation between these phenotypes and body size measurements. In particular, depression score was correlated (r_{GM}=1) with weight, BMI waist or hip circumference, consistent with its brain-morphometricity lowered to 0 when controlling for body size. In addition, depression score was also correlated (r_{GM}=-1) with activity levels and acceleration force, but also with poultry or cheese intake, happiness in own's health, diabetes and time spent watching television (r_{GM}=1), variable themselves strongly associated with measures of body shape (**Figure S7, section 3.2**).

3.4.Associations with grey-matter structure of specific cortical and subcortical regions

We investigated the brain-morphometricity of traits by estimating the association with grey-matter structure of specific cortical and subcortical regions (Desikan-Killiany atlas (Desikan et al., 2006)). All phenotypes were corrected for height, weight and BMI in addition to the baseline covariates. Associations with BMI and other body size variables under the baseline model are also presented. In this post-hoc analysis, we used Bonferroni correction to account for the number of tests performed (significance threshold of 0.05/(164*39)=7.2e-6 in the UKB, 1.2e-5 in the HCP).

In the UKB, the largest associations were observed between age of the participants and subcortical volumes (R^2 ranging between 0.22 and 0.35 for subcortical thickness, 0.20 0.38 for subcortical area), but most cortical regions were also significantly associated with age, albeit to a lesser extent (R^2 in the 0.0083-0.15 range for cortical thickness, 0.0048-0.15 range for cortical surface area). Next, significant ROI associations included sex, associated with all subcortical volumes (R^2 in the 0.0049-0.024 range for thickness, 0.0058-0.027 for area) and with many cortical regions (R^2 in the 0.0011-0.0076 range for cortical thickness, 0.0019-0.014 for cortical surface area) (**Figure S8** and **Dataset S7**). Maternal smoking around

564 birth was further associated with 28 ROI, mostly located in the occipital and temporal lobes
565 (R^2 in the 0.013-0.026 range with cortical thickness, R^2 in 0.014-0.071 with cortical surface
566 and R^2 in the 0.010-0.039 range with subcortical structure). In addition, we found significant
567 associations between cognition domains and structure of thalamus, putamen, pallidum and
568 hippocampus (R^2 in the 0.0043-0.024 range). Notably, fluid intelligence was associated with
569 all aspects of thalamus anatomy (left and right, thickness and surface area) while the other
570 cognition domains considered were associated with some aspects of thalamus structure. No
571 association between cognition and cortical structure survived multiple testing correction.
572 Diabetes diagnosis correlated with (left) superior frontal surface area (R^2=0.054), as
573 well as with thalamus, putamen, and pallidum thickness (R^2 ranging between 0.0067 and
574 0.015), or thalamus and hippocampus surface (R^2 in the 0.0061-0.014 range). Alcohol intake
575 was associated with left thalamus thickness (R^2=0.018) while smoking status and past
576 tobacco use were associated with thalamus, caudate, putamen and pallidum thickness, as
577 well as with thalamus surface area (R^2 in the 0.007-0.020 range). Finally, we also observed
578 small associations between cortical or subcortical regions and overall health rating, time
579 spend watching TV, body fat percentage and physiological measurements (**Figure S8**).
580 Using the replication UKB sample, we replicated 633 out of the 975 significant ROI-
581 trait associations ($p<0.05/975$). Most associations were found with age, sex and body size
582 variables, though we also replicated associations between subcortical volumes and hand
583 grip strength or time spent watching TV (**Dataset S8**). In addition, the magnitude of the
584 associations with age, and body size were greatly similar between discovery and replication
585 analyses (**Figure S9**). For sex, we observed larger ROI associations in the UKB replication
586 sample (**Figure S9**), consistent with the larger brain-morphometricity observed in this
587 sample (**Figure S6**).

In the HCP sample, age was associated with thickness (R^2 in the 0.020-0.049 range) and surface area (R^2 ranging between 0.067-0.10) throughout the cortex, as well as with subcortical structure (R^2 in the 0.016-0.087 range). Sex was associated with cortical thickness of the lateral orbitofrontal cortex (R^2 in the 0.059-0.073 range), as well as with subcortical structure (R^2 in the 0.042-0.19 range). In addition, we found large associations between cocaine, opiate or hallucinogens use and surface area of several cortical regions located in the temporal lobe (fusiform, superior temporal, insula), frontal (pars-triangularis, pars-opercularis, caudal-middle frontal), parietal (supramarginal, superior and inferior-parietal, precuneus) or in the cingulate (R^2 in the 0.25-1.00 range for cocaine test, R^2 in the 0.43-0.46 range for opiates, R^2 in 0.25-0.56 for number of times used hallucinogens). However, the small numbers and possible outliers in the vertex-wise measurements make such associations prone to false positives. Alcohol consumption was also associated with surface are of the frontal cortex (right rostral middle frontal, paracentral and precentral gyri, R^2 in the 0.28-0.36 range). No other association survived multiple testing correction (**Figure S10 and Dataset S9**).

Body size variables were strongly associated with subcortical structure under the baseline model (R^2 ranging between 0.010-0.059 for height, R^2 between 0.048-0.30 for the others) and to a lesser extent with cortical surface area (R^2 between 0.0078-0.026 for height, R^2 between 0.0061-0.060 for the others) and cortical thickness (R^2 in 0.0039 0.016 for height, R^2 in 0.0017 0.045 for the others). The associations between grey-matter structure and body size were pervasive (72/164 significant ROIs associations with height, 109 with waist circumference, 105 with BMI) (**Figure S11, Dataset S10**), suggesting that when acting as confounders height, weight or BMI could lead to false positives in many brain regions.

612

613 *3.5.Ten-fold cross-validation in the UKB and prediction into the UKB replication sample*

614 For each UKB participant, we calculated grey-matter scores relative to phenotypes
615 showing significant brain-morphometricity, by estimating the marginal association between
616 each vertex and the trait of interest. As in previous sections, we used height, weight and
617 BMI controlled for baseline covariates; and further regressed out body size for all other
618 phenotypes. We evaluated the prediction accuracy of the grey-matter BLUP scores by
619 computing their correlations with the observed values (10-fold cross validation design).

620 Most grey-matter scores significantly correlated (positively) with their corresponding
621 phenotypes (significance threshold of 0.05/39=1.2e-3, **Table 1, S3, Figure 3**). Albeit
622 significant, prediction accuracy was overall low (typically r<0.10, including r=0.11 for sex,
623 r<0.09 with cognition, r=0.08 for alcohol intake, r=0.06 with smoking status) except for age
624 (r=0.60), and maternal smoking around birth (r=0.26) whose grey-matter score correlated
625 more strongly with the observed values. We found similar prediction results in the UKB
626 replication sample, with 29 associations reaching significance at p<1.2e-3 (**Table 1, S3**).
627 Prediction accuracy was on par for most traits, though greater in the replication sample for
628 age and sex (**Figure 3, Table 1, S3**), consistent with a larger training sample being used and
629 larger morphomometricity observed in the replication set (**Figure S6**).

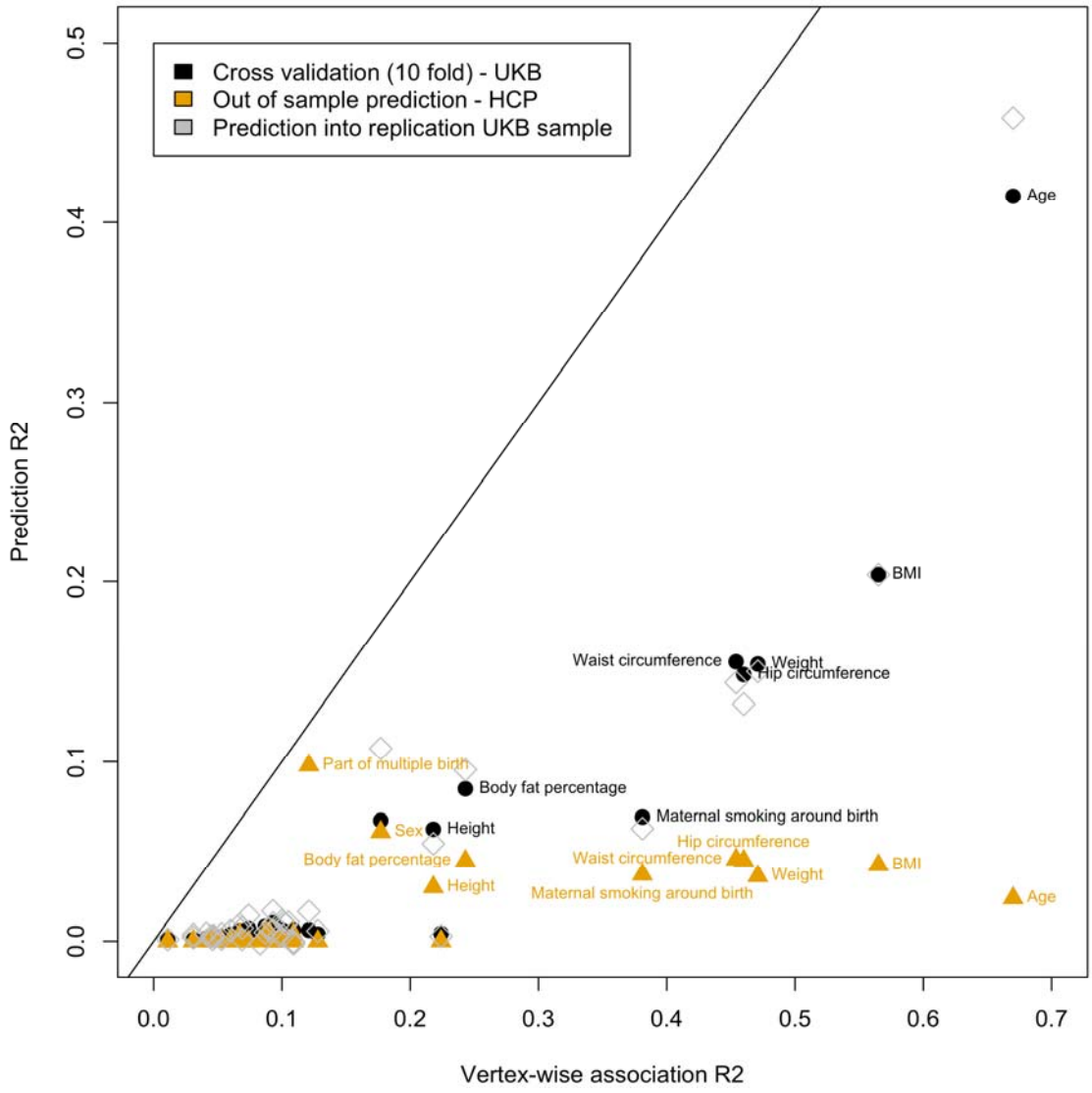
630 When not correcting for body size, 56/58 BLUP scores significantly correlated with
631 the observed values in the 10-fold cross validation and 42 associations replicated using the
632 UKB replication sample (p<0.05/58, See **FigureS12** and **DatasetS11**). Predicted age
633 correlated with chronological age (r=0.72 in the discovery, r=0.70 in the replication), while
634 predicted sex also strongly associated with the observed value (AUC of 0.90 and 0.89). Grey-
635 matter scores of body shape (under the baseline covariates) were also significantly

636 correlated with the observed values (r=0.25 for height, r=0.29 for body fat percentage,
637 r=0.39 for weight and hip or waist circumference, r=0.45 for BMI). Finally, grey-matter
638 scores of BMI correlated positively with depression symptom count (r=0.10, p-value<1e-14),
639 as expected from the brain-morphometricity of depression being limited the covariation
640 with body size. It even outperformed the grey-matter score built from the depression score
641 itself (r=0.05, p-value<1e.5).
642
643 *3.6.Out of sample prediction – application in the HCP sample*
644 Out of sample prediction validates that the morphometric associations are
645 generalizable to independent brain images, beyond population and scanner differences. We
646 trained our prediction models on the UKB discovery cohort and calculated grey matter
647 scores for each HCP participant. We tested the association between predicted value (brain
648 scores) and the observed phenotype in the HCP. For traits only available in the UKB (e.g.
649 waist circumference) we used a proxy in the HCP (e.g. BMI).
650 Grey matter scores for age, sex, and being a twin significantly correlated with the
651 observed values (r_{age}=0.15, r_{sex}=0.25, $r_{twin\text{-}status}$=0.31, p-value significant after multiple testing
652 correction) (**Table 1, S3** and **Figure 3**). Grey-matter score for maternal smoking around birth
653 correlated with smoking status (r=0.19). None of the other grey-matter scores significantly
654 correlated with a similar HCP variable.
655 Without correcting for body size, 19 BLUP scores correlated to corresponding
656 variables (**Dataset S11, Figure S12**). For example, scores for BMI, body fat percentage, hip or
657 waist circumference also correlated positively with BMI (r=0.21, p-value<1.2e-3), while
658 scores for height and weight also correlated with the observed phenotypes (r_{Height}=0.17,

659 r_{Weight}=0.19). Finally, scores build from diet items or quantifying activity levels significantly

660 predicted BMI in the HCP.

661



662

663 **Figure 3: In sample and out of sample prediction accuracy as a function of the total**

664 **association R^2.**

665 Labels highlight some of the significant prediction having the greatest accuracy. As

666 predicted by the theory, the prediction accuracy is capped by the total association R^2 (points

667 below the diagonal). In addition, out of sample prediction results in a lower prediction

accuracy than in-sample prediction. We hypothesise that the low prediction accuracy of age in the HCP is due to the much younger age range of the HCP participants, compared to the UKB). Participants born from multiple pregnancy appear better identified (predicted) in the HCP than within the UKB sample, which is due to a greater proportion of females and twins in the HCP compared to the UKB, as well as greater morphometricity in the HCP. Such mechanism has been discussed in the field of genetic and solutions exist to correct results for differences in prevalence between samples (Lee et al., 2012). We reported the AUC in **Table 1** (for discrete variables) as it is independent of the proportion of twins and males, thus differences in AUC likely reflect differences in morphometricity between the UKB and HCP samples.

678 **Table 1:** Summary of the prediction accuracy (R^2) of the BLUP grey-matter scores. We constructed BLUP scores for the 39 UKB variables
679 showing significant morphometricity and evaluated their predictive power in the UKB (10 fold-cross validation) and HCP sample. When the
680 phenotype corresponding to the grey-matter score was not available in the HCP, we chose the closest available (e.g. waist circumference grey-
681 matter score evaluated against BMI). We evaluate the prediction accuracy by fitting GLM controlling for height, weight and BMI as well as for
682 the baseline covariates (acquisition, age, sex and head size); except for (#) denoting associations not controlling for height, weight and BMI.
683 Rows in bold indicate significant association after correcting for multiple testing (p<0.05/39=1.3e-3) both in and out of sample. This reduced
684 table only shows prediction results significant in all 3 scenarios, see **Table S3** for full table of results.

	In sample prediction (UKB)				Prediction into UKB replication				Out of sample prediction (HCP)				
	r	**pvalue**	**R^2**	**AUC (SE)**	**r**	**pvalue**	**R^2**	**AUC (SE)**	**HCP variable predicted**	**r**	**pvalue**	**R^2**	**AUC (SE)**
Age	0.64	0.0e+00	0.41		0.68	0.0e+00	0.46		Age	0.15	3.1e-08	0.024	
Sex	0.26	0.0e+00	0.067	0.58 (0.0059)	0.33	9.8e-305	0.11	0.8 (0.0064)	Sex	-0.25	8.0e-42	0.061	0.68 (0.016)
Part of multiple birth	0.078	4.1e-14	0.0061	0.66 (0.022)	0.13	1.5e-03	0.016	0.72 (0.065)	Being a twin	0.31	1.1e-28	0.098	0.69 (0.016)
Body fat percentage#	0.29	0.0e+00	0.085		0.31	7.7e-190	0.095		BMI	0.21	5.6e-13	0.045	
Waist circumference#	0.39	0.0e+00	0.16		0.38	2.0e-205	0.14		BMI	0.21	3.5e-13	0.046	
BMI#	0.45	0.0e+00	0.2		0.45	7.4e-235	0.20		BMI	0.21	2.4e-12	0.042	
Hip circumference#	0.38	0.0e+00	0.15		0.36	7.3e-143	0.13		BMI	0.21	5.2e-13	0.045	
Height#	0.25	6.5e-	0.062		0.23	2.6e-132	0.054		Height	0.17	1.8e-	0.03	

		318									17		
Weight#	0.39	0.0e+00	0.15		0.39	5.8e-231	0.15		Weight	0.19	1.2e-12	0.036	
Maternal smoking around birth	0.26	9.8e-132	0.069	0.66 (0.0067)	0.25	1.7e-08	0.063	0.65 (0.027)	FTND score	0.19	8.9e-04	0.037	

685

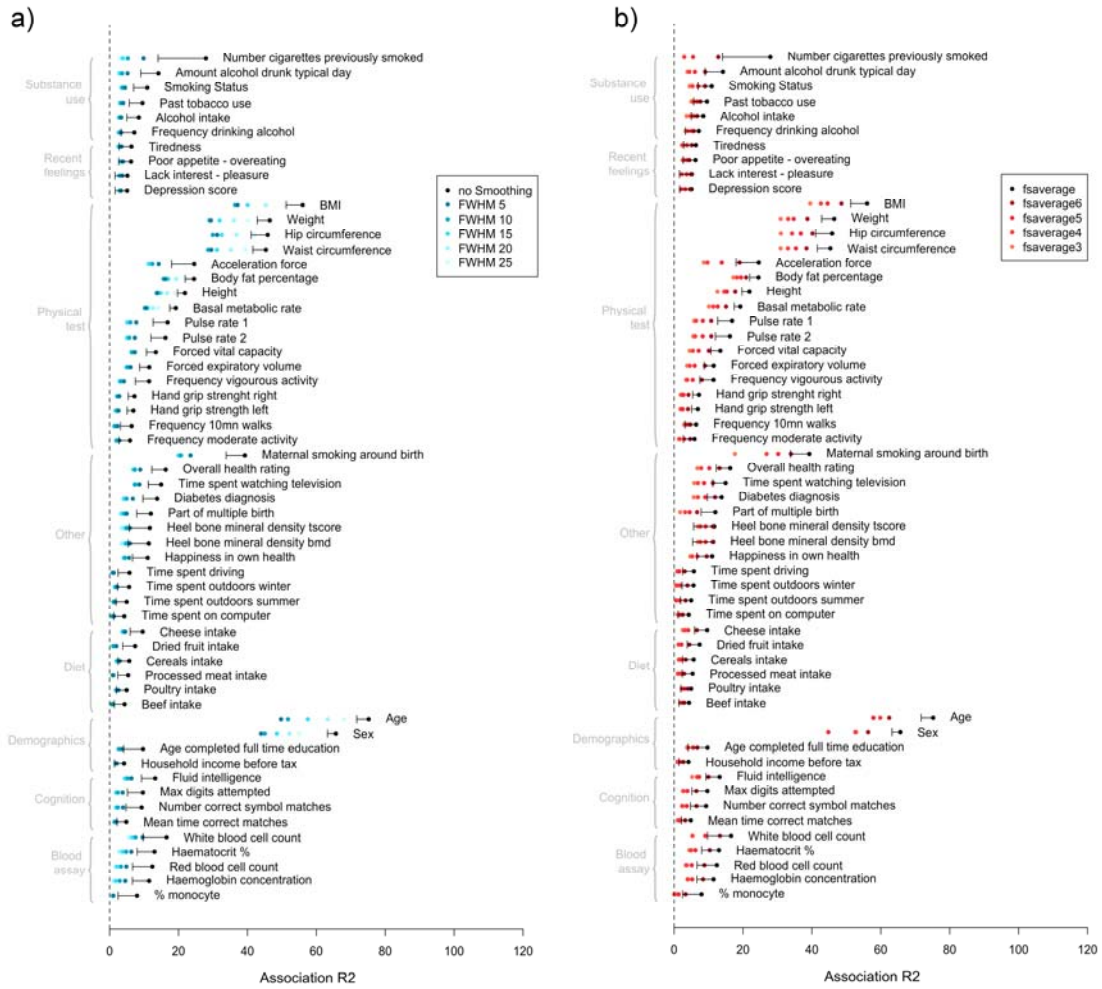
686

687

3.7.Best cortical processing

We compared the brain-morphometricity estimates obtained by varying the cortical processing options: smoothing of the cortical meshes and applying coarser meshes available in FreeSurfer (see **2.10**). We performed this analysis on the UKB discovery and replication samples as the large SE of the estimates in HCP would limit the interpretation of the results. We used baseline covariates as in **Figure 1**. We found that applying smoothing (5-25mm) or reducing the cortical mesh complexity always led to a lower point estimate of brain morphometricity in the UKB discovery (**Figure 4**) and replication (**SFigure 13, Datasets S12-13 for full tables)** samples. As such, the fsaverage cortical mesh with no smoothing can be deemed "best" processing for all phenotypes considered.

In addition, we compared results from **Figure 1** to those from Region-Of-Interest (ROI) based processing (taking the average of each cortical or subcortical region, here ENIGMA processing). We found that the vertex-wise approach always yielded greater association R^2, thus retained more information than a ROI based dimension reduction (**Figure S14**).



704

705 **Figure 4: Comparison of brain-morphometricity estimates varying cortical processing**

706 **options in FreeSurfer.**

707 The reduction of brain-morphometricity as a function of mesh smoothing is presented on

708 the left panel (a), while the right panel (b) shows the effect of reducing the cortical mesh

709 complexity. The black bar indicates the lower bound of the 95% confidence interval of the

710 fsaverage-no smoothing estimate (identical to results presented in **Figure1**). Note that

711 brain-morphometricity estimates below the 95%CI lower bound cannot be deemed

712 significantly lower. Rather the 95%CI are presented for context and to remind that all

713 estimate from **Figure 1** do not have the same SE.

714

4. Discussion

We report the associations between vertex-wise measurements of grey-matter structure and a large set of phenotypes capturing aspects of demographics, physical capacities, substance use, psychiatry, lifestyle and stress/traumas **(Figure 1)**. In addition, we introduced the concept of between-trait grey-matter correlation (**Figure 2, Figure S7**) that quantifies the proportion of brain markers shared between two traits. We demonstrated the versatility of our vertex-wise LMM approach by identifying specific cortical and subcortical regions (**Figure S8, S10, S11**) associated with the phenotypes of interest. Finally, we derived BLUP (Best Linear Unbiased Predictor) grey-matter scores and demonstrated their significant predictive abilities in the UKB discovery sample (10-fold cross validation) and in completely independent samples (HCP, and replication UKB, **Table 2, S3, Figure 3**).

Our vertex-wise analyses retained the complexity of the cortical ribbon and subcortical structure, leading to larger associations compared to the standard ROI based data reduction (**Figure S14**). Similarly, reducing the cortical complexity via local averages (smoothing) or halving the number of vertices also led to reduced brain-morphometricity estimates for all phenotypes considered (**Figure 4, S13**). These results indicate that grey-matter scores from "fsaverage-no smoothing" cortical measurements can achieve greater level of prediction, but may require larger training samples to counterbalance their increased complexity (Dudbridge, 2013).

In both the UKB and HCP samples, the largest brain-morphometricity (**Figure 1**) were found between grey-matter structure and age, sex and with measures of body size (height, weight, BMI, body fat percentage, waist or hip circumference). In our post-hoc LMM-ROI analysis, we found those phenotypes to be associated with most cortical and subcortical regions (**Figure S8-11**). Our results for sex are consistent with results from the

740 UKB first release (N=5,216, using ROI average (Ritchie et al., 2018)), while several studies
741 have previously reported associations between BMI and several grey-matter measurements
742 (Cole et al., 2013; Gupta et al., 2015; Kurth et al., 2013; Masouleh et al., 2016; Medic et al.,
743 2016; Opel et al., 2017). Despite such large and widespread pattern of association between
744 body-size and grey-matter we did not observe significant brain-morphometricity for (self-
745 reported) anorexia, bulimia or binge eating though the small numbers (<30 cases in the
746 UKB, **Dataset S1**) limit the interpretability of the results.

747 We observed moderate to small associations (R^2<0.4) between grey-matter and
748 substance use (tobacco and alcohol), maternal smoking around birth, blood assay results,
749 education and income level, diet, depression score and symptoms, twin-status as well as
750 cognition domains (**Figure 1**). The latter replicated and expanded the result of an analogous
751 analysis on an early release of the HCP (N=150) and the ADNI dataset (Sabuncu et al., 2016).
752 We note that handedness was only weakly (R^2_{UKB}=0.04, R^2_{HCP}<0.001, not significant)
753 associated with cortical or subcortical grey-matter coherent with the conflicting results
754 reviewed in (Jin Kang et al., 2017). Our results indicate that individuals that display similar
755 grey-matter structure tend to also be similar in term of age, sex, body size, cognition,
756 activity levels, substance use and lifestyle. We did not detect significant association
757 between grey-matter morphometry and psychiatric diagnoses (lifetime self-reported), sleep
758 phenotypes or lifetime stress/traumas **(Dataset S3)** despite previous morphometricity
759 reports from case-control samples of autism, schizophrenia and ADHD (Sabuncu et al.,
760 2016).

761 When controlling for height, weight and BMI in the analyses, many of the
762 associations became non-significant: such as those between grey-matter and diet, activity
763 levels or depression score/symptoms (R^2<0.04; **Dataset S3**). Furthermore, we did not detect

any significant association between grey-matter structure and other depression related phenotypes (e.g. self-reported diagnosis by a doctor, MDD case-control status as used by the Psychiatric Genetic Consortium (Wray et al., 2018), and neuroticism; **Dataset S3**). Our findings shed a new light on previously published results, as even the largest case-control international initiatives (e.g. ENIGMA-MDD (Schmaal et al., 2016a; Schmaal et al., 2016b)) may reflect, at least in part, variance shared between depression and BMI (such as the causal effect of BMI on depression(Wray et al., 2018)). Understanding the relationship between brain and depression may call to analyse brain regions or features not extracted in the current processing (e.g. brain stem and cerebellum) or features collected from another type of images (e.g. Diffusion Weighted Images (DWI), fMRI).

To summarise, body size is associated with large, widespread variations of grey-matter structure (**Figure 1, Figure S11**) and more work is needed to understand its contribution to published results linking grey-matter anatomy to psychiatric disorders (MDD, bipolar, schizophrenia and substance use are associated with BMI (Luppino et al., 2010; McElroy and Keck, 2012; Rajan and Menon, 2017; Saarni et al., 2009; Wray et al., 2018)) or sexually dimorphic traits (likely associated with height and weight). In addition, body size may be differently associated with the phenome across countries or age groups, which may limit the replication of findings and predictive abilities of body size dependent scores. Note that the possible confounding effects of body size are exacerbated in small case-control samples, leading to increased chances of false positive associations (Button et al., 2013; Ioannidis, 2005). Body size being associated to many brain regions (**Figure S11**), such confounding effect could lead to widespread cortical or subcortical false positives.

In subsequent association and prediction analyses, we made a conservative choice to correct for height, weight and BMI. This meant that we likely reported conservative

788 estimates of brain-morphometricity and fewer significant grey-matter correlations,
789 predictive grey-matter scores or trait-ROI associations (see **Figures S7, S12** and **Dataset S3**
790 for uncorrected results). The large covariation of body-size with the phenome (at least with
791 the variables we selected) is still of interest but it may be more powerful to study directly
792 BMI for example. This is exemplified by the greater prediction accuracy achieved by a BMI
793 grey-matter score (vs. depression specific score) when predicting depression score. Such
794 behaviour can be anticipated based on the large r_{GM} between BMI and depression score
795 (**Figure S7**), combined to the larger brain-morphometricity of BMI (Dudbridge, 2013).
796 Finally, our conservative approach should remind us to be careful when interpreting
797 associations. For example, one should not conclude about actionable links between diet and
798 depression based on the large and significant grey-matter correlations, as it might be
799 mediated by body size. Though, directionality of the associations will need to be established
800 to conclude in this case.

801 We estimated between-trait grey-matter correlation (**Figure 2, Dataset S5, S6**) that
802 quantifies the proportion of brain markers shared between two traits and found significant
803 relationships between cognition domains, between tobacco and alcohol consumption or
804 between measures of fitness. Large grey-matter correlations between seemingly unrelated
805 traits (e.g. fluid IQ and cheese intake [replicated], waist circumference and pulse rate or
806 overall health rating) raise questions about the nature of the relationships between those
807 variables (causality, true positive or confounded association?). Note that r_{GM} would also
808 capture correlated measurement errors between traits, for example due to head motion or
809 other sources of noise in MRI acquisition.

810 We further characterised the brain-morphometricity by identifying specific cortical
811 and subcortical regions (ROI) associated with our phenotypes (**Figure S8-S11**). In the UKB,

812 smoking status was associated with thickness and surface of the thalamus (left and right),
813 although we also found associations with the caudate and pallidum. Previous studies have
814 reported association between tobacco usage and volume of left thalamus (Gallinat et al.,
815 2006; Gillespie et al., 2018; Hanlon et al., 2016), which might be due to faster age related
816 volume loss in smokers (Durazzo et al., 2017). We did not replicate other cortical or
817 subcortical associations previously reported (Gallinat et al., 2006; Hanlon et al., 2016; Prom-
818 Wormley et al., 2015). Alcohol intake was also associated with left thalamus thickness in the
819 UKB, consistent with the significant grey-matter correlation (**Figure 2**) between the two
820 traits. The thalamus has been implicated in alcohol-related neurological complications (e.g.
821 Korsakoff's syndrome)(Pitel et al., 2015) but may also be associated with regular alcohol
822 usage (Cardenas et al., 2007; Pitel et al., 2015) or alcohol use disorder (van Holst et al.,
823 2012). Maternal smoking around birth was further associated with the thalamus, putamen,
824 hippocampus and pallidum, as well as temporal and occipital ROIs. In addition, diagnosis of
825 diabetes was associated with area of the left superior-frontal cortex (**Dataset S8, Figure S8**).
826 Nervous system complications of diabetes (sometimes labelled diabetic encephalopathy)
827 are widely accepted (Mijnhout et al., 2006) but little is known about the specific brain
828 regions associated with the condition (Moheet et al., 2015).
829 Finally, we derived and evaluated BLUP (Best Linear Unbiased Predictor) grey-matter
830 scores for each individual and the 39 phenotypes showing brain-morphometricity in the UKB
831 (after correcting for body size). The prediction accuracy above what is expected by chance
832 confirmed that the traits associations with grey-matter structure are transferable to
833 independent samples and even samples imaged on a different scanner with different
834 demographics (e.g. HCP, **Table 2, S3, Figure 3**). Overall, the prediction accuracy was below a
835 few percent (of variance) except for age, sex, being a twin, maternal smoking at birth and

836 body size measurements (**Table 2, Figure 3**). Grey-matter score for maternal smoking
837 around birth predicted FTND score in the HCP sample suggesting that passive and active
838 smoking may be associated with similar grey-matter morphology. Our ability to predict (in
839 part) the twin status of participants (**Table 2**) suggests that twins' grey-matter structure may
840 be more similar than average even if the twins are not from the same family.
841 Other methods allow to derive prediction from a large number of brain features (e.g.
842 penalised regression, or deep learning) though direct comparison with prediction accuracy
843 from previous publications is limited by the use of different samples, MRI scanners,
844 processing options, input data and prediction algorithm. To note, BLUP is computationally
845 efficient as it does not require estimation of hyper-parameters (as in penalised regression).
846 Similar to polygenic risk scores (Dudbridge, 2013), the prediction R^2 of grey-matter BLUP
847 scores increases with the training sample size and is capped by the association R^2 with all
848 vertices (**Figure 3**). Future application of the grey-matter scores include studying correlates
849 of brain age (Cole, 2017; Cole et al., 2017; Liem et al., 2017), body size and substance use,
850 especially in samples where this information was not collected.
851 To note, most of the results observed in the UKB discovery sample (brain-
852 morphometricity, r_{GM}, ROI based associations) replicated in an independent UKB sample
853 (replication). On the other hand, the UKB and HCP samples differed in term of data
854 collected, age range**,** country of origin, MRI acquisition, processing and participants'
855 recruitment, which might explain some of the differences in results (brain-morphometricity
856 of cognition for example).
857 In the UKB, we chose to add the T2 FLAIR (when available) to improve pial
858 reconstruction in the FreeSurfer processing, though the effect of such option and the
859 possible differences with T1w only processing is not well described in the literature. We

observed a large difference in total cortical thickness between participants processed either way (**Figure S2**). This warrants further investigation though it is unlikely to have impacted the results presented here. Indeed, our QC step excluded more than 80% of the 400-odd participants processed using T1w only, likely because they showed outlying brains compared to the T1w+T2 FLAIR processing. In addition, availability of T2 FLAIR was not associated with any of the phenotypes. Finally, the replication of the UKB associations and the out of sample prediction suggest that our results are robust to the presence of these few outliers.

The HCP comprises many twin pairs (thus, non-independent observations), though we modelled the grey-matter relatedness in all analyses, which should account for the grey-matter resemblances arising from shared genetics or environment. A bias due to twins is unlikely as our results on the full HCP sample yielded always similar (e.g. Fluid IQ) or lower (e.g. attention) brain-morphometricity estimates than reported by Sabuncu et al., who selected 1 subject per family (Sabuncu et al., 2016). Finally, the grey-matter similarity of twins was greater than average but in line with the similarity seen between unrelated individuals (**Appendix S2**), which led us not to exclude twin pairs from the analyses (contrary to what is seen/done in genetics).

Due to recruitment choices, the UKB and HCP samples do not contain many psychiatric cases (outside of the highly prevalent MDD, see (Fry et al., 2017) on the healthy volunteer bias in the UKB) and cannot replace the large case-control initiatives (e.g. ENIGMA disease groups). Despite using some of the largest imaging samples available, our ROI based and grey-matter score analyses suffered from limited statistical power, though more data is currently being collected by the UKB.

To complement our analyses (limited to young and older adults), more work is required to understand the relationship between grey-matter morphology and the phenome during development (e.g. in children or adolescents) as well as in specific age/disease groups (Rosenberg et al., 2018), or using different scanners or processing options (e.g. 1.5 Tesla MRI, scanning time, FSL or SPM processing (Flandin and Friston, 2008; Jenkinson et al., 2012)). Note that all associations reported here must be interpreted carefully as they may be causes or consequences of the disorder or trait, or a result of the pervasive pleiotropy underlying human complex phenotypes (Solovieff et al., 2013). Future application or LMM include determining the best MRI image processing for a trait (i.e. the processing options that maximise the association R^2; e.g. **Figure 4**) by extending our analysis to other measures of grey-matter structure (e.g. voxel-based morphometry (Wright et al., 1995)).

We have released the scripts used in image processing and LMM analyses to facilitate replication and dissemination of the results (see **URLs**). We have also released BLUP weights to allow meta-analyses or application of the grey-matter scores in independent cohorts.

5. URLs

Summary-level data (BLUP weights) and vertex membership in the Desikan atlas: http://cnsgenomics.com/data.html ; OSCA: http://cnsgenomics.com/software/osca/ ; ENIGMA protocols: http://enigma.ini.usc.edu/protocols/imaging-protocols/ ;

6. Author contributions

PMV, NRW, JY and BCD designed the analyses. FZ and JY developed the OSCA software. YH and BCD created the plots. KK, LY and ZZ assisted BCD with the UKB phenotypic

908 and genetic data, including download, formatting and curation. LS downloaded and
909 processed the HCP MRI images under MJW supervision. BCD downloaded and processed the
910 UKB MRI images. BCD performed the analyses and wrote the manuscript. All the authors
911 reviewed the manuscript.
912

913 **7. Acknowledgements**

914 This research was supported by the Australian National Health and Medical Research
915 Council (1078037, 1078901, 1113400, 1161356 and 1107258), the Australian Research
916 Council (FT180100186 and FL180100072), and the Sylvia & Charles Viertel Charitable
917 Foundation.

918 Informed consent was obtained from all UK Biobank participants. Procedures are
919 controlled by a dedicated Ethics and Guidance Council
920 (http://www.ukbiobank.ac.uk/ethics), with the Ethics and Governance Framework available
921 at http://www.ukbiobank.ac.uk/wp-content/uploads/2011/05/EGF20082.pdf. IRB approval
922 was also obtained from the North West Multi-centre Research Ethics Committee. This
923 research has been conducted using the UK Biobank Resource under Application Number
924 12505.

925 Informed consent was obtained from all HCP participants. HCP Data were provided
926 by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David
927 Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centres
928 that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Centre for
929 Systems Neuroscience at Washington University.

930 We used R(R Development Core Team, 2012) (v3.3.3) for analyses not performed
931 using OSCA (Zhang et al., 2019) and for plots. We used the colour-blind friendly R palette

(http://jfly.iam.u-tokyo.ac.jp/color/), *qqman*(Turner, 2014) for QQ-plots, *ggplot2*(Wickham, 2009) and *ggsignif*(Ahlmann-Eltze, 2017) for circular bar plots, *corrplot(Wei and Simko, 2017)* for correlation matrix plots, *ukbtools(Hanscombe, 2017)* to facilitate UKB phenotype manipulation. Other packages used to assist analyses and data handling include *FactoMineR*(Husson et al., 2015; Husson et al., 2009), *Hmisc(Harrell, 2017), rowr(Varrichio, 2016), pwr(Champely, 2017), XML(Temple and the CRAN Team, 2017), tidyverse(Wickham, 2017a), dplyr(Wickham and Francois, 2015), readr(Wickham, 2017b), reshape2(Wickham, 2007) and rmarkdown(Allaire, 2018).*

We would like to thank Allan McRae, the Institute of Molecular Bioscience (IMB) and the Research Computing Centre (RCC) IT teams at the University of Queensland for their support with high performance computing, data handling, storage and processing.

8. Competing financial Interests statement

The authors declare no conflict of interests.

References

Achenbach, T.M., 2009. Achenbach system of empirically based assessment (ASEBA): Development, findings, theory, and applications. University of Vermont, Research Center of Children, Youth & Families.

Achenbach, T.M., Dumenci, L., Rescorla, L.A., 2003. Ratings of Relations Between DSM-IV Diagnostic Categories and Items of the Adult Self-Report (ASR) and Adult Behavior Checklist (ABCL).

Ahlmann-Eltze, C., 2017. ggsignif: Significance Bars for 'ggplot2'.

Allaire, J.X., Yihui.; McPherson, Jonathan.; Luraschi, Javier.; Ushey, Kevin.; Atkins, Aron.; Wickham, Hadley.; Cheng, Joe.; Chang, Winston. , 2018. rmarkdown: Dynamic Documents for R.

Bates, D., Machler, M., Bolker, B., Walker, S., 2015. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software 67, 1-48.

Bijma, P., Bastiaansen, J.W., 2014. Standard error of the genetic correlation: how much data do we need to estimate a purebred-crossbred genetic correlation? Genetics Selection Evolution 46, 79.

Button, K.S., Ioannidis, J.P.A., Mokrysz, C., Nosek, B.A., Flint, J., Robinson, E.S.J., Munafo, M.R., 2013. Power failure: why small sample size undermines the reliability of neuroscience. Nature Reviews Neuroscience 14, 365-376.

Buysse, D.J., Reynolds, C.F., 3rd, Monk, T.H., Berman, S.R., Kupfer, D.J., 1989. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 28, 193-213.

Cardenas, V.A., Studholme, C., Gazdzinski, S., Durazzo, T.C., Meyerhoff, D.J., 2007. Deformation-based morphometry of brain changes in alcohol dependence and abstinence. Neuroimage 34, 879-887.

Champely, S., 2017. pwr: Basic Functions for Power Analysis.

Cole, J.H., 2017. Neuroimaging-derived brain-age: an ageing biomarker? Aging (Albany NY) 9, 1861-1862.

Cole, J.H., Boyle, C.P., Simmons, A., Cohen-Woods, S., Rivera, M., McGuffin, P., Thompson, P.M., Fu, C.H., 2013. Body mass index, but not FTO genotype or major depressive disorder, influences brain structure. Neuroscience 252, 109-117.

Cole, J.H., Poudel, R.P.K., Tsagkrasoulis, D., Caan, M.W.A., Steves, C., Spector, T.D., Montana, G., 2017. Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. Neuroimage 163, 115-124.

Crainiceanu, C.M., Ruppert, D., 2004. Likelihood ratio tests in linear mixed models with one variance component. Journal of the Royal Statistical Society Series B-Statistical Methodology 66, 165-185.

Desikan, R.S., Segonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., Killiany, R.J., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 31, 968-980.

Dudbridge, F., 2013. Power and predictive accuracy of polygenic risk scores. PLoS Genet 9, e1003348.

Durazzo, T.C., Meyerhoff, D.J., Yoder, K.K., Murray, D.E., 2017. Cigarette smoking is associated with amplified age-related volume loss in subcortical brain regions. Drug and Alcohol Dependence 177, 228-236.

993 Edens, E.L., Glowinski, A.L., Pergadia, M.L., Lessov-Schlaggar, C.N., Bucholz, K.K., 2010.
994 Nicotine Addiction in Light Smoking African American Mothers. Journal of Addiction
995 Medicine 4, 55-60.
996 Fischl, B., 2012. FreeSurfer. Neuroimage 62, 774-781.
997 Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Segonne, F., Salat, D.H., Busa, E.,
998 Seidman, L.J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., Dale, A.M.,
999 2004. Automatically parcellating the human cerebral cortex. Cerebral Cortex 14, 11-22.
1000 Flandin, G., Friston, K.J., 2008. Statistical parametric mapping (SPM). Scholarpedia 3(4),
1001 6232.
1002 Fry, A., Littlejohns, T.J., Sudlow, C., Doherty, N., Adamska, L., Sprosen, T., Collins, R., Allen,
1003 N.E., 2017. Comparison of Sociodemographic and Health-Related Characteristics of UK
1004 Biobank Participants With Those of the General Population. American Journal of
1005 Epidemiology 186, 1026-1034.
1006 Gallinat, J., Meisenzahl, E., Jacobsen, L.K., Kalus, P., Bierbrauer, J., Kienast, T., Witthaus, H.,
1007 Leopold, K., Seifert, F., Schubert, F., Staedtgen, M., 2006. Smoking and structural brain
1008 deficits: a volumetric MR investigation. European Journal of Neuroscience 24, 1744-1750.
1009 Gillespie, N.A., Neale, M.C., Bates, T.C., Eyler, L.T., Fennema-Notestine, C., Vassileva, J.,
1010 Lyons, M.J., Prom-Wormley, E.C., McMahon, K.L., Thompson, P.M., Zubicaray, G., Hickie,
1011 I.B., McGrath, J.J., Strike, L.T., Rentería, M.E., Panizzon, M.S., Martin, N.G., Franz, C.E.,
1012 Kremen, W.S., Wright, M.J., 2018. Testing associations between cannabis use and
1013 subcortical volumes in two large population-based samples. Addiction 0.
1014 Glasser, M.F., Sotiropoulos, S.N., Wilson, J.A., Coalson, T.S., Fischl, B., Andersson, J.L., Xu, J.,
1015 Jbabdi, S., Webster, M., Polimeni, J.R., Van Essen, D.C., Jenkinson, M., Consortium, W.U.-
1016 M.H., 2013. The minimal preprocessing pipelines for the Human Connectome Project.
1017 Neuroimage 80, 105-124.
1018 Goddard, M.E., Wray, N.R., Verbyla, K., Visscher, P.M., 2009. Estimating Effects and Making
1019 Predictions from Genome-Wide Marker Data. Statistical Science 24, 517-529.
1020 Gupta, A., Mayer, E.A., Sanmiguel, C.P., Van Horn, J.D., Woodworth, D., Ellingson, B.M.,
1021 Fling, C., Love, A., Tillisch, K., Labus, J.S., 2015. Patterns of brain structural connectivity
1022 differentiate normal weight from overweight subjects. Neuroimage-Clinical 7, 506-517.
1023 Gutman, B.A., Madsen, S.K., Toga, A.W., Thompson, P.M., 2013. A Family of Fast Spherical
1024 Registration Algorithms for Cortical Shapes. In: Shen, L., Liu, T., Yap, P.-T., Huang, H., Shen,
1025 D., Westin, C.-F. (Eds.), Multimodal Brain Image Analysis: Third International Workshop,
1026 MBIA 2013, Held in Conjunction with MICCAI 2013, Nagoya, Japan, September 22, 2013,
1027 Proceedings. Springer International Publishing, Cham, pp. 246-257.
1028 Gutman, B.A., Wang, Y.L., Rajagopalan, P., Toga, A.W., Thompson, P.M., 2012. Shape
1029 Matching with Medial Curves and 1-D Group-Wise Registration. 2012 9th Ieee International
1030 Symposium on Biomedical Imaging (Isbi), 716-719.
1031 Hanlon, C.A., Owens, M.M., Joseph, J.E., Zhu, X., George, M.S., Brady, K.T., Hartwell, K.J.,
1032 2016. Lower subcortical gray matter volume in both younger smokers and established
1033 smokers relative to non-smokers. Addiction Biology 21, 185-195.
1034 Hanscombe, K., 2017. ukbtools: Manipulate and Explore UK Biobank Data.
1035 Harrell, F.E.J., 2017. Hmisc: Harrell Miscellaneous.
1036 Henderson, C.R., 1950. Estimation of Genetic Parameters. Annals of Mathematical Statistics
1037 21, 309-310.

1038 Henderson, C.R., 1975. Best Linear Unbiased Estimation and Prediction under a Selection
1039 Model. Biometrics 31, 423-447.
1040 Husson, F., Josse, J., Le, S., Mazet, J., 2015. FactoMineR: Multivariate Exploratory Data
1041 Analysis and Data Mining.
1042 Husson, F., Josse, J., Pagès, L., 2009. FactoMineR, An R package dedicated to exploratory
1043 multivariate analysis.
1044 Ioannidis, J.P., 2005. Why most published research findings are false. PLoS Med 2, e124.
1045 Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the
1046 robust and accurate linear registration and motion correction of brain images. Neuroimage
1047 17, 825-841.
1048 Jenkinson, M., Beckmann, C.F., Behrens, T.E.J., Woolrich, M.W., Smith, S.M., 2012. FSL.
1049 Neuroimage 62, 782-790.
1050 Jin Kang, S., A Kang, K., Jang, H., Youn Lee, J., Il Lee, K., Seok Kwoen, M., Soo Kim, J., Park,
1051 K.M., 2017. Brain morphology according to age, sex, and handedness.
1052 Kurth, F., Levitt, J.G., Phillips, O.R., Luders, E., Woods, R.P., Mazziotta, J.C., Toga, A.W., Narr,
1053 K.L., 2013. Relationships between gray matter, body mass index, and waist circumference in
1054 healthy adults. Human Brain Mapping 34, 1737-1746.
1055 Lee, S.H., Yang, J., Goddard, M.E., Visscher, P.M., Wray, N.R., 2012. Estimation of pleiotropy
1056 between complex diseases using single-nucleotide polymorphism-derived genomic
1057 relationships and restricted maximum likelihood. Bioinformatics 28, 2540-2542.
1058 Liem, F., Varoquaux, G., Kynast, J., Beyer, F., Kharabian Masouleh, S., Huntenburg, J.M.,
1059 Lampe, L., Rahim, M., Abraham, A., Craddock, R.C., Riedel-Heller, S., Luck, T., Loeffler, M.,
1060 Schroeter, M.L., Witte, A.V., Villringer, A., Margulies, D.S., 2017. Predicting brain-age from
1061 multimodal imaging data captures cognitive impairment. Neuroimage 148, 179-188.
1062 Luppino, F.S., de Wit, L.M., Bouvy, P.F., Stijnen, T., Cuijpers, P., Penninx, B.W.J.H., Zitman,
1063 F.G., 2010. Overweight, Obesity, and Depression A Systematic Review and Meta-analysis of
1064 Longitudinal Studies. Archives of General Psychiatry 67, 220-229.
1065 Marcus, D.S., Harms, M.P., Snyder, A.Z., Jenkinson, M., Wilson, J.A., Glasser, M.F., Barch,
1066 D.M., Archie, K.A., Burgess, G.C., Ramaratnam, M., Hodge, M., Horton, W., Herrick, R.,
1067 Olsen, T., McKay, M., House, M., Hileman, M., Reid, E., Harwell, J., Coalson, T., Schindler, J.,
1068 Elam, J.S., Curtiss, S.W., Van Essen, D.C., Consortium, W.U.-M.H., 2013. Human Connectome
1069 Project informatics: quality control, database services, and data visualization. Neuroimage
1070 80, 202-219.
1071 Marcus, D.S., Harwell, J., Olsen, T., Hodge, M., Glasser, M.F., Prior, F., Jenkinson, M.,
1072 Laumann, T., Curtiss, S.W., Van Essen, D.C., 2011. Informatics and data mining tools and
1073 strategies for the human connectome project. Front Neuroinform 5, 4.
1074 Masouleh, S.K., Arelin, K., Horstmann, A., Lampe, L., Kipping, J.A., Luck, T., Riedel-Heller,
1075 S.G., Schroeter, M.L., Stumvoll, M., Villringer, A., Witte, A.V., 2016. Higher body mass index
1076 in older adults is associated with lower gray matter volume: implications for memory
1077 performance. Neurobiology of Aging 40, 1-10.
1078 McElroy, S.L., Keck, P.E., 2012. Obesity in Bipolar Disorder: An Overview. Current Psychiatry
1079 Reports 14, 650-658.
1080 Medic, N., Ziauddeen, H., Ersche, K.D., Farooqi, I.S., Bullmore, E.T., Nathan, P.J., Ronan, L.,
1081 Fletcher, P.C., 2016. Increased body mass index is associated with specific regional
1082 alterations in brain structure. International Journal of Obesity 40, 1177-1182.

1083 Mijnhout, G.S., Scheltens, P., Diamant, M., Biessels, G.J., Wessels, A.M., Simsek, S., Snoek,
1084 F.J., Heine, R.J., 2006. Diabetic encephalopathy: A concept in need of a definition.
1085 Diabetologia 49, 1447-1448.
1086 Miller, K.L., Alfaro-Almagro, F., Bangerter, N.K., Thomas, D.L., Yacoub, E., Xu, J., Bartsch, A.J.,
1087 Jbabdi, S., Sotiropoulos, S.N., Andersson, J.L., Griffanti, L., Douaud, G., Okell, T.W., Weale, P.,
1088 Dragonu, I., Garratt, S., Hudson, S., Collins, R., Jenkinson, M., Matthews, P.M., Smith, S.M.,
1089 2016. Multimodal population brain imaging in the UK Biobank prospective epidemiological
1090 study. Nature Neuroscience 19, 1523-1536.
1091 Moheet, A., Mangia, S., Seaquist, E.R., 2015. Impact of diabetes on cognitive function and
1092 brain structure. Ann N Y Acad Sci 1353, 60-71.
1093 Mugler, J.P., 3rd, Bao, S., Mulkern, R.V., Guttmann, C.R., Robertson, R.L., Jolesz, F.A.,
1094 Brookeman, J.R., 2000. Optimized single-slab three-dimensional spin-echo MR imaging of
1095 the brain. Radiology 216, 891-899.
1096 Mugler, J.P., 3rd, Brookeman, J.R., 1990. Three-dimensional magnetization-prepared rapid
1097 gradient-echo imaging (3D MP RAGE). Magn Reson Med 15, 152-157.
1098 Opel, N., Redlich, R., Kaehler, C., Grotegerd, D., Dohm, K., Heindel, W., Kugel, H.,
1099 Thalamuthu, A., Koutsouleris, N., Arolt, V., Teuber, A., Wersching, H., Baune, B.T., Berger, K.,
1100 Dannlowski, U., 2017. Prefrontal gray matter volume mediates genetic risks for obesity.
1101 Molecular Psychiatry 22, 703-710.
1102 Patterson, H.D., Thompson, R., 1971. Recovery of Inter-Block Information when Block Sizes
1103 are Unequal. Biometrika 58, 545-554.
1104 Peters, M.J., Joehanes, R., Pilling, L.C., Schurmann, C., Conneely, K.N., Powell, J., Reinmaa, E.,
1105 Sutphin, G.L., Zhernakova, A., Schramm, K., Wilson, Y.A., Kobes, S., Tukiainen, T.,
1106 Consortium, N.U., Ramos, Y.F., Goring, H.H., Fornage, M., Liu, Y., Gharib, S.A., Stranger, B.E.,
1107 De Jager, P.L., Aviv, A., Levy, D., Murabito, J.M., Munson, P.J., Huan, T., Hofman, A.,
1108 Uitterlinden, A.G., Rivadeneira, F., van Rooij, J., Stolk, L., Broer, L., Verbiest, M.M., Jhamai,
1109 M., Arp, P., Metspalu, A., Tserel, L., Milani, L., Samani, N.J., Peterson, P., Kasela, S., Codd, V.,
1110 Peters, A., Ward-Caviness, C.K., Herder, C., Waldenberger, M., Roden, M., Singmann, P.,
1111 Zeilinger, S., Illig, T., Homuth, G., Grabe, H.J., Volzke, H., Steil, L., Kocher, T., Murray, A.,
1112 Melzer, D., Yaghootkar, H., Bandinelli, S., Moses, E.K., Kent, J.W., Curran, J.E., Johnson, M.P.,
1113 Williams-Blangero, S., Westra, H.J., McRae, A.F., Smith, J.A., Kardia, S.L., Hovatta, I., Perola,
1114 M., Ripatti, S., Salomaa, V., Henders, A.K., Martin, N.G., Smith, A.K., Mehta, D., Binder, E.B.,
1115 Nylocks, K.M., Kennedy, E.M., Klengel, T., Ding, J., Suchy-Dicey, A.M., Enquobahrie, D.A.,
1116 Brody, J., Rotter, J.I., Chen, Y.D., Houwing-Duistermaat, J., Kloppenburg, M., Slagboom, P.E.,
1117 Helmer, Q., den Hollander, W., Bean, S., Raj, T., Bakhshi, N., Wang, Q.P., Oyston, L.J., Psaty,
1118 B.M., Tracy, R.P., Montgomery, G.W., Turner, S.T., Blangero, J., Meulenbelt, I., Ressler, K.J.,
1119 Yang, J., Franke, L., Kettunen, J., Visscher, P.M., Neely, G.G., Korstanje, R., Hanson, R.L.,
1120 Prokisch, H., Ferrucci, L., Esko, T., Teumer, A., van Meurs, J.B., Johnson, A.D., 2015. The
1121 transcriptional landscape of age in human peripheral blood. Nat Commun 6, 8570.
1122 Pinheiro, J., Bates, D., 2000. Mixed-Effects Models in S and S-PLUS. Springer New York.
1123 Pitel, A.L., Segobin, S.H., Ritz, L., Eustache, F., Beaunieux, H., 2015. Thalamic abnormalities
1124 are a cardinal feature of alcohol-related brain dysfunction. Neurosci Biobehav Rev 54, 38-45.
1125 Prom-Wormley, E., Maes, H.H.M., Schmitt, J.E., Panizzon, M.S., Xian, H., Eyler, L.T., Franz,
1126 C.E., Lyons, M.J., Tsuang, M.T., Dale, A.M., Fennema-Notestine, C., Kremen, W.S., Neale,
1127 M.C., 2015. Genetic and Environmental Contributions to the Relationships Between Brain
1128 Structure and Average Lifetime Cigarette Use. Behavior Genetics 45, 157-170.

R Development Core Team, 2012. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rajan, T.M., Menon, V., 2017. Psychiatric disorders and obesity: A review of association studies. Journal of Postgraduate Medicine 63, 182-190.

Ritchie, S.J., Cox, S.R., Shen, X., Lombardo, M.V., Reus, L.M., Alloza, C., Harris, M.A., Alderson, H.L., Hunter, S., Neilson, E., Liewald, D.C.M., Auyeung, B., Whalley, H.C., Lawrie, S.M., Gale, C.R., Bastin, M.E., McIntosh, A.M., Deary, I.J., 2018. Sex Differences in the Adult Human Brain: Evidence from 5216 UK Biobank Participants. Cereb Cortex 28, 2959-2975.

Robinson, G.K., 1991. That BLUP is a Good Thing: The Estimation of Random Effects. Statistical Science 6, 15-32.

Robinson, M.R., Kleinman, A., Graff, M., Vinkhuyzen, A.A.E., Couper, D., Miller, M.B., Peyrot, W.J., Abdellaoui, A., Zietsch, B.P., Nolte, I.M., van Vliet-Ostaptchouk, J.V., Snieder, H., The LifeLines Cohort, S., Genetic Investigation of Anthropometric Traits, c., Medland, S.E., Martin, N.G., Magnusson, P.K.E., Iacono, W.G., McGue, M., North, K.E., Yang, J., Visscher, P.M., 2017. Genetic evidence of assortative mating in humans. 1, 0016.

Rosenberg, M.D., Casey, B.J., Holmes, A.J., 2018. Prediction complements explanation in understanding the developing brain. Nat Commun 9, 589.

Roshchupkin, G.V., Gutman, B.A., Vernooij, M.W., Jahanshad, N., Martin, N.G., Hofman, A., McMahon, K.L., van der Lee, S.J., van Duijn, C.M., de Zubicaray, G.I., Uitterlinden, A.G., Wright, M.J., Niessen, W.J., Thompson, P.M., Ikram, M.A., Adams, H.H.H., 2016. Heritability of the shape of subcortical brain structures in the general population. Nat Commun 7, 13738.

Saarni, S.E., Saarni, S.I., Fogelholm, M., Heliovaara, M., Perala, J., Suvisaari, J., Lonnqvist, J., 2009. Body composition in psychotic disorders: a general population survey. Psychological Medicine 39, 801-810.

Sabuncu, M.R., Ge, T., Holmes, A.J., Smoller, J.W., Buckner, R.L., Fischl, B., Initia, A.D.N., 2016. Morphometricity as a measure of the neuroanatomical signature of a trait. Proceedings of the National Academy of Sciences of the United States of America 113, E5749-E5756.

Sartor, C.E., Bucholz, K.K., Nelson, E.C., Madden, P.A.F., Lynskey, M.T., Heath, A.C., 2011. Reporting Bias in the Association Between Age at First Alcohol Use and Heavy Episodic Drinking. Alcoholism-Clinical and Experimental Research 35, 1418-1425.

Schmaal, L., Hibar, D.P., Samann, P.G., Hall, G.B., Baune, B.T., Jahanshad, N., Cheung, J.W., van Erp, T.G., Bos, D., Ikram, M.A., Vernooij, M.W., Niessen, W.J., Tiemeier, H., Hofman, A., Wittfeld, K., Grabe, H.J., Janowitz, D., Bulow, R., Selonke, M., Volzke, H., Grotegerd, D., Dannlowski, U., Arolt, V., Opel, N., Heindel, W., Kugel, H., Hoehn, D., Czisch, M., Couvy-Duchesne, B., Renteria, M.E., Strike, L.T., Wright, M.J., Mills, N.T., de Zubicaray, G.I., McMahon, K.L., Medland, S.E., Martin, N.G., Gillespie, N.A., Goya-Maldonado, R., Gruber, O., Kramer, B., Hatton, S.N., Lagopoulos, J., Hickie, I.B., Frodl, T., Carballedo, A., Frey, E.M., van Velzen, L.S., Penninx, B.W., van Tol, M.J., van der Wee, N.J., Davey, C.G., Harrison, B.J., Mwangi, B., Cao, B., Soares, J.C., Veer, I.M., Walter, H., Schoepf, D., Zurowski, B., Konrad, C., Schramm, E., Normann, C., Schnell, K., Sacchet, M.D., Gotlib, I.H., MacQueen, G.M., Godlewska, B.R., Nickson, T., McIntosh, A.M., Papmeyer, M., Whalley, H.C., Hall, J., Sussmann, J.E., Li, M., Walter, M., Aftanas, L., Brack, I., Bokhan, N.A., Thompson, P.M., Veltman, D.J., 2016a. Cortical abnormalities in adults and adolescents with major depression based on brain scans from 20 cohorts worldwide in the ENIGMA Major Depressive Disorder Working Group. Mol Psychiatry.

Schmaal, L., Veltman, D.J., van Erp, T.G., Samann, P.G., Frodl, T., Jahanshad, N., Loehrer, E., Tiemeier, H., Hofman, A., Niessen, W.J., Vernooij, M.W., Ikram, M.A., Wittfeld, K., Grabe, H.J., Block, A., Hegenscheid, K., Volzke, H., Hoehn, D., Czisch, M., Lagopoulos, J., Hatton, S.N., Hickie, I.B., Goya-Maldonado, R., Kramer, B., Gruber, O., Couvy-Duchesne, B., Renteria, M.E., Strike, L.T., Mills, N.T., de Zubicaray, G.I., McMahon, K.L., Medland, S.E., Martin, N.G., Gillespie, N.A., Wright, M.J., Hall, G.B., MacQueen, G.M., Frey, E.M., Carballedo, A., van Velzen, L.S., van Tol, M.J., van der Wee, N.J., Veer, I.M., Walter, H., Schnell, K., Schramm, E., Normann, C., Schoepf, D., Konrad, C., Zurowski, B., Nickson, T., McIntosh, A.M., Papmeyer, M., Whalley, H.C., Sussmann, J.E., Godlewska, B.R., Cowen, P.J., Fischer, F.H., Rose, M., Penninx, B.W., Thompson, P.M., Hibar, D.P., 2016b. Subcortical brain alterations in major depressive disorder: findings from the ENIGMA Major Depressive Disorder working group. Mol Psychiatry 21, 806-812.

Self, S.G., Liang, K.Y., 1987. Asymptotic Properties of Maximum-Likelihood Estimators and Likelihood Ratio Tests under Nonstandard Conditions. Journal of the American Statistical Association 82, 605-610.

Solovieff, N., Cotsapas, C., Lee, P.H., Purcell, S.M., Smoller, J.W., 2013. Pleiotropy in complex traits: challenges and strategies. Nat Rev Genet 14, 483-495.

Stein, J.L., Medland, S.E., Vasquez, A.A., Hibar, D.P., Senstad, R.E., Winkler, A.M., Toro, R., Appel, K., Bartecek, R., Bergmann, O., Bernard, M., Brown, A.A., Cannon, D.M., Chakravarty, M.M., Christoforou, A., Domin, M., Grimm, O., Hollinshead, M., Holmes, A.J., Homuth, G., Hottenga, J.J., Langan, C., Lopez, L.M., Hansell, N.K., Hwang, K.S., Kim, S., Laje, G., Lee, P.H., Liu, X., Loth, E., Lourdusamy, A., Mattingsdal, M., Mohnke, S., Maniega, S.M., Nho, K., Nugent, A.C., O'Brien, C., Papmeyer, M., Putz, B., Ramasamy, A., Rasmussen, J., Rijpkema, M., Risacher, S.L., Roddey, J.C., Rose, E.J., Ryten, M., Shen, L., Sprooten, E., Strengman, E., Teumer, A., Trabzuni, D., Turner, J., van Eijk, K., van Erp, T.G., van Tol, M.J., Wittfeld, K., Wolf, C., Woudstra, S., Aleman, A., Alhusaini, S., Almasy, L., Binder, E.B., Brohawn, D.G., Cantor, R.M., Carless, M.A., Corvin, A., Czisch, M., Curran, J.E., Davies, G., de Almeida, M.A., Delanty, N., Depondt, C., Duggirala, R., Dyer, T.D., Erk, S., Fagerness, J., Fox, P.T., Freimer, N.B., Gill, M., Goring, H.H., Hagler, D.J., Hoehn, D., Holsboer, F., Hoogman, M., Hosten, N., Jahanshad, N., Johnson, M.P., Kasperaviciute, D., Kent, J.W., Jr., Kochunov, P., Lancaster, J.L., Lawrie, S.M., Liewald, D.C., Mandl, R., Matarin, M., Mattheisen, M., Meisenzahl, E., Melle, I., Moses, E.K., Muhleisen, T.W., Nauck, M., Nothen, M.M., Olvera, R.L., Pandolfo, M., Pike, G.B., Puls, R., Reinvang, I., Renteria, M.E., Rietschel, M., Roffman, J.L., Royle, N.A., Rujescu, D., Savitz, J., Schnack, H.G., Schnell, K., Seiferth, N., Smith, C., Steen, V.M., Valdes Hernandez, M.C., Van den Heuvel, M., van der Wee, N.J., Van Haren, N.E., Veltman, J.A., Volzke, H., Walker, R., Westlye, L.T., Whelan, C.D., Agartz, I., Boomsma, D.I., Cavalleri, G.L., Dale, A.M., Djurovic, S., Drevets, W.C., Hagoort, P., Hall, J., Heinz, A., Jack, C.R., Jr., Foroud, T.M., Le Hellard, S., Macciardi, F., Montgomery, G.W., Poline, J.B., Porteous, D.J., Sisodiya, S.M., Starr, J.M., Sussmann, J., Toga, A.W., Veltman, D.J., Walter, H., Weiner, M.W., Alzheimer's Disease Neuroimaging, I., Consortium, E., Consortium, I., Saguenay Youth Study, G., Bis, J.C., Ikram, M.A., Smith, A.V., Gudnason, V., Tzourio, C., Vernooij, M.W., Launer, L.J., DeCarli, C., Seshadri, S., Cohorts for, H., Aging Research in Genomic Epidemiology, C., Andreassen, O.A., Apostolova, L.G., Bastin, M.E., Blangero, J., Brunner, H.G., Buckner, R.L., Cichon, S., Coppola, G., de Zubicaray, G.I., Deary, I.J., Donohoe, G., de Geus, E.J., Espeseth, T., Fernandez, G., Glahn, D.C., Grabe, H.J., Hardy, J., Hulshoff Pol, H.E., Jenkinson, M., Kahn, R.S., McDonald, C., McIntosh, A.M., McMahon, F.J., McMahon, K.L., Meyer-Lindenberg, A., Morris, D.W., Muller-Myhsok, B., Nichols, T.E., Ophoff, R.A., Paus, T., Pausova, Z., Penninx,

B.W., Potkin, S.G., Samann, P.G., Saykin, A.J., Schumann, G., Smoller, J.W., Wardlaw, J.M., Weale, M.E., Martin, N.G., Franke, B., Wright, M.J., Thompson, P.M., Enhancing Neuro Imaging Genetics through Meta-Analysis, C., 2012. Identification of common variants associated with human hippocampal and intracranial volumes. Nat Genet 44, 552-561.

Stram, D.O., Lee, J.W., 1994. Variance components testing in the longitudinal mixed effects model. Biometrics 50, 1171-1177.

Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., Liu, B., Matthews, P., Ong, G., Pell, J., Silman, A., Young, A., Sprosen, T., Peakman, T., Collins, R., 2015. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 12, e1001779.

Temple, D.L., the CRAN Team, R., 2017. XML: Tools for Parsing and Generating XML Within R and S-Plus.

Thompson, P.M., Stein, J.L., Medland, S.E., Hibar, D.P., Vasquez, A.A., Renteria, M.E., Toro, R., Jahanshad, N., Schumann, G., Franke, B., Wright, M.J., Martin, N.G., Agartz, I., Alda, M., Alhusaini, S., Almasy, L., Almeida, J., Alpert, K., Andreasen, N.C., Andreassen, O.A., Apostolova, L.G., Appel, K., Armstrong, N.J., Aribisala, B., Bastin, M.E., Bauer, M., Bearden, C.E., Bergmann, O., Binder, E.B., Blangero, J., Bockholt, H.J., Boen, E., Bois, C., Boomsma, D.I., Booth, T., Bowman, I.J., Bralten, J., Brouwer, R.M., Brunner, H.G., Brohawn, D.G., Buckner, R.L., Buitelaar, J., Bulayeva, K., Bustillo, J.R., Calhoun, V.D., Cannon, D.M., Cantor, R.M., Carless, M.A., Caseras, X., Cavalleri, G.L., Chakravarty, M.M., Chang, K.D., Ching, C.R., Christoforou, A., Cichon, S., Clark, V.P., Conrod, P., Coppola, G., Crespo-Facorro, B., Curran, J.E., Czisch, M., Deary, I.J., de Geus, E.J., den Braber, A., Delvecchio, G., Depondt, C., de Haan, L., de Zubicaray, G.I., Dima, D., Dimitrova, R., Djurovic, S., Dong, H., Donohoe, G., Duggirala, R., Dyer, T.D., Ehrlich, S., Ekman, C.J., Elvsashagen, T., Emsell, L., Erk, S., Espeseth, T., Fagerness, J., Fears, S., Fedko, I., Fernandez, G., Fisher, S.E., Foroud, T., Fox, P.T., Francks, C., Frangou, S., Frey, E.M., Frodl, T., Frouin, V., Garavan, H., Giddaluru, S., Glahn, D.C., Godlewska, B., Goldstein, R.Z., Gollub, R.L., Grabe, H.J., Grimm, O., Gruber, O., Guadalupe, T., Gur, R.E., Gur, R.C., Goring, H.H., Hagenaars, S., Hajek, T., Hall, G.B., Hall, J., Hardy, J., Hartman, C.A., Hass, J., Hatton, S.N., Haukvik, U.K., Hegenscheid, K., Heinz, A., Hickie, I.B., Ho, B.C., Hoehn, D., Hoekstra, P.J., Hollinshead, M., Holmes, A.J., Homuth, G., Hoogman, M., Hong, L.E., Hosten, N., Hottenga, J.J., Hulshoff Pol, H.E., Hwang, K.S., Jack, C.R., Jr., Jenkinson, M., Johnston, C., Jonsson, E.G., Kahn, R.S., Kasperaviciute, D., Kelly, S., Kim, S., Kochunov, P., Koenders, L., Kramer, B., Kwok, J.B., Lagopoulos, J., Laje, G., Landen, M., Landman, B.A., Lauriello, J., Lawrie, S.M., Lee, P.H., Le Hellard, S., Lemaitre, H., Leonardo, C.D., Li, C.S., Liberg, B., Liewald, D.C., Liu, X., Lopez, L.M., Loth, E., Lourdusamy, A., Luciano, M., Macciardi, F., Machielsen, M.W., Macqueen, G.M., Malt, U.F., Mandl, R., Manoach, D.S., Martinot, J.L., Matarin, M., Mather, K.A., Mattheisen, M., Mattingsdal, M., Meyer-Lindenberg, A., McDonald, C., McIntosh, A.M., McMahon, F.J., McMahon, K.L., Meisenzahl, E., Melle, I., Milaneschi, Y., Mohnke, S., Montgomery, G.W., Morris, D.W., Moses, E.K., Mueller, B.A., Munoz Maniega, S., Muhleisen, T.W., Muller-Myhsok, B., Mwangi, B., Nauck, M., Nho, K., Nichols, T.E., Nilsson, L.G., Nugent, A.C., Nyberg, L., Olvera, R.L., Oosterlaan, J., Ophoff, R.A., Pandolfo, M., Papalampropoulou-Tsiridou, M., Papmeyer, M., Paus, T., Pausova, Z., Pearlson, G.D., Penninx, B.W., Peterson, C.P., Pfennig, A., Phillips, M., Pike, G.B., Poline, J.B., Potkin, S.G., Putz, B., Ramasamy, A., Rasmussen, J., Rietschel, M., Rijpkema, M., Risacher, S.L., Roffman, J.L., Roiz-Santianez, R., Romanczuk-Seiferth, N., Rose, E.J., Royle, N.A., Rujescu, D., Ryten, M., Sachdev, P.S., Salami, A., Satterthwaite, T.D., Savitz, J., Saykin, A.J., Scanlon, C., Schmaal, L., Schnack, H.G., Schork, A.J., Schulz, S.C., Schur, R., Seidman, L.,

Shen, L., Shoemaker, J.M., Simmons, A., Sisodiya, S.M., Smith, C., Smoller, J.W., Soares, J.C., Sponheim, S.R., Sprooten, E., Starr, J.M., Steen, V.M., Strakowski, S., Strike, L., Sussmann, J., Samann, P.G., Teumer, A., Toga, A.W., Tordesillas-Gutierrez, D., Trabzuni, D., Trost, S., Turner, J., Van den Heuvel, M., van der Wee, N.J., van Eijk, K., van Erp, T.G., van Haren, N.E., van 't Ent, D., van Tol, M.J., Valdes Hernandez, M.C., Veltman, D.J., Versace, A., Volzke, H., Walker, R., Walter, H., Wang, L., Wardlaw, J.M., Weale, M.E., Weiner, M.W., Wen, W., Westlye, L.T., Whalley, H.C., Whelan, C.D., White, T., Winkler, A.M., Wittfeld, K., Woldehawariat, G., Wolf, C., Zilles, D., Zwiers, M.P., Thalamuthu, A., Schofield, P.R., Freimer, N.B., Lawrence, N.S., Drevets, W., Alzheimer's Disease Neuroimaging Initiative, E.C.I.C.S.Y.S.G., 2014. The ENIGMA Consortium: large-scale collaborative analyses of neuroimaging and genetic data. Brain Imaging Behav 8, 153-182.

Thompson, R., 1973. Estimation of Variance and Covariance Components with an Application When Records Are Subject to Culling. Biometrics 29, 527-550.

Turner, S.D., 2014. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots.

Tyrrell, J., Jones, S.E., Beaumont, R., Astley, C.M., Lovell, R., Yaghootkar, H., Tuke, M., Ruth, K.S., Freathy, R.M., Hirschhorn, J.N., Wood, A.R., Murray, A., Weedon, M.N., Frayling, T.M., 2016. Height, body mass index, and socioeconomic status: mendelian randomisation study in UK Biobank. Bmj-British Medical Journal 352.

Van Essen, D.C., Glasser, M.F., Dierker, D.L., Harwell, J., Coalson, T., 2012a. Parcellations and hemispheric asymmetries of human cerebral cortex analyzed on surface-based atlases. Cereb Cortex 22, 2241-2262.

Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E., Yacoub, E., Ugurbil, K., Consortium, W.U.-M.H., 2013. The WU-Minn Human Connectome Project: an overview. Neuroimage 80, 62-79.

Van Essen, D.C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T.E., Bucholz, R., Chang, A., Chen, L., Corbetta, M., Curtiss, S.W., Della Penna, S., Feinberg, D., Glasser, M.F., Harel, N., Heath, A.C., Larson-Prior, L., Marcus, D., Michalareas, G., Moeller, S., Oostenveld, R., Petersen, S.E., Prior, F., Schlaggar, B.L., Smith, S.M., Snyder, A.Z., Xu, J., Yacoub, E., Consortium, W.U.-M.H., 2012b. The Human Connectome Project: a data acquisition perspective. Neuroimage 62, 2222-2231.

van Holst, R.J., de Ruiter, M.B., van den Brink, W., Veltman, D.J., Goudriaan, A.E., 2012. A voxel-based morphometry study comparing problem gamblers, alcohol abusers, and healthy controls. Drug and Alcohol Dependence 124, 142-148.

Varrichio, C., 2016. rowr: Row-Based Functions for R Objects.

Vilhjalmsson, B.J., Yang, J., Finucane, H.K., Gusev, A., Lindstrom, S., Ripke, S., Genovese, G., Loh, P.R., Bhatia, G., Do, R., Hayeck, T., Won, H.H., Kathiresan, S., Pato, M., Pato, C., Tamimi, R., Stahl, E., Zaitlen, N., Pasaniuc, B., Belbin, G., Kenny, E.E., Schierup, M.H., De Jager, P., Patsopouos, N.A., Mc Carroll, S., Daly, M., Purce, S., Chasman, D., Neale, B., Goddard, M., Visscher, P.M., Kraft, P., Patterson, N., Price, A.L., Consortium, P.G., Inherited, D.B.R., 2015. Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. American Journal of Human Genetics 97, 576-592.

Visscher, P.M., 1998. On the sampling variance of intraclass correlations and genetic correlations. Genetics 149, 1605-1614.

Visscher, P.M., Hemani, G., Vinkhuyzen, A.A.E., Chen, G.B., Lee, S.H., Wray, N.R., Goddard, M.E., Yang, J., 2014. Statistical Power to Detect Genetic (Co)Variance of Complex Traits Using SNP Data in Unrelated Samples. Plos Genetics 10.

Wei, T., Simko, V., 2017. R package "corrplot": Visualization of a Correlation Matrix.
Wickham, H., 2007. Reshaping data with the reshape package. Journal of Statistical Software 21, 1-20.
Wickham, H., 2009. Elegant Graphics for Data Analysis. Springer-Verlag, New York.
Wickham, H., 2017a. tidyverse: Easily Install and Load 'Tidyverse' Packages.
Wickham, H., Francois, R., 2015. dplyr: A Grammar of Data Manipulation.
Wickham, H.H., J.; Francois, R., 2017b. readr: Read Rectangular Text Data.
Wray, N.R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E.M., Abdellaoui, A., Adams, M.J., Agerbo, E., Air, T.M., Andlauer, T.M.F., Bacanu, S.A., Baekvad-Hansen, M., Beekman, A.F.T., Bigdeli, T.B., Binder, E.B., Blackwood, D.R.H., Bryois, J., Buttenschon, H.N., Bybjerg-Grauholm, J., Cai, N., Castelao, E., Christensen, J.H., Clarke, T.K., Coleman, J.I.R., Colodro-Conde, L., Couvy-Duchesne, B., Craddock, N., Crawford, G.E., Crowley, C.A., Dashti, H.S., Davies, G., Deary, I.J., Degenhardt, F., Derks, E.M., Direk, N., Dolan, C.V., Dunn, E.C., Eley, T.C., Eriksson, N., Escott-Price, V., Kiadeh, F.H.F., Finucane, H.K., Forstner, A.J., Frank, J., Gaspar, H.A., Gill, M., Giusti-Rodriguez, P., Goes, F.S., Gordon, S.D., Grove, J., Hall, L.S., Hannon, E., Hansen, C.S., Hansen, T.F., Herms, S., Hickie, I.B., Hoffmann, P., Homuth, G., Horn, C., Hottenga, J.J., Hougaard, D.M., Hu, M., Hyde, C.L., Ising, M., Jansen, R., Jin, F., Jorgenson, E., Knowles, J.A., Kohane, I.S., Kraft, J., Kretzschmar, W.W., Krogh, J., Kutalik, Z., Lane, J.M., Li, Y., Li, Y., Lind, P.A., Liu, X., Lu, L., MacIntyre, D.J., MacKinnon, D.F., Maier, R.M., Maier, W., Marchini, J., Mbarek, H., McGrath, P., McGuffin, P., Medland, S.E., Mehta, D., Middeldorp, C.M., Mihailov, E., Milaneschi, Y., Milani, L., Mill, J., Mondimore, F.M., Montgomery, G.W., Mostafavi, S., Mullins, N., Nauck, M., Ng, B., Nivard, M.G., Nyholt, D.R., O'Reilly, P.F., Oskarsson, H., Owen, M.J., Painter, J.N., Pedersen, C.B., Pedersen, M.G., Peterson, R.E., Pettersson, E., Peyrot, W.J., Pistis, G., Posthuma, D., Purcell, S.M., Quiroz, J.A., Qvist, P., Rice, J.P., Riley, B.P., Rivera, M., Saeed Mirza, S., Saxena, R., Schoevers, R., Schulte, E.C., Shen, L., Shi, J., Shyn, S.I., Sigurdsson, E., Sinnamon, G.B.C., Smit, J.H., Smith, D.J., Stefansson, H., Steinberg, S., Stockmeier, C.A., Streit, F., Strohmaier, J., Tansey, K.E., Teismann, H., Teumer, A., Thompson, W., Thomson, P.A., Thorgeirsson, T.E., Tian, C., Traylor, M., Treutlein, J., Trubetskoy, V., Uitterlinden, A.G., Umbricht, D., Van der Auwera, S., van Hemert, A.M., Viktorin, A., Visscher, P.M., Wang, Y., Webb, B.T., Weinsheimer, S.M., Wellmann, J., Willemsen, G., Witt, S.H., Wu, Y., Xi, H.S., Yang, J., Zhang, F., eQtlgen, andMe, Arolt, V., Baune, B.T., Berger, K., Boomsma, D.I., Cichon, S., Dannlowski, U., de Geus, E.C.J., DePaulo, J.R., Domenici, E., Domschke, K., Esko, T., Grabe, H.J., Hamilton, S.P., Hayward, C., Heath, A.C., Hinds, D.A., Kendler, K.S., Kloiber, S., Lewis, G., Li, Q.S., Lucae, S., Madden, P.F.A., Magnusson, P.K., Martin, N.G., McIntosh, A.M., Metspalu, A., Mors, O., Mortensen, P.B., Muller-Myhsok, B., Nordentoft, M., Nothen, M.M., O'Donovan, M.C., Paciga, S.A., Pedersen, N.L., Penninx, B., Perlis, R.H., Porteous, D.J., Potash, J.B., Preisig, M., Rietschel, M., Schaefer, C., Schulze, T.G., Smoller, J.W., Stefansson, K., Tiemeier, H., Uher, R., Volzke, H., Weissman, M.M., Werge, T., Winslow, A.R., Lewis, C.M., Levinson, D.F., Breen, G., Borglum, A.D., Sullivan, P.F., Major Depressive Disorder Working Group of the Psychiatric Genomics, C., 2018. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. Nat Genet 50, 668-681.
Wright, I.C., McGuire, P.K., Poline, J.B., Travere, J.M., Murray, R.M., Frith, C.D., Frackowiak, R.S., Friston, K.J., 1995. A voxel-based method for the statistical analysis of gray and white matter density applied to schizophrenia. Neuroimage 2, 244-252.
Yang, J., Benyamin, B., McEvoy, B.P., Gordon, S., Henders, A.K., Nyholt, D.R., Madden, P.A., Heath, A.C., Martin, N.G., Montgomery, G.W., Goddard, M.E., Visscher, P.M., 2010.

1364 Common SNPs explain a large proportion of the heritability for human height. Nat Genet 42,
1365 565-569.
1366 Yang, J., Lee, S.H., Goddard, M.E., Visscher, P.M., 2011. GCTA: a tool for genome-wide
1367 complex trait analysis. American Journal of Human Genetics 88, 76-82.
1368 Zhang, F., Chen, W., Zhu, Z., Zhang, Q., Nabais, M.F., Qi, T., Deary, I.J., Wray, N.R., Visscher,
1369 P.M., McRae, A.F., Yang, J., 2019. OSCA: a tool for omic-data-based complex trait analysis.
1370 Genome Biology 20, 107.
1371

1372