



HAL
open science

A unified framework for association and prediction from vertex-wise grey-matter structure

Baptiste Couvy-duchesne, Lachlan T Strike, Futao Zhang, Yan Holtz, Zhili Zheng, Kathryn E Kemper, Loïc Yengo, Olivier Colliot, Margaret Wright, Naomi Wray, et al.

► To cite this version:

Baptiste Couvy-duchesne, Lachlan T Strike, Futao Zhang, Yan Holtz, Zhili Zheng, et al.. A unified framework for association and prediction from vertex-wise grey-matter structure. *Human Brain Mapping*, 2020, 10.1002/hbm.25109 . hal-02904458

HAL Id: hal-02904458

<https://hal.science/hal-02904458>

Submitted on 22 Jul 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 A unified framework for association and prediction from vertex-wise
2 grey-matter structure

3 *Running title: Association and prediction from grey-matter*

4 Baptiste Couvy-Duchesne, PhD¹, Lachlan T. Strike, PhD², Futao Zhang, PhD¹, Yan Holtz,
5 MS,^{1,2}, Zhili Zheng, MD^{1,3}, Kathryn E. Kemper, PhD¹, Loic Yengo, PhD¹, Olivier Colliot^{4,5,6,7,8},
6 Margaret J. Wright, PhD^{2,9}, Naomi R. Wray, PhD^{1,2,*}, Jian Yang, PhD^{1,3,*}, Peter M. Visscher,
7 PhD^{1,2,*}

8

9 1: Institute for Molecular Bioscience, the University of Queensland, 4072 St Lucia, QLD,
10 Australia

11 2: Queensland Brain Institute, the University of Queensland, 4072 St Lucia, QLD, Australia

12 3: Institute for Advanced Research, Wenzhou Medical University, Wenzhou, Zhejiang
13 325027, China

14 4 : Inria, ARAMIS Project-team, 75013, Paris, France

15 5 : Institut du Cerveau et de la Moelle épinière, 75013, Paris, France

16 6 : Inserm, U1127, 75013, Paris, France

17 7 : CNRS, UMR 7225, 75013, Paris, France

18 8 : Sorbonne Universitee, 75013, Paris, France

19 9: Centre for Advanced Imaging, the University of Queensland, 4072 St Lucia, QLD, Australia

20 *: These authors jointly supervised this work.

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

24 **1. Acknowledgements**

25 This research was supported by the Australian National Health and Medical Research
26 Council (1078037, 1078901, 1113400, 1161356 and 1107258), the Australian Research
27 Council (FT180100186 and FL180100072), the Sylvia & Charles Viertel Charitable
28 Foundation, as well as the Agence Nationale de la Recherche as part of the "Investissements
29 d'avenir" program, reference ANR-19-P3IA-0001 (PRAIRIE 3IA Institute) and reference ANR-
30 10-IAIHU-06 (Agence Nationale de la Recherche-10-IA Institut Hospitalo-Universitaire-6).

31 Informed consent was obtained from all UK Biobank participants. Procedures are
32 controlled by a dedicated Ethics and Guidance Council
33 (<http://www.ukbiobank.ac.uk/ethics>), with the Ethics and Governance Framework available
34 at <http://www.ukbiobank.ac.uk/wp-content/uploads/2011/05/EGF20082.pdf>. IRB approval
35 was also obtained from the North West Multi-centre Research Ethics Committee. This
36 research has been conducted using the UK Biobank Resource under Application Number
37 12505.

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

43 We used R(R Development Core Team, 2012) (v3.3-3.6) for analyses not performed
44 using OSCA (Zhang et al., 2019) and for plots. We used colour-blind friendly R palettes in the
45 *Viridis* and *scales* (H. Wickham, 2015) packages, *qqman* (Turner, 2014) for QQ-plots, *ggplot2*
46 (H. Wickham, 2009) and *ggsignif* (Ahlmann-Eltze, 2017) for circular bar plots, *corrplot* (Wei
47 & Simko, 2017) for correlation matrix plots, *ukbtools* (Hanscombe, 2017) to facilitate UKB

48 phenotype manipulation, *bisgstatsr* (Privé, Aschard, Ziyatdinov, & Blum, 2018) for Lasso
49 regression, *meta* for meta-analyses (Balduzzi, Rücker, & Schwarzer, 2019). Other packages
50 used to assist analyses and data handling include *Hmisc* (Harrell, 2017), *rowr* (Varrichio,
51 2016), *pwr* (Champely, 2017), *XML* (Temple & the CRAN Team, 2017), *tidyverse* (H.
52 Wickham, 2017), *dplyr* (Hadley. Wickham & Francois, 2015), *readr* (H. H. Wickham, J.;
53 Francois, R., 2017), *reshape2* (H. Wickham, 2007), *RMarkdown* (Allaire, 2018), and *epuRate*
54 (Holtz, 2020).

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

58

59 **2. Abstract**

60 The recent availability of large-scale neuroimaging cohorts facilitates deeper
61 characterisation of the relationship between phenotypic and brain architecture variation in
62 humans. Here, we investigate the association (previously coined morphometricity) of a
63 phenotype with all 652,283 vertex-wise measures of cortical and subcortical morphology in
64 a large data set from the UK Biobank (UKB; N=9,497 for discovery, N=4,323 for replication)
65 and the Human Connectome Project (N=1,110). We used a linear mixed model with the
66 brain measures of individuals fitted as random effects with covariance relationships
67 estimated from the imaging data. We tested 167 behavioural, cognitive, psychiatric or
68 lifestyle phenotypes and found significant morphometricity for 58 phenotypes (spanning
69 substance use, blood assay results, education or income level, diet, depression, and
70 cognition domains), 23 of which replicated in the UKB replication set or the HCP. We then
71 extended the model for a bivariate analysis to estimate grey-matter correlation between

72 phenotypes, which revealed that body size (i.e., height, weight, BMI, waist and hip
73 circumference, body fat percentage) could account for a substantial proportion of the
74 morphometricity (confirmed using a conditional analysis), providing possible insight into
75 previous MRI case-control results for psychiatric disorders where case status is associated
76 with body mass index. Our LMM framework also allowed to predict some of the associated
77 phenotypes from the vertex-wise measures, in two independent samples. Finally, we
78 demonstrated additional new applications of our approach (i) Region Of Interest (ROI)
79 analysis that retain the vertex-wise complexity; (ii) comparison of the information retained
80 by different MRI processings.

81

82 **3. Keywords**

83 Brain MRI, Morphometricity, mixed models, association, prediction, grey-matter correlation

84

85 **4. Introduction**

86 The field of MRI studies is at a turning point owing to the recent availability of large data
87 sets to researchers, including the UKB (Miller et al., 2016) and HCP (Van Essen et al., 2013;
88 Van Essen, Ugurbil, et al., 2012) samples. These datasets promote the replication of
89 previous findings, but also the identification of small(er) associations and the expansion of
90 the range of phenotypes available for study (e.g. psychiatric symptoms and lifestyle factors).
91 Furthermore, the boost in statistical power may allow the simultaneous use of all the brain
92 complexity data of current MRI acquisitions rather than relying on data reduction
93 techniques (e.g. the region-of-interest [ROI] approach). In addition, these community
94 samples can complement the typical case-control paradigm by identifying confounders of
95 MRI analyses or by studying related traits (e.g. cognition domains relevant in Alzheimer's
96 disease). However, "big-data" neuroimaging offers a number of statistical challenges (on top

97 of the obvious computing ones)(Smith & Nichols, 2018): i) the curse of dimensionality (the
98 number of tests may increase faster than the sample size) which requires efficient methods
99 and appropriate control of multiple testing; ii) the possibility that (small) associations result
100 from confounding (via another variable, or acquisition noise); iii) the difficulty to generate
101 prediction from complex datasets.

102 Here, we propose a linear mixed model (LMM), efficiently implemented to tackle several
103 of these big-data neuroimaging challenges. Our approach allows performing association and
104 prediction analyses on tens of thousands of participants with more than 650,000 vertex-
105 wise morphological measurements of grey-matter structure per individual. Specifically, we
106 overcame the curse of dimensionality by estimating the total correlation of all cortical and
107 subcortical measurement at vertices with a phenotype of interest (previously coined
108 morphometricity (Sabuncu et al., 2016), here we prefer the more specific brain-
109 morphometricity). Using the same framework, we also estimate the total association of a
110 trait with differently processed MRI images as well with subset of the vertex-wise data
111 corresponding to specific brain features, hemispheres or regions of interest (ROIs). We
112 further introduce multi-trait LMMs that can quantify shared morphometricity between
113 traits (grey-matter correlation). Grey-matter correlation can help generate hypotheses
114 about putative confounders (that may be regressed out in a conditional analysis) or about
115 the origin of brain-morphometricity. Finally, we show how the same LMMs can be used to
116 construct grey-matter scores that achieve brain MRI-based prediction in independent
117 samples. As such, our approach unifies association and prediction analyses, in order to
118 unravel the brain-phenome relationships (Rosenberg, Casey, & Holmes, 2018) in big-data
119 neuroimaging.

120 To demonstrate the applicability and usefulness of our methods, we analysed two of the
121 largest MRI datasets available (UKB [split into discovery N=9,888 and replication N=4,561]
122 and HCP [N=1,110]) and considered a wide range of phenotypes spanning demographics,
123 blood cell composition, diet, psychiatric and traumatic history, physical capacities, and
124 substance use. We discuss our results in the context of the recent commentary article of
125 Smith & Nichols (Smith & Nichols, 2018). We have released our image processing and
126 analysis software/scripts as well as all summary statistics to facilitate replication and re-use
127 of the results.

128

129 **5. Materials and Methods**

130 *5.1. UK Biobank (UKB) sample(s)*

131 The UKB participants were unselected volunteers from the United Kingdom (Sudlow
132 et al., 2015) living near the imaging centres (Manchester for 96.5% of our sample, Newcastle
133 for the remaining 3.5%). Exclusion criteria included: presence of metal implant, recent
134 surgery and health conditions problematic for MRI imaging (e.g. hearing, breathing
135 problems or extreme claustrophobia) (Miller et al., 2016). MRI acquisition parameters have
136 been reported previously (Miller et al., 2016) and are summarised in **Appendix S1**.

137 We split the available UKB data into a discovery and replication sample based on
138 their imaging date. The discovery sample consisted of 9,497 adults aged 62.5 on average
139 (SD=7.5, range 44.6–79.6) and comprised 52.4% of female participants (see **Appendix S2** for
140 details of processing and QC; **Dataset S1** for description of excluded participants). The UKB
141 replication sample (N=4,323) was on average 63.1 years old (SD=7.46, range 46.1-80.3) with
142 52.1% of females (see **Dataset S1, Appendix S2**).

143 We included 168 variables grouped in several categories: demographics, cognition,
144 physical test, psychiatry, recent feelings, stress and traumas, substance use, miscellaneous,

145 brain measurements, blood assay and diet (see **Dataset S2** for details). When longitudinal
146 observations were available for a participant, we used the one collected as part of the
147 imaging assessment or the closest in time to that.

148

149 *5.2. UKB image processing*

150 We processed the T1w and T2w images using FreeSurfer 6.0 (Fischl, 2012) to extract
151 cortical surface area and thickness, and we used the ENIGMA-shape protocols to measure
152 the structure of 7 subcortical volumes (hippocampus, putamen, amygdala, thalamus,
153 caudate, pallidum and accumbens) (Boris A. Gutman, Madsen, Toga, & Thompson, 2013; B.
154 A. Gutman, Wang, Rajagopalan, Toga, & Thompson, 2012). In FreeSurfer, we processed T1w
155 and T2w together to enhance the tissue segmentation, hence a more precise skull stripping
156 and pial surfaces definition. When the T2w was not acquired or not usable, we processed
157 the T1w image by itself.

158 We retained the full cortical information by using the (“fsaverage”) cortical mesh for
159 cortical thickness and surface area. This corresponded to about ~149,900 cortical vertices
160 for each hemisphere and modality. In addition, we extracted subcortical radial thickness and
161 log Jacobian determinant (surface deformation, somewhat analogous to a relative surface
162 area (Roshchupkin et al., 2016)) for 13,560 vertices across the 7 subcortical volumes (Boris
163 A. Gutman et al., 2013). Overall, the imaging data used in the analyses comprised 652,283
164 vertex measurements per individual: 299,009 for cortical thickness, 299,034 for cortical
165 surface area, 27,120 for subcortical thickness and 27,120 for subcortical curvature.

166 For comparison with ROI based processing, we extracted cortical thickness and
167 surface area of 34 cortical regions (Desikan et al., 2006; Fischl et al., 2004) and volumes of
168 the subcortical structures (ENIGMA processing). To further the comparison of processing

169 options, we extracted cortical measurements from smoothed fsaverage meshes (fwhm 5,
170 10, 15, 20 and 25mm) as well as (unsmoothed) coarser meshes provided by FreeSurfer:
171 fsaverage6 (149,091 vertices across all hemispheres and modalities), fsaverage5 (37,455
172 vertices), fsaverage4 (9,457 vertices) and fsaverage3 (2,414 vertices).

173

174 *5.3. Human Connectome Project (HCP) sample*

175 HCP participants were recruited from ongoing longitudinal studies of the Missouri
176 Family Study and had to be between 22 and 35 years of age. Inclusion and exclusion criteria
177 have been described previously (Van Essen, Ugurbil, et al., 2012) (see **Appendix S1** for the
178 MRI acquisition parameters, **Appendix S2** for QC). As per the HCP “1200 Subjects data
179 release” (1st of March 2017), 1,113 participants were scanned on the 3T MRI and underwent
180 extensive behavioural testing. Participants were mostly (54.4%) females and were 28.8
181 years old on average (SD=3.7, range 22–37). The sample comprised 286 monozygotic twins
182 (138 complete pairs) and 169 dizygotic twins (78 complete pairs). In addition, siblings and
183 half siblings of twins were also recruited which resulted in 445 distinct families in the
184 sample.

185 For the HCP sample, we included 161 variables, some of which were also available in
186 the UKB (e.g. demographics, cognition, physical assessment, blood assay or psychiatry). We
187 also included: personality, emotion, mental health assessment (Semi-Structured Assessment
188 for the Genetics of Alcoholism (SSAGA) and Adult Self Report (ASR) (Thomas M Achenbach,
189 2009; T. M. Achenbach, Dumenci, & Rescorla, 2003)), detailed cognition, Pittsburgh sleep
190 index (PSQI) (Buysse, Reynolds, Monk, Berman, & Kupfer, 1989), or results from the urine
191 drug tests (see **Dataset S2**).

192

193 *5.4. Image processing in the HCP*

194 The FreeSurfer processing was performed by the HCP team (Glasser et al., 2013;
195 Marcus et al., 2013; Van Essen, Glasser, Dierker, Harwell, & Coalson, 2012) using an optimal
196 combination of automated and manual steps (**Appendix S3**). We downloaded the
197 segmented images (Marcus et al., 2011) and performed the ENIGMA-shape analysis (Boris
198 A. Gutman et al., 2013; B. A. Gutman et al., 2012) to extract vertex-wise measurements of
199 the subcortical volumes. As with the UKB sample, a total of 652,283 vertex measurements
200 were extracted for each individual.

201

202 *5.5. Covariates used*

203 Our baseline model included commonly used covariates in MRI analyses: acquisition
204 variables (UKB imaging wave, processing with T1w or with combined T1w+T2w), age, sex,
205 and head size (intra-cranial volume (ICV) as well as left and right total cortical surface area
206 and cortical thickness that correspond to the measurements used here). In a follow-up
207 analysis, we included other covariates such as height, weight and BMI to evaluate their
208 confounding effect on the reported associations. As some of the covariates are correlated
209 we report the adjusted R^2 (from linear regression in R3.3.3 (R Development Core Team,
210 2012)) calculated by adding progressively the covariates (same order as above). The
211 associations with covariates was highly concordant between the two UKB samples (**Figure**
212 **S1**).

213

214 *5.6. Linear mixed models for association and prediction*

215 We aimed to estimate the proportion of variance of a trait captured by brain
216 features, which Sabuncu et al., called “morphometricity” (Sabuncu et al., 2016). To do so we
217 consider the following linear mixed model (Figure 1) that allows estimating the association
218 between a phenotype and M vertices even when M is greater than the sample size (N):

219
$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{b} + \mathbf{e} \quad (1)$$

220 where $\mathbf{Y}_{N,1}$ is the phenotype considered with N the number of observations, $\mathbf{X}_{N,c}$ is a matrix
 221 of c covariates (as such does not include any vertex variable), $\boldsymbol{\beta}_{c,1}$ is a vector of fixed effects,
 222 \mathbf{b} is a vector of random effects with $\mathbf{b} \sim \mathcal{N}(0, \mathbf{B}\sigma_b^2)$ and \mathbf{e} is a vector of error terms with
 223 $\mathbf{e} \sim \mathcal{N}(0, \mathbf{I}\sigma_e^2)$. In this formulation $\mathbf{I}_{N,N}$ is the identity matrix as we assume the error terms to
 224 be independent and identically distributed. $\mathbf{B}_{N,N}$ is a matrix of variance-covariance between
 225 individuals calculated from all vertex measurements, which we will refer to as the brain
 226 relatedness matrix (BRM, **Figure 1**). Off-diagonal elements of the BRM reflect the grey-
 227 matter similarity between two individuals (see **Appendix S4**). Finally, σ_e^2 and σ_b^2 are the
 228 variance components for the random effects \mathbf{e} and \mathbf{b} . For context, this model is analogous
 229 to that used in complex trait genetics to estimate SNP-based heritability, where a Genetic
 230 Relatedness Matrix (GRM) replaces the BRM (Yang et al., 2010; Yang, Lee, Goddard, &
 231 Visscher, 2011), or that used to estimate the proportion variance in a phenotype captured
 232 by all DNA methylation or gene expression measures of the genome (Zhang et al., 2019).

233 The element i,j of the BRM can be calculated as the inner product of brain measurements of
 234 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

235 individual i centred and standardised over all individuals, $z_{j,m}$ represents the value of vertex
 236 m for individual j centred and standardised over all individuals, M is the total number of

237 vertices or brain features included. In matrix notation, $\mathbf{B} = \frac{\mathbf{Z}\mathbf{Z}'}{M}$ with $\mathbf{Z}_{N,M}$ being a matrix of
 238 the centred and standardised brain observations. We estimated the proportion of the trait

239 variance captured by the grey-matter measurements as: $R^2 = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_e^2}$ (**Figure 1**) using the

240 REstricted Maximum Likelihood (REML)(Patterson & Thompson, 1971) implemented in
 241 OSCA (OmicS-data-based Complex trait Analysis) (Zhang et al., 2019). We tested whether

242 the morphometricity was different from 0 using a likelihood ratio test (see **Appendix S5** for
243 details).

244 We extended the LMM above to jointly estimate the variance accounted for by the
245 different modalities (cortical thickness, cortical area, subcortical thickness, subcortical area).

$$246 \quad \mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{b}_1 + \mathbf{b}_2 + \mathbf{b}_3 + \mathbf{b}_4 + \mathbf{e}$$

247 with $\mathbf{b}_i \sim \mathcal{N}(0, \mathbf{B}_i \sigma_{b_i}^2)$, $i \in \llbracket 1; 4 \rrbracket$, and all other parameters left unchanged. Each \mathbf{B}_i is
248 constructed from the vertex-wise measurements of a single modality, with $\sigma_{b_i}^2$ the
249 corresponding association and $\frac{\sigma_{b_1}^2 + \sigma_{b_2}^2 + \sigma_{b_3}^2 + \sigma_{b_4}^2}{\sigma_{b_1}^2 + \sigma_{b_2}^2 + \sigma_{b_3}^2 + \sigma_{b_4}^2 + \sigma_e^2}$ the brain-morphometricity.

250 Next, we sought to estimate the correlation between two traits that is attributable
251 to the same grey-matter variation, which we call grey-matter correlation r_{GM} (**Figure 1**). This
252 can be achieved by fitting a bivariate LMM, a direct extension of the models presented
253 above (Thompson, 1973). We restricted our bivariate analysis to variables that were
254 significantly associated with grey-matter structure. We derived the residual correlations (r_E)
255 from the phenotypic (r) and grey-matter correlations estimated by GCTA (Genome-wide
256 Complex Trait Analysis) (Yang et al., 2011) (option not yet included in OSCA). We calculated
257 its SE using the delta method (**Appendix S6** and (Bijma & Bastiaansen, 2014; Lee, Yang,
258 Goddard, Visscher, & Wray, 2012; Visscher, 1998)).

259 We detailed power calculations for the LMMs (**Appendix S7**, (Visscher et al., 2014)),
260 which showed that in the UKB discovery sample we had good power to detect a small
261 morphometricity ($R^2 > 2.2\%$) but only a moderate grey-matter correlation ($r_{GM} > 0.35$).
262 Statistical power was a lot reduced in the HCP due to the smaller sample size.

263 We demonstrated two further utilities of LMMs for neuroimaging data analyses.
264 First, we conducted post-hoc analyses to test the associations with each modality and each

265 cortical (Desikan et al., 2006) or subcortical structure. We used BRMs specific to each region
266 and brain measurement (**Figure 1**), which bridges the gap between ROI and vertex-wise
267 analyses. Second, we define as “best” processing the MRI cortical processing that maximises
268 the association with a trait of interest, from the minimal number of features (vertices).
269 Thus, we evaluated which of our FreeSurfer processing (fsaverage – no smoothing;
270 fsaverage – smoothing fwhm5, 10, 15, 20, 25; fsaverage6, 5, 4, 3 – no smoothing; ENIGMA
271 ROI processing) maximised the brain-morphometricity for all the UKB traits (See **Appendix**
272 **S2** for details about QC). As the ENIGMA processing only consists of 150 measurements, we
273 used linear models (multiple regression and adjusted R^2) to estimate the brain-
274 morphometricity.

275 Finally, we derived brain prediction scores using the Best Linear Unbiased Predictors
276 (BLUP, **Figure 1**) (Henderson, 1950, 1975; G. K. Robinson, 1991) and evaluated them in the
277 UKB discovery sample using a 10-fold cross-validation design. In addition, we derived BLUP
278 brain prediction scores constructed from the UKB discovery sample, and applied them to
279 the UKB replication and HCP participants to evaluate the “out of sample” predictive
280 performance. BLUP estimates the predicted values of the random effects (**b or Zu**, see (1)
281 and **Figure 1**) (Goddard, Wray, Verbyla, & Visscher, 2009; G. K. Robinson, 1991). In short,
282 BLUP scores integrate the correlations between vertices to derive weights that correspond
283 to the joint effects of all the vertices (**Figure 1**). BLUP have desirable statistical properties:
284 they are unbiased and are best predictors in the sense that they minimise the mean square
285 error in the class of linear unbiased predictors (Henderson, 1975; G. K. Robinson, 1991),
286 leading to more accurate prediction than other linear predictors (M. R. Robinson et al.,
287 2017; Vilhjalmsjon et al., 2015).

288

289 *5.7. Prediction accuracy of BLUP versus LASSO*

290 We compared prediction accuracy achieved by BLUP scores to that of LASSO (least
291 absolute shrinkage and selection operator) (Tibshirani, 1996) for phenotypes with significant
292 brain-morphometricity (baseline covariates). LASSO penalises vertices coefficients of the
293 linear regression, leading to select a subset of vertices (and their weights) that maximise
294 prediction accuracy. We used the LASSO function implemented in the bigstatsr R package
295 (Privé et al., 2018) and estimated the hyper-parameter using cross-model selection and
296 averaging on 5 folds within the UKB discovery sample. For each grey-matter score, we
297 reported the prediction R^2 on the UKB replication sample and tested the difference in
298 prediction using a Wilcoxon test on the absolute errors of the BLUP and LASSO predictors.

299

300 *5.8. Data and code availability statement*

301 Data used in this manuscript are held and distributed by the HCP and UKB teams. We
302 have released the scripts used in image processing and LMM analyses to facilitate
303 replication and dissemination of the results (see **URLs**). We have also released BLUP weights
304 to allow meta-analyses or application of the grey-matter scores in independent cohorts.

305

306 **6. Results**

307 *6.1. Associations between phenotypes and all grey-matter structure vertices*

308 For the phenotypes of interest, we summarised in circular barplots (**Figure 2**) the
309 proportion of phenotypic variance associated with all 652,283 vertex-wise grey-matter
310 measures (brain-morphometricity, R^2) as well as with baseline covariates (see **Methods**).
311 **Figure 2** shows only the results that were significant after Bonferroni correction
312 ($p_{\text{UKB_discovery}} < 2.8e-4$ and $p_{\text{HCP}} < 2.9e-4$). The full results are available in **Dataset S3, S4** (see
313 **Figure S2** for positive control associations with global measures of the brain).

314 Grey-matter structure was strongly associated ($R^2 > 0.40$) with age, sex, as well as
315 weight, BMI waist and hip circumference but also with maternal smoking around birth
316 ($R^2 = 0.39$) and number of cigarettes previously smoked ($R^2 = 0.27$) (**Figure 2**). We identified
317 many other phenotypes significantly associated with grey-matter structure (**Figure 2**,
318 **Dataset S3**) including other measures of build (e.g. height, body fat percentage, basal
319 metabolic rate), substance use (e.g. amount of alcohol drank each day), household income
320 level and education level, strength (e.g. hand grip, acceleration), cognition (e.g. fluid IQ),
321 blood assay (e.g. white blood cell count), diet (cheese intake), but also perhaps more
322 surprisingly with being a twin or overall health rating. We also found associations with
323 clinical phenotypes such as diabetes, depression score and depression symptoms. We
324 replicated 23 of the 58 associations listed above in the UKB replication sample ($p < 0.05/58$;
325 **Figure S3, Dataset S4**). We did not detect any significant association between grey-matter
326 structure and other psychiatric variables (diagnoses and symptoms), self-reported stresses
327 and traumas, or neuroticism (**Dataset S3**). The interested readers may also find the
328 morphometricity estimates for the full UKB sample (inverse-variance weighted meta-
329 analysis) in **Dataset S3**.

330 In the UKB (discovery), results and conclusions did not change regardless of fitting a
331 single random effect or several random effects each corresponding to one of the grey-
332 matter modalities (**Figure S4**). In the HCP, we observed 3 extra significant associations
333 between grey-matter structure and cocaine (urine test), self-reported number of times used
334 cocaine or hallucinogens. Similar to the association found with opiate (urine test), these
335 results warrant replication due to the small number of positive participants. Finally, the HCP
336 results did not change when excluding related individuals (**Appendix S8**).

337

338 6.2. *Adjustment for possible confounders*

339 The large associations between grey-matter structure and height, weight, BMI, waist
340 and hip circumference (**Figure 2**) led us to perform a sensitivity analysis to evaluate their
341 contribution to the brain-morphometricity of the traits studied. We repeated the analysis
342 further controlling for height, weight and BMI, which yielded lower R^2 estimates (**Figure S5**)
343 and fewer significant associations with grey-matter structure. Thus, when correcting for
344 height in the UKB, 4 of the 58 associations with grey-matter structure did not remain
345 significant: household income, monocyte percentage, beef intake, and time spent using
346 computer, **Dataset S3**). Such finding is consistent with the reported association between
347 body size and income or socio-economic status in the UKB (Tyrrell et al., 2016). When
348 further correcting for weight and BMI another 14 associations did not remain significant
349 including educational attainment, frequency drinking alcohol, most diet items (cereal, dried
350 fruits, poultry, processed meat), time spent driving, red blood cell count, frequency of walks
351 and small exercise. Notably, the brain-morphometricity of the depression score could be
352 completely explained by differences in weight and BMI ($R^2_{\text{baseline}}=0.050$, $SE=0.018$;
353 $R^2_{\text{baseline+height}}=0.048$, $SE=0.017$, $R^2_{\text{baseline+height+BMI+weight}}<0.001$, $SE=0.007$), and none of the
354 associations between grey-matter structure and depression symptoms remained significant
355 conditioning on weight and BMI (Tiredness, Anhedonia, Poor appetite- overeating,
356 $R^2_{\text{baseline+height+BMI+weight}}<0.014$). Yet, even after controlling for body size, we still detected a
357 significant morphometricity for cheese intake as well as time watching TV (**Dataset S3**),
358 suggesting that these behaviours are associated with brain structure irrespective of body
359 size. The morphometricity estimates in the UKB replication sample aligned with those from
360 the discovery sample ($cor=0.90$), except for age and sex that showed larger associations
361 with grey-matter structure in the replication analysis (**Figure S6**). In the HCP dataset, after

362 controlling for body size, four of the 27 associations did not remain significant (**Dataset S4**)
363 though we had limited power to detect associations smaller than R^2 of 0.2 in this sample
364 (see **Appendix S7**).

365 In light of these results, we chose a conservative approach to control for body size
366 variables in the main text, though the analyses using baseline covariates can be found in the
367 supplementary. We acknowledge (see discussion) that this may be overly conservative, by
368 implicitly making strong assumptions about body size acting as a confounding factor. On the
369 other hand, it avoids reporting associations that may be fully or in part caused by
370 differences in body shape.

371

372 *6.3. Grey-matter correlations*

373 We estimated grey-matter correlation (r_{GM}) between the phenotypes that showed
374 significant brain-morphometricity in the univariate analyses (Figure 2). r_{GM} can be
375 interpreted as the correlation between the grey-matter vertices associations with each trait.
376 We controlled for height, weight and BMI on top of the baseline covariates, leaving a
377 conservative set of 35 UKB (18 HCP) phenotypes (**Figure 3; Datasets S5 [UKB], S6 [HCP]**). In
378 the UKB, we observed significant positive grey-matter correlations between cognition
379 domains, substance use phenotypes or between measures of physical activity (**Figure 3**). In
380 addition, we found unexpected large grey-matter correlations. For example, cheese intake
381 and forced expiratory volume were both correlated ($r_{GM}=1.0$, $SE=0.11$) with fluid
382 intelligence, and waist circumference was correlated with overall health rating and pulse
383 rate ($r_{GM}>0.67$). Overall, 9 out of the 26 significant correlations replicated in the UKB
384 replication sample ($p<0.05/26$ i.e. $p<1.9e-3$, **Table S1**). In the HCP, we also observed positive
385 grey-matter correlations between cognition domains or between the two tobacco related

386 phenotypes. Though, unlike in the UKB, we found a significant r_{GM} between IQ dimensions
387 and education level (**Figure 3, Dataset S6**).

388 For completeness, we estimated grey-matter correlations under the baseline model
389 (**Figure S7**), which reveals many large grey-matter correlations between measures of body
390 size and diet, blood assay, activity levels and depression symptoms and score. These results
391 further highlight that in the phenome, the brain-morphometricity of some traits may be
392 accounted for by the covariation between these phenotypes and body size measurements.
393 In particular, depression score was correlated ($r_{GM}=1$) with weight, BMI waist or hip
394 circumference, consistent with its brain-morphometricity lowered to 0 when controlling for
395 body size (**Figure S7**).

396

397 *6.4. Associations with grey-matter structure of specific cortical and subcortical regions*

398 We investigated the brain-morphometricity of traits by estimating the association
399 with grey-matter structure of specific cortical (Desikan et al., 2006) and subcortical regions,
400 correcting for multiple testing (Bonferroni significance threshold of $0.05/(164*39)=7.2e-6$ in
401 the UKB, $1.2e-5$ in the HCP). We found many significant ROIs associations with UKB
402 phenotypes, including age, sex, maternal smoking around birth, fluid intelligence, diabetes
403 or substance use (**Figure S8 and Dataset S7**). In particular, the associations between grey-
404 matter structure and body size were pervasive (72/164 significant ROIs associations with
405 height, 109 with waist circumference, 105 with BMI) (**Figure S9, Dataset S8**), suggesting that
406 when acting as confounders height, weight or BMI could lead to false positives in many
407 brain regions. We replicated 633 out of the 975 significant ROI-trait associations
408 ($p<0.05/975$, see **Dataset S9 for results on UKB replication sample**). Most replicated
409 associations were found with age, sex and body size variables, though we also replicated

410 associations between subcortical volumes and hand grip strength or time spent watching TV
411 (**Dataset S7-9**). Overall, some of the trait-ROIs associations were partially redundant as
412 indicated by a sum of R^2 (over ROIs) greater than the morphometricity (see **Appendix S9** for
413 detailed results and discussion, **Dataset S10** for results in HCP).

414

415 *6.5. Better cortical processing*

416 We compared the brain-morphometricity estimates obtained by varying the cortical
417 processing options: smoothing of the cortical meshes and applying coarser FreeSurfer
418 meshes. We found that applying smoothing (5-25mm) or reducing the cortical mesh
419 complexity always led to a lower point estimate of brain morphometricity in the UKB
420 discovery (**Figure 5**) and replication (**Figure S10, Datasets S11-12** for full tables) samples.
421 These differences were significant for a handful of variables (incl. age, sex, maternal
422 smoking or body size) using a stringent definition of significance based on overlapping
423 confidence intervals (**Table S3**). Thus, the fsaverage cortical mesh with no smoothing may
424 be deemed a better processing approach for at least some of the phenotypes considered.
425 Similarly, we found that the vertex-wise approach always yielded greater association R^2 ,
426 thus retained more information than a ROI based dimension reduction (**Figure S11**).

427

428 *6.6. Ten-fold cross-validation in the UKB and prediction into the UKB replication sample*

429 For each UKB participant, we calculated (BLUP) grey-matter scores relative to
430 phenotypes showing significant brain-morphometricity. As in sections above, for height,
431 weight and BMI we controlled for baseline covariates and further regressed out body size
432 for all other phenotypes.

433 In the 10-fold cross-validation analysis, most grey-matter scores significantly
434 correlated (positively) with their corresponding phenotypes (significance threshold of
435 $0.05/39=1.2e-3$, **Table 1, S3, Figure 4**). Albeit significant, prediction accuracy was overall low
436 (typically $r<0.10$, including $r=0.11$ for sex, $r<0.09$ with cognition, $r=0.08$ for alcohol intake,
437 $r=0.06$ with smoking status) except for age ($r=0.60$) and maternal smoking around birth
438 ($r=0.26$). We found similar prediction results in the UKB replication sample, with 29
439 associations reaching significance at $p<1.2e-3$ (**Table 1, S3**). Prediction accuracy into the UKB
440 replication sample was on par for most traits, though slightly greater for age and sex
441 compared to the cross-validation results (**Figure 4, Table 1, S3**). This is consistent with a
442 larger training sample being used and larger morphometricity observed in the replication
443 set (**Figure S6**).

444 When not correcting for body size, 56/58 BLUP scores significantly correlated with
445 the observed values in the 10-fold cross validation and 42 associations replicated using the
446 UKB replication sample ($p<0.05/58$, See **FigureS12** and **DatasetS13**). Predicted age
447 correlated with chronological age ($r=0.72$ in the discovery, $r=0.70$ in the replication), while
448 predicted sex also strongly associated with the observed value (AUC of 0.90 and 0.89). Grey-
449 matter scores of body shape (under the baseline covariates) were also significantly
450 correlated with the observed values ($r=0.25$ for height, $r=0.29$ for body fat percentage,
451 $r=0.39$ for weight and hip or waist circumference, $r=0.45$ for BMI). Finally, grey-matter
452 scores of BMI correlated positively with depression symptom count ($r=0.10$, $p\text{-value}<1e-14$),
453 as expected from the brain-morphometricity of depression being limited by the covariation
454 with body size. It even outperformed the grey-matter score built from the depression score
455 itself ($r=0.05$, $p\text{-value}<1e.5$).

456 BLUP achieved similar to superior prediction accuracy (R^2) compared to LASSO brain
457 scores across the 58 phenotypes (**Figure S13, Dataset S13**). BLUP significantly outperformed
458 LASSO (Wilcoxon test on absolute errors, $p < 0.05/58$) in predicting hip circumference,
459 alcohol intake and number of correct symbol matches (cognition).

460

461 *6.7. Out of sample prediction – application in the HCP sample*

462 Out of sample prediction validates that the morphometric associations are
463 generalizable to independent brain images, beyond population and scanner differences. For
464 traits only available in the UKB (e.g. waist circumference) we used a proxy in the HCP (e.g.
465 BMI). Grey matter scores for age, sex, and being a twin significantly correlated with the
466 observed values ($r_{\text{age}}=0.15$, $r_{\text{sex}}=0.25$, $r_{\text{twin-status}}=0.31$, **Table 1, Table S3** and **Figure 4**). Grey-
467 matter score for maternal smoking around birth correlated with smoking status ($r=0.19$).
468 None of the other grey-matter scores significantly correlated with a similar HCP variable.

469 Without correcting for body size, 19 BLUP scores correlated to corresponding
470 variables (**Dataset S13, Figure S12**). For example, scores for BMI, body fat percentage, hip or
471 waist circumference also correlated positively with BMI ($r=0.21$, $p\text{-value} < 1.2e-3$), while
472 scores for height and weight also correlated with the observed phenotypes ($r_{\text{Height}}=0.17$,
473 $r_{\text{Weight}}=0.19$). Finally, scores build from diet items or quantifying activity levels significantly
474 predicted BMI in the HCP.

475

476 **7. Discussion**

477 We have introduced a set of analyses, that rely on linear mixed models (LMMs, **Figure 1**) to
478 perform association and prediction, while being suited to tackle the challenges of big-data in
479 neuroimaging (Smith & Nichols, 2018). We have demonstrated their applications in two of
480 the largest MRI cohorts available for research (UKB (Miller et al., 2016) and HCP (Van Essen

481 et al., 2013)) using a fine-grained processing of anatomical MRI that consisted in >650,000
482 grey-matter measurements per individual. In LMMs, the overall effect of the high
483 dimensional vertex-wise measures is modelled by a single random effect, with a variance-
484 covariance structure calculated from the vertex-wise data: the brain relatedness matrix
485 (BRM, **Figure 1**). BRM off-diagonal elements represent the relative global similarity between
486 grey-matter structure of two people. The model is equivalent to fitting all vertices as a set of
487 random effects, constraining the association effect sizes to be normally distributed (**Figure**
488 **1**), which can be seen as an extension of multiple regression when the number of variables
489 exceeds the number of participants. This framework allows estimating new sample
490 characteristics such as the total association (morphometricity (Sabuncu et al., 2016))
491 between a phenotype and vertex-wise brain data or grey-matter correlations that quantify
492 how much phenotypes may be similarly associated with grey-matter. In addition, it offers to
493 build performant brain-based predictors that do not require hyper-parameter estimation.

494 Our analyses replicated and extended previous morphometricity reports (Sabuncu et
495 al., 2016) (**Figure 2, Dataset S3, S4**). We have demonstrated that our methods produce
496 robust, replicable results (**Figure S3, S6, S10, Table S1, S3**) that were partly transferrable on
497 a completely independent sample (the HCP) despite large differences between the samples
498 (**Table S2, Dataset S13**). We have shown additional utilities of this LMM framework such as
499 the ROI based association test that retained the vertex-wise complexity of a brain region
500 (**Fig S8-S8, Appendix S9, Dataset S7-S10**), rather than summarising them by a single average
501 measure, effectively bridging the gap between ROI/atlas based and vertex-wise analyses.
502 Our results aligned with previously published associations with sex (Ritchie et al., 2018), BMI
503 (Cole et al., 2013; Gupta et al., 2015; Kurth et al., 2013; Masouleh et al., 2016; Medic et al.,
504 2016; Opel et al., 2017) or substance use (Cardenas, Studholme, Gazdzinski, Durazzo, &

505 Meyerhoff, 2007; Gallinat et al., 2006; Gillespie et al., 2018; Hanlon et al., 2016; Pitel,
506 Segobin, Ritz, Eustache, & Beaunieux, 2015) (see details in **Appendix S9**). We showed
507 another application of LMMs for big-data neuroimaging: to compare the amount of
508 information retained by different MRI image processing. We found that using the most
509 complex cortical mesh (“fsaverage”) with no smoothing maximised the brain-
510 morphometricity across all phenotypes studied, though further statistical testing of the
511 difference is required (**Table S3**). This suggests there is meaningful information in fine
512 grained grey-matter data that is lost when performing local averages (via smoothing,
513 coarser mesh or average over a ROI). More work is needed to compare our surface-based
514 approach (Fischl, 2012) to volume based processing (Flandin & Friston, 2008), or evaluating
515 the putative added value of including the T2w image (on top of the T1w). To finish on
516 processing, in the UKB we combined vertex-wise data estimated from T1w and T1w+T2w
517 which is meant to improve grey-matter segmentation, though few studies quantified it
518 (Lindroth et al., 2019). Here, we confirmed a difference in cortical thickness between
519 processing groups (Lindroth et al., 2019)(**Figure S2**), though our data driven QC (**Appendix**
520 **S2**) excluded 80% of the 400-odd participants processed using T1w only (flagged as outliers).
521 We corrected for processing type in the analyses and the good replication of the UKB
522 associations (**Figure S3, S6, Table S1**) in addition to the out of sample prediction (**Figure 4**)
523 suggest that our results are robust.

524 Beyond the large morphometricity estimates found for age and sex, BMI, weight,
525 waist and hip circumference, and both passive and active smoking (**Figure 2, Dataset S3**), we
526 found many small(er) associations, with a wide array of phenotypes, including some more
527 unexpected ones (e.g. self-reported diet, being a twin, happiness with one’s health, blood
528 assay results (**Figure 2, Dataset S3**)). Such findings may echo the concerns raised by Smith

529 and Nichols about the presence of many (small) confounded associations in big-data
530 neuroimaging (Smith & Nichols, 2018). The fact that we replicated the morphometricity in
531 another UKB sample, does not completely rule out a confounding effect, as the same bias
532 (e.g. healthy bias in recruitment (Fry et al., 2017)) may be present.

533 We illustrated this concern using the example of body size (BMI, weight, height),
534 which showed large, replicated morphometricity and was available in both cohorts. We
535 evaluated its contribution to the reported morphometricity by performing conditional
536 analyses and bivariate LMMs. Both approaches yielded the same conclusions: a large
537 fraction of the morphometricity detected was attributable to body size (**Figure S7, Dataset**
538 **S3**). However, co-variation does not necessarily imply confounding (which requires
539 establishing direction of effects) but may instead point to intermediate phenotypes or arise
540 from the pervasive pleiotropy across the human phenome (Solovieff, Cotsapas, Lee, Purcell,
541 & Smoller, 2013). In addition, we are dealing with associations, meaning that (except for
542 exposures such as age or sex) the trait-vertex associations responsible for the
543 morphometricity may be a cause and/or a consequence of the phenotypes. For example,
544 there is a known association between BMI and depression, with evidence of pleiotropy, but
545 also of a causal effect of BMI on depression (Wray et al., 2018), though we do not know
546 which of BMI and grey-matter structure cause the other and cannot label body size a
547 confounder. Thus, a conservative interpretation is that morphometricity of the depression
548 score is limited to the shared variation with body size in the UKB (**Dataset S3, Figure S7**).
549 However, our findings shed a new light on previously published results, as even the largest
550 case-control international initiatives (e.g. ENIGMA-MDD (Schmaal, Hibar, et al., 2016;
551 Schmaal, Veltman, et al., 2016)) may reflect, at least in part, variance shared between
552 depression and BMI (Cole et al., 2013). More work is needed to understand body size

553 contribution to published results linking grey-matter anatomy to psychiatric disorders
554 (MDD, bipolar, schizophrenia and substance use are all associated with BMI (Luppino et al.,
555 2010; McElroy & Keck, 2012; Rajan & Menon, 2017; Saarni et al., 2009; Wray et al., 2018))
556 or sexually dimorphic traits (likely associated with height and weight). In addition, body size
557 may be differently associated with the phenome across countries or age groups, which may
558 limit the replication of findings and predictive abilities of body size dependent scores.
559 Finally, the possible confounding effects of body size are exacerbated in small case-control
560 samples, leading to increased chances of false positive associations (Button et al., 2013;
561 Ioannidis, 2005). Note that body size being associated to many brain regions (**Figure S9**), its
562 confounding effect could lead to widespread cortical or subcortical false positives.

563 A different example may be that of cheese intake, previously given as an absurd
564 example of putative association likely confounded by socio-economic status (Smith &
565 Nichols, 2018), and for which we found a significant (replicated) morphometricity, even
566 after correcting for body size (**Figure 2, Dataset S3**). Consistent with the hypothesis of a
567 confounded association, our bivariate analysis identified large r_{GM} ($r_{GM}=1$) between cheese
568 intake and household income level, or fluid intelligence (**Figure S7**), though the latter was
569 significant only when controlling for body-size (**Figure 3**). Thus, grey-matter correlation may
570 allow hypothesis-generation about the origin of the morphometricity signal and could help
571 better identifying putative confounders. If confirmed, such confounded morphometricity
572 would not translate into significant prediction into the general population (where IQ may
573 not be associated with cheese intake) or into another sample/country with different dietary
574 habits.

575 Beyond the confounding/mediating effect between phenotypes one should also be
576 wary of known MRI acquisition artefacts and confounds (Smith & Nichols, 2018). Here, we

577 focused on well-studied MRI modalities (T1w and T2w) which may be among the least
578 sensitive to artefacts, especially that acquisition was performed on a single MRI machine
579 and processed using standard image processing pipelines (Fischl, 2012). However, we
580 detected significant morphometricity for pulse rate, bone mineral density and indirect
581 measures of breathing rate/depth (Smith & Nichols, 2018), even after controlling for body
582 size (**Dataset S3**). In addition, the large grey-matter correlations between pulse rate and
583 overall health rating, and between forced expiratory volume and fluid intelligence (**Figure**
584 **3**), suggest they might indeed act as confounders. More work is needed to extend the list of
585 acquisition confounders studied (e.g. head motion), and more power is needed to detect
586 finer grained r_{GM} with the phenotypes of interest (**Appendix S7**). Finally, note that r_{GM} would
587 also capture correlated measurement errors between traits (e.g. when 2 traits are
588 associated with head motion).

589 Next, we constructed BLUP scores that estimate the random effects of the LMM and
590 demonstrated their predictive abilities in independent samples (UKB replication and to a
591 lower extent in the HCP, [**Figure 4, S12, Table 1, S2, Dataset S11**], which differs in term of
592 scanner and sample composition). In addition to its statistical properties (unbiased, best
593 predictor in the class of linear predictors), we demonstrated that BLUP achieved similar (to
594 greater) prediction accuracy compared to LASSO based prediction, while being more
595 computationally efficient than most traditional machine learning approaches as it does not
596 require hyper-parameter estimation. Note that prediction relates naturally to association
597 which is apparent from our model formulation (**Figure 1**). Thus, the morphometricity value
598 represents the upper asymptote achievable in linear prediction (**Figure 4**, (Dudbridge,
599 2013)); in addition, grey-matter correlation indicates when transfer learning is possible
600 between the 2 variables. The limited prediction accuracy currently prevents BLUP scores

601 being used in the clinical settings. However, they open the way to new analyses on samples
602 already collected, for which information was not or could not be collected. Further
603 application of our BLUP grey-matter scores include studying correlates of brain age or
604 predicted age difference (difference between predicted and chronological age) (Cole, 2017;
605 Cole et al., 2017; Liem et al., 2017).

606 Despite good statistical power in theory (**Appendix S7**), the low numbers for some of
607 the binary variables may explain the lack of associations found with psychiatry (e.g.
608 schizophrenia, ADHD), stresses, traumas (**Dataset S3**), which would have to be confirmed
609 using a larger UKB sample or case-control samples (see results in (Sabuncu et al., 2016)).
610 Similarly, a lot of the trait variance remains unaccounted for by the grey-matter structure
611 variation (**Figure 2**) which calls to study brain regions not extracted here (e.g. brain stem,
612 cerebellum), other processing options (e.g. volume based processing), or MRI images
613 (diffusion weighted, fMRI) to further characterise the phenotypes.

614 Our approach is suited to studying other MRI contrasts and even multiple MRI
615 modalities at once by fitting several random effect components (**Figure S4**). In addition, the
616 efficient implementation in the OSCA software (Zhang et al., 2019) means the analyses are
617 scalable to the future full UKB sample of 100,000 participants, which should improve power
618 and BLUP prediction accuracy (Dudbridge, 2013). Beyond the global or regional associations
619 reported here, future analyses should aim at identifying the vertices that contribute to the
620 morphometricity. An existing method is mass-univariate vertex-wise analysis, though this
621 comes as a huge increase of multiple testing burden and may still be underpowered with
622 the current sample sizes (Smith & Nichols, 2018).

623

624 **8. URLs**

625 Summary-level data (BLUP weights): https://cnsgenomics.com/content/data_and
626 <https://cloudstor.aarnet.edu.au/plus/s/T1gyJyQsF6wTMjF>; Code used for the analyses and
627 plots is downloadable at <https://github.com/baptisteCD/Brain-LMM> and viewable at
628 <https://baptistecd.github.io/Brain-LMM/index.html> ; OSCA:
629 <http://cnsgenomics.com/software/osca/> ; ENIGMA processing protocols:
630 <http://enigma.ini.usc.edu/protocols/imaging-protocols/> ;

631

632 **9. Author contributions**

633 PMV, NRW, JY and BCD designed the analyses. FZ and JY developed the OSCA
634 software. YH and BCD created the plots. KK, LY and ZZ assisted BCD with the UKB phenotypic
635 and genetic data, including download, formatting and curation. LS downloaded and
636 processed the HCP MRI images under MJW supervision. BCD downloaded and processed the
637 UKB MRI images. BCD performed the analyses and wrote the manuscript. All the authors
638 reviewed the manuscript.

639

640 **10. Competing financial Interests statement**

641 The authors declare no conflict of interests.

642

643 **References**

644 Achenbach, T. M. (2009). Achenbach system of empirically based assessment (ASEBA):
645 Development, findings, theory, and applications: University of Vermont, Research
646 Center of Children, Youth & Families.
647 Achenbach, T. M., Dumenci, L., & Rescorla, L. A. (2003). Ratings of Relations Between DSM-
648 IV Diagnostic Categories and Items of the Adult Self-Report (ASR) and Adult Behavior
649 Checklist (ABCL).
650 Ahlmann-Eltze, C. (2017). ggsignif: Significance Bars for 'ggplot2'. Retrieved from
651 <https://cran.r-project.org/package=ggsignif>
652 Allaire, J. X., Yihui.; McPherson, Jonathan.; Luraschi, Javier.; Ushey, Kevin.; Atkins, Aron.;
653 Wickham, Hadley.; Cheng, Joe.; Chang, Winston. . (2018). rmarkdown: Dynamic
654 Documents for R. Retrieved from <https://cran.r-project.org/package=rmarkdown>

655 Balduzzi, S., Rücker, G., & Schwarzer, G. (2019). How to perform a meta-analysis with R: a
656 practical tutorial. *Evidence Based Mental Health*, 22(4), 153. doi:10.1136/ebmental-
657 2019-300117

658 Bijma, P., & Bastiaansen, J. W. (2014). Standard error of the genetic correlation: how much
659 data do we need to estimate a purebred-crossbred genetic correlation? *Genetics
660 Selection Evolution*, 46(1), 79. doi:10.1186/s12711-014-0079-z

661 Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., &
662 Munafo, M. R. (2013). Power failure: why small sample size undermines the
663 reliability of neuroscience. *Nature Reviews Neuroscience*, 14(5), 365-376.
664 doi:10.1038/nrn3475

665 Buysse, D. J., Reynolds, C. F., 3rd, Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The
666 Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and
667 research. *Psychiatry Res*, 28(2), 193-213.

668 Cardenas, V. A., Studholme, C., Gazdzinski, S., Durazzo, T. C., & Meyerhoff, D. J. (2007).
669 Deformation-based morphometry of brain changes in alcohol dependence and
670 abstinence. *Neuroimage*, 34(3), 879-887. doi:10.1016/j.neuroimage.2006.10.015

671 Champely, S. (2017). pwr: Basic Functions for Power Analysis. Retrieved from [https://cran.r-](https://cran.r-project.org/package=pwr)
672 [project.org/package=pwr](https://cran.r-project.org/package=pwr)

673 Cole, J. H. (2017). Neuroimaging-derived brain-age: an ageing biomarker? *Aging (Albany
674 NY)*, 9(8), 1861-1862. doi:10.18632/aging.101286

675 Cole, J. H., Boyle, C. P., Simmons, A., Cohen-Woods, S., Rivera, M., McGuffin, P., . . . Fu, C. H.
676 (2013). Body mass index, but not FTO genotype or major depressive disorder,
677 influences brain structure. *Neuroscience*, 252, 109-117.
678 doi:10.1016/j.neuroscience.2013.07.015

679 Cole, J. H., Poudel, R. P. K., Tsagkrasoulis, D., Caan, M. W. A., Steves, C., Spector, T. D., &
680 Montana, G. (2017). Predicting brain age with deep learning from raw imaging data
681 results in a reliable and heritable biomarker. *Neuroimage*, 163, 115-124.
682 doi:10.1016/j.neuroimage.2017.07.059

683 Desikan, R. S., Segonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., . . . Killiany, R.
684 J. (2006). An automated labeling system for subdividing the human cerebral cortex
685 on MRI scans into gyral based regions of interest. *Neuroimage*, 31(3), 968-980.
686 doi:10.1016/j.neuroimage.2006.01.021

687 Dudbridge, F. (2013). Power and predictive accuracy of polygenic risk scores. *PLoS Genet*,
688 9(3), e1003348. doi:10.1371/journal.pgen.1003348

689 Fischl, B. (2012). FreeSurfer. *Neuroimage*, 62(2), 774-781.
690 doi:10.1016/j.neuroimage.2012.01.021

691 Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Segonne, F., Salat, D. H., . . . Dale, A.
692 M. (2004). Automatically parcellating the human cerebral cortex. *Cerebral Cortex*,
693 14(1), 11-22. doi:10.1093/cercor/bhg087

694 Flandin, G., & Friston, K. J. (2008). Statistical parametric mapping (SPM). *Scholarpedia*, 3(4),
695 6232.

696 Fry, A., Littlejohns, T. J., Sudlow, C., Doherty, N., Adamska, L., Sprosen, T., . . . Allen, N. E.
697 (2017). Comparison of Sociodemographic and Health-Related Characteristics of UK
698 Biobank Participants With Those of the General Population. *American Journal of
699 Epidemiology*, 186(9), 1026-1034. doi:10.1093/aje/kwx246

700 Gallinat, J., Meisenzahl, E., Jacobsen, L. K., Kalus, P., Bierbrauer, J., Kienast, T., . . . Staedtgen,
701 M. (2006). Smoking and structural brain deficits: a volumetric MR investigation.

702 *European Journal of Neuroscience*, 24(6), 1744-1750. doi:10.1111/j.1460-
703 9568.2006.05050.x

704 Gillespie, N. A., Neale, M. C., Bates, T. C., Eyer, L. T., Fennema - Notestine, C., Vassileva, J., .
705 . . Wright, M. J. (2018). Testing associations between cannabis use and subcortical
706 volumes in two large population - based samples. *Addiction*, 0(ja).
707 doi:doi:10.1111/add.14252

708 Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., . . .
709 Consortium, W. U.-M. H. (2013). The minimal preprocessing pipelines for the Human
710 Connectome Project. *Neuroimage*, 80, 105-124.
711 doi:10.1016/j.neuroimage.2013.04.127

712 Goddard, M. E., Wray, N. R., Verbyla, K., & Visscher, P. M. (2009). Estimating Effects and
713 Making Predictions from Genome-Wide Marker Data. *Statistical Science*, 24(4), 517-
714 529. doi:10.1214/09-Sts306

715 Gupta, A., Mayer, E. A., Sanmiguel, C. P., Van Horn, J. D., Woodworth, D., Ellingson, B. M., . .
716 . Labus, J. S. (2015). Patterns of brain structural connectivity differentiate normal
717 weight from overweight subjects. *Neuroimage-Clinical*, 7, 506-517.
718 doi:10.1016/j.nicl.2015.01.005

719 Gutman, B. A., Madsen, S. K., Toga, A. W., & Thompson, P. M. (2013). A Family of Fast
720 Spherical Registration Algorithms for Cortical Shapes. In L. Shen, T. Liu, P.-T. Yap, H.
721 Huang, D. Shen, & C.-F. Westin (Eds.), *Multimodal Brain Image Analysis: Third
722 International Workshop, MBIA 2013, Held in Conjunction with MICCAI 2013, Nagoya,
723 Japan, September 22, 2013, Proceedings* (pp. 246-257). Cham: Springer International
724 Publishing.

725 Gutman, B. A., Wang, Y. L., Rajagopalan, P., Toga, A. W., & Thompson, P. M. (2012). Shape
726 Matching with Medial Curves and 1-D Group-Wise Registration. *2012 9th IEEE
727 International Symposium on Biomedical Imaging (Isbi)*, 716-719.

728 Hanlon, C. A., Owens, M. M., Joseph, J. E., Zhu, X., George, M. S., Brady, K. T., & Hartwell, K.
729 J. (2016). Lower subcortical gray matter volume in both younger smokers and
730 established smokers relative to non-smokers. *Addiction Biology*, 21(1), 185-195.
731 doi:10.1111/adb.12171

732 Hanscombe, K. (2017). ukbtools: Manipulate and Explore UK Biobank Data. Retrieved from
733 kenhanscombe/ukbtools (github)

734 Harrell, F. E. J. (2017). Hmisc: Harrell Miscellaneous. Retrieved from [https://cran.r-](https://cran.r-project.org/package=Hmisc)
735 [project.org/package=Hmisc](https://cran.r-project.org/package=Hmisc)

736 Henderson, C. R. (1950). Estimation of Genetic Parameters. *Annals of Mathematical
737 Statistics*, 21(2), 309-310.

738 Henderson, C. R. (1975). Best Linear Unbiased Estimation and Prediction under a Selection
739 Model. *Biometrics*, 31(2), 423-447. doi:Doi 10.2307/2529430

740 Holtz, Y. (2020). epuRate: A clean template for R Markdown documents. Retrieved from
741 <https://github.com/holtzy/epuRate>

742 Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Med*, 2(8), e124.
743 doi:10.1371/journal.pmed.0020124

744 Kurth, F., Levitt, J. G., Phillips, O. R., Luders, E., Woods, R. P., Mazziotta, J. C., . . . Narr, K. L.
745 (2013). Relationships between gray matter, body mass index, and waist
746 circumference in healthy adults. *Human Brain Mapping*, 34(7), 1737-1746.
747 doi:10.1002/hbm.22021

- 748 Lee, S. H., Yang, J., Goddard, M. E., Visscher, P. M., & Wray, N. R. (2012). Estimation of
749 pleiotropy between complex diseases using single-nucleotide polymorphism-derived
750 genomic relationships and restricted maximum likelihood. *Bioinformatics*, *28*(19),
751 2540-2542. doi:10.1093/bioinformatics/bts474
- 752 Liem, F., Varoquaux, G., Kynast, J., Beyer, F., Kharabian Masouleh, S., Huntenburg, J. M., . . .
753 Margulies, D. S. (2017). Predicting brain-age from multimodal imaging data captures
754 cognitive impairment. *Neuroimage*, *148*, 179-188.
755 doi:<https://doi.org/10.1016/j.neuroimage.2016.11.005>
- 756 Lindroth, H., Nair, V. A., Stanfield, C., Casey, C., Mohanty, R., Wayer, D., . . . Sanders, R. D.
757 (2019). Examining the identification of age-related atrophy between T1 and T1 + T2-
758 FLAIR cortical thickness measurements. *Sci Rep*, *9*(1), 11288. doi:10.1038/s41598-
759 019-47294-2
- 760 Luppino, F. S., de Wit, L. M., Bouvy, P. F., Stijnen, T., Cuijpers, P., Penninx, B. W. J. H., &
761 Zitman, F. G. (2010). Overweight, Obesity, and Depression A Systematic Review and
762 Meta-analysis of Longitudinal Studies. *Archives of General Psychiatry*, *67*(3), 220-229.
763 doi:DOI 10.1001/archgenpsychiatry.2010.2
- 764 Marcus, D. S., Harms, M. P., Snyder, A. Z., Jenkinson, M., Wilson, J. A., Glasser, M. F., . . .
765 Consortium, W. U.-M. H. (2013). Human Connectome Project informatics: quality
766 control, database services, and data visualization. *Neuroimage*, *80*, 202-219.
767 doi:10.1016/j.neuroimage.2013.05.077
- 768 Marcus, D. S., Harwell, J., Olsen, T., Hodge, M., Glasser, M. F., Prior, F., . . . Van Essen, D. C.
769 (2011). Informatics and data mining tools and strategies for the human connectome
770 project. *Front Neuroinform*, *5*, 4. doi:10.3389/fninf.2011.00004
- 771 Masouleh, S. K., Arelin, K., Horstmann, A., Lampe, L., Kipping, J. A., Luck, T., . . . Witte, A. V.
772 (2016). Higher body mass index in older adults is associated with lower gray matter
773 volume: implications for memory performance. *Neurobiology of Aging*, *40*, 1-10.
774 doi:10.1016/j.neurobiolaging.2015.12.020
- 775 McElroy, S. L., & Keck, P. E. (2012). Obesity in Bipolar Disorder: An Overview. *Current*
776 *Psychiatry Reports*, *14*(6), 650-658. doi:10.1007/s11920-012-0313-8
- 777 Medic, N., Ziauddeen, H., Ersche, K. D., Farooqi, I. S., Bullmore, E. T., Nathan, P. J., . . .
778 Fletcher, P. C. (2016). Increased body mass index is associated with specific regional
779 alterations in brain structure. *International Journal of Obesity*, *40*(7), 1177-1182.
780 doi:10.1038/ijo.2016.42
- 781 Miller, K. L., Alfaro-Almagro, F., Bangerter, N. K., Thomas, D. L., Yacoub, E., Xu, J., . . . Smith,
782 S. M. (2016). Multimodal population brain imaging in the UK Biobank prospective
783 epidemiological study. *Nature Neuroscience*, *19*(11), 1523-1536.
784 doi:10.1038/nn.4393
- 785 Opel, N., Redlich, R., Kaehler, C., Grotegerd, D., Dohm, K., Heindel, W., . . . Dannlowski, U.
786 (2017). Prefrontal gray matter volume mediates genetic risks for obesity. *Molecular*
787 *Psychiatry*, *22*(5), 703-710. doi:10.1038/mp.2017.51
- 788 Patterson, H. D., & Thompson, R. (1971). Recovery of Inter-Block Information when Block
789 Sizes are Unequal. *Biometrika*, *58*(3), 545-554. doi:10.2307/2334389
- 790 Pitel, A. L., Segobin, S. H., Ritz, L., Eustache, F., & Beaunieux, H. (2015). Thalamic
791 abnormalities are a cardinal feature of alcohol-related brain dysfunction. *Neurosci*
792 *Biobehav Rev*, *54*, 38-45. doi:10.1016/j.neubiorev.2014.07.023

793 Privé, F., Aschard, H., Ziyatdinov, A., & Blum, M. G. B. (2018). Efficient analysis of large-scale
794 genome-wide data with two R packages: bigstatsr and bigsnpr. *Bioinformatics*
795 (*Oxford, England*), 34(16), 2781-2787. doi:10.1093/bioinformatics/bty185

796 R Development Core Team. (2012). R: A Language and Environment for Statistical
797 Computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from
798 <http://www.r-project.org/>

799 Rajan, T. M., & Menon, V. (2017). Psychiatric disorders and obesity: A review of association
800 studies. *Journal of Postgraduate Medicine*, 63(3), 182-190.
801 doi:10.4103/jpgm.JPGM_712_16

802 Ritchie, S. J., Cox, S. R., Shen, X., Lombardo, M. V., Reus, L. M., Alloza, C., . . . Deary, I. J.
803 (2018). Sex Differences in the Adult Human Brain: Evidence from 5216 UK Biobank
804 Participants. *Cereb Cortex*, 28(8), 2959-2975. doi:10.1093/cercor/bhy109

805 Robinson, G. K. (1991). That BLUP is a Good Thing: The Estimation of Random Effects.
806 *Statistical Science*, 6(1), 15-32.

807 Robinson, M. R., Kleinman, A., Graff, M., Vinkhuyzen, A. A. E., Couper, D., Miller, M. B., . . .
808 Visscher, P. M. (2017). Genetic evidence of assortative mating in humans. 1, 0016.
809 doi:10.1038/s41562-016-0016

810 Rosenberg, M. D., Casey, B. J., & Holmes, A. J. (2018). Prediction complements explanation
811 in understanding the developing brain. *Nat Commun*, 9(1), 589. doi:10.1038/s41467-
812 018-02887-9

813 Roshchupkin, G. V., Gutman, B. A., Vernooij, M. W., Jahanshad, N., Martin, N. G., Hofman,
814 A., . . . Adams, H. H. H. (2016). Heritability of the shape of subcortical brain
815 structures in the general population. *Nat Commun*, 7, 13738.
816 doi:10.1038/ncomms13738

817 Saarni, S. E., Saarni, S. I., Fogelholm, M., Heliövaara, M., Perala, J., Suvisaari, J., & Lonnqvist,
818 J. (2009). Body composition in psychotic disorders: a general population survey.
819 *Psychological Medicine*, 39(5), 801-810. doi:10.1017/S0033291708004194

820 Sabuncu, M. R., Ge, T., Holmes, A. J., Smoller, J. W., Buckner, R. L., Fischl, B., & Initia, A. D. N.
821 (2016). Morphometricity as a measure of the neuroanatomical signature of a trait.
822 *Proceedings of the National Academy of Sciences of the United States of America*,
823 113(39), E5749-E5756. doi:10.1073/pnas.1604378113

824 Schmaal, L., Hibar, D. P., Samann, P. G., Hall, G. B., Baune, B. T., Jahanshad, N., . . . Veltman,
825 D. J. (2016). Cortical abnormalities in adults and adolescents with major depression
826 based on brain scans from 20 cohorts worldwide in the ENIGMA Major Depressive
827 Disorder Working Group. *Mol Psychiatry*. doi:10.1038/mp.2016.60

828 Schmaal, L., Veltman, D. J., van Erp, T. G., Samann, P. G., Frodl, T., Jahanshad, N., . . . Hibar,
829 D. P. (2016). Subcortical brain alterations in major depressive disorder: findings from
830 the ENIGMA Major Depressive Disorder working group. *Mol Psychiatry*, 21(6), 806-
831 812. doi:10.1038/mp.2015.69

832 Smith, S. M., & Nichols, T. E. (2018). Statistical Challenges in "Big Data" Human
833 Neuroimaging. *Neuron*, 97(2), 263-268. doi:10.1016/j.neuron.2017.12.018

834 Solovieff, N., Cotsapas, C., Lee, P. H., Purcell, S. M., & Smoller, J. W. (2013). Pleiotropy in
835 complex traits: challenges and strategies. *Nat Rev Genet*, 14(7), 483-495.
836 doi:10.1038/nrg3461

837 Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., . . . Collins, R. (2015). UK
838 biobank: an open access resource for identifying the causes of a wide range of

839 complex diseases of middle and old age. *PLoS Med*, 12(3), e1001779.
840 doi:10.1371/journal.pmed.1001779

841 Temple, D. L., & the CRAN Team, R. (2017). XML: Tools for Parsing and Generating XML
842 Within R and S-Plus. Retrieved from <https://cran.r-project.org/package=XML>

843 Thompson, R. (1973). Estimation of Variance and Covariance Components with an
844 Application When Records Are Subject to Culling. *Biometrics*, 29(3), 527-550. doi:Doi
845 10.2307/2529174

846 Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal*
847 *Statistical Society. Series B (Methodological)*, 58(1), 267-288.

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

850 Tyrrell, J., Jones, S. E., Beaumont, R., Astley, C. M., Lovell, R., Yaghootkar, H., . . . Frayling, T.
851 M. (2016). Height, body mass index, and socioeconomic status: mendelian
852 randomisation study in UK Biobank. *Bmj-British Medical Journal*, 352. doi:ARTN i582
853 10.1136/bmj.i582

854 Van Essen, D. C., Glasser, M. F., Dierker, D. L., Harwell, J., & Coalson, T. (2012). Parcellations
855 and hemispheric asymmetries of human cerebral cortex analyzed on surface-based
856 atlases. *Cereb Cortex*, 22(10), 2241-2262. doi:10.1093/cercor/bhr291

857 Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., Ugurbil, K., &
858 Consortium, W. U.-M. H. (2013). The WU-Minn Human Connectome Project: an
859 overview. *Neuroimage*, 80, 62-79. doi:10.1016/j.neuroimage.2013.05.041

860 Van Essen, D. C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T. E., Bucholz, R., . . .
861 Consortium, W. U.-M. H. (2012). The Human Connectome Project: a data acquisition
862 perspective. *Neuroimage*, 62(4), 2222-2231. doi:10.1016/j.neuroimage.2012.02.018

863 Varrichio, C. (2016). rowr: Row-Based Functions for R Objects. Retrieved from [https://cran.r-](https://cran.r-project.org/package=rowr)
864 [project.org/package=rowr](https://cran.r-project.org/package=rowr)

865 Vilhjalmsson, B. J., Yang, J., Finucane, H. K., Gusev, A., Lindstrom, S., Ripke, S., . . . Inherited,
866 D. B. R. (2015). Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk
867 Scores. *American Journal of Human Genetics*, 97(4), 576-592.
868 doi:10.1016/j.ajhg.2015.09.001

869 Visscher, P. M. (1998). On the sampling variance of intraclass correlations and genetic
870 correlations. *Genetics*, 149(3), 1605-1614.

871 Visscher, P. M., Hemani, G., Vinkhuyzen, A. A. E., Chen, G. B., Lee, S. H., Wray, N. R., . . .
872 Yang, J. (2014). Statistical Power to Detect Genetic (Co)Variance of Complex Traits
873 Using SNP Data in Unrelated Samples. *Plos Genetics*, 10(4). doi:ARTN e1004269
874 10.1371/journal.pgen.1004269

875 Wei, T., & Simko, V. (2017). R package "corrplot": Visualization of a Correlation Matrix.
876 Retrieved from <https://github.com/taiyun/corrplot>

877 Wickham, H. (2007). Reshaping data with the reshape package. *Journal of Statistical*
878 *Software*, 21(12), 1-20.

879 Wickham, H. (2009). *Elegant Graphics for Data Analysis*. New York: Springer-Verlag.

880 Wickham, H. (2015). scales: Scale functions for graphics. Retrieved from [http://cran.r-](http://cran.r-project.org/package=scales)
881 [project.org/package=scales](http://cran.r-project.org/package=scales)

882 Wickham, H. (2017). tidyverse: Easily Install and Load 'Tidyverse' Packages. Retrieved from
883 <https://cran.r-project.org/package=tidyverse>

884 Wickham, H., & Francois, R. (2015). dplyr: A Grammar of Data Manipulation.

885 Wickham, H. H., J.; Francois, R. (2017). readr: Read Rectangular Text Data. Retrieved from
886 <https://cran.r-project.org/package=readr>
887 Wray, N. R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E. M., Abdellaoui, A., . . .
888 Major Depressive Disorder Working Group of the Psychiatric Genomics, C. (2018).
889 Genome-wide association analyses identify 44 risk variants and refine the genetic
890 architecture of major depression. *Nat Genet*, 50(5), 668-681. doi:10.1038/s41588-
891 018-0090-3
892 Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., . . . Visscher,
893 P. M. (2010). Common SNPs explain a large proportion of the heritability for human
894 height. *Nat Genet*, 42(7), 565-569. doi:10.1038/ng.608
895 Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: a tool for genome-wide
896 complex trait analysis. *American Journal of Human Genetics*, 88(1), 76-82.
897 doi:10.1016/j.ajhg.2010.11.011
898 Zhang, F., Chen, W., Zhu, Z., Zhang, Q., Nabais, M. F., Qi, T., . . . Yang, J. (2019). OSCA: a tool
899 for omic-data-based complex trait analysis. *Genome Biology*, 20(1), 107.
900 doi:10.1186/s13059-019-1718-z
901

902 **Table 1:** Summary of the prediction accuracy (R^2) of the BLUP grey-matter scores.

903 *We constructed BLUP scores for the 39 UKB variables showing significant morphometricity and evaluated their predictive power in the UKB (10*
904 *fold-cross validation) and HCP sample. When the phenotype corresponding to the grey-matter score was not available in the HCP, we chose the*
905 *closest available (e.g. waist circumference grey-matter score evaluated against BMI). We evaluate the prediction accuracy by fitting GLM*
906 *controlling for height, weight and BMI as well as for the baseline covariates (acquisition, age, sex and head size); except for (#) denoting*
907 *associations not controlling for height, weight and BMI. Rows in bold indicate significant association after correcting for multiple testing*
908 *($p < 0.05/39 = 1.3e-3$) both in and out of sample. This reduced table only shows prediction results significant in all 3 scenarios, see **Table S2** for full*
909 *table of results. We reported the AUC (for discrete variables) as it is independent of the proportion of twins and males, thus differences in AUC*
910 *likely reflect differences in morphometricity between the UKB and HCP samples.*

	In sample prediction (UKB)				Prediction into UKB replication				Out of sample prediction (HCP)				
	r	p-value	R^2	AUC (SE)	r	p-value	R^2	AUC (SE)	HCP variable predicted	r	p-value	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-318	0.062		0.23	2.6e-132	0.054		Height	0.17	1.8e-17	0.03	
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	

911

912 **Figure 1: Summary of the analyses and methods**

913 The data used in the analyses are detailed at the top, and include phenotypes, covariates and vertex-
914 wise measurements. Calculation of the brain relatedness matrix from the vertex-wise data is also
915 described. The two LMM formulations are equivalent, though the first one emphasises how
916 morphometricity relates to the brain relatedness matrix and the second emphasises the joint effects
917 of all vertex-wise measurements. The boxes at the bottom provide brief interpretations of the new
918 methods and concepts. Morphometricity estimate may be restricted to vertex-wise measurements
919 within a ROI, to estimate the phenotype-ROI association that takes into account all local variations
920 within the ROI. Morphometricity may also be used to compare the information retained by different
921 MRI image processing (“Best processing”). Grey-matter correlation measures the shared
922 morphometricity between traits, or in other words the correlation between the vertex-phenotypes
923 associations for each phenotype. BLUP scores are predictors of the random effect, and make the
924 hypothesis that the vertex-phenotypes marginal joint associations weights are normally distributed.
925 BLUP prediction, expressed in proportion of phenotypic variance, is capped by the morphometricity.

926
927 **Figure 2: Circular barplot of the associations (R^2) between phenotypes and grey-matter structure**
928 **vertices (morphometricity)**

929 For clarity, we only plotted the significant associations in the UKB discovery (panel **a**) and HCP
930 sample (panel **b**). We applied Bonferroni correcting to account for multiple testing in each sample.
931 The black bars represent the 95% confidence intervals of the morphometricity estimates. For
932 context, we also present the association R^2 between phenotypes and covariates of the baseline
933 model, as per the legend under the barplot. As some covariates may be correlated, the R^2 was
934 calculated by adding progressively the covariates in that order: acquisition and processing variables
935 (labelled “other”), age, sex and head size (ICV, total cortical thickness and surface area). Age and sex
936 were not included as covariates when studying them as phenotypes. See **Dataset S3-4** for full
937 results. See **Figure S1** for positive control associations.

938 **Figure 3: Matrices of grey-matter correlations (upper diagonals) and residual correlations (lower**
939 **diagonals) between all the variables showing significant morphometricity after controlling for**
940 **baseline covariates, as well as height, weight and BMI.**

941 Panel (a) shows the results for the UKB and panel (b) the HCP results. We excluded phenotypes used
942 as covariates (age, sex, head and body size) as regressing them out makes them orthogonal (i.e. not
943 associated) with the remaining traits. We used conservative significance thresholds of
944 $0.05/(35*34)=4.2e-5$ for UKB and $0.05/(18*17)=1.6e-4$ for HCP that account for the total number of
945 correlations performed in each sample. Correlations significant after multiple testing correction are
946 indicated by a star. Blocks circled in black indicate the different phenotype categories used
947 previously (see **Figure 2**). r_{GM} is a measure of the shared brain-morphometricity between 2 traits.
948 Contrasting r_{GM} and residual correlation (r_E) indicate how much of the phenotypic co-variance is
949 attributable to individual's resemblance in term of grey-matter structure vs. other factors (brain or
950 non-brain resemblances).

951

952 **Figure 4: In sample and out of sample prediction accuracy as a function of the total association R^2 .**

953 Labels highlight some of the significant prediction having the greatest accuracy. As predicted by the
954 theory, the prediction accuracy is capped by the total association R^2 (points below the diagonal). We
955 limited the prediction analysis to phenotypes showing a significant brain-morphometricity in the
956 UKB discovery sample.

957

958 **Figure 5: Comparison of brain-morphometricity estimates varying cortical processing options in**
959 **FreeSurfer.**

960 The reduction of brain-morphometricity as a function of mesh smoothing is presented on the left
961 panel (a), while the right panel (b) shows the effect of reducing the cortical mesh complexity. The
962 black bar indicates the lower bound of the 95% confidence interval of the fsaverage-no smoothing
963 estimate (identical to results presented in **Figure2**, except that covariates R^2 are not plotted here to

964 simplify the figure). Brain-morphometricity estimates below the 95%CI lower bound cannot be
965 deemed significantly lower. Rather the 95%CI are presented for context and to remind that all
966 estimate from **Figure 2** do not have the same SE. See **Table S3** for a conservative list of phenotypes
967 with significantly reduced morphometricity compared to fsaverage – no smoothing.
968