nature portfolio

Corresponding author(s): Warrier

Last updated by author(s): Jun 20, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Data for this project came from the UK Biobank and ABCD datasets
Data analysis	We used standard Freesurfer workflow to our imaging phenotypes. These are described in https://github.com/ucam-department-of-psychiatry/UKB; https://github.com/ucam-department-of-psychiatry/ABCD. Genetic quality control has been described in warrier/ABCD_geneticQC (github.com); vwarrier/Imaging_genetics_analyses (github.com). Softwares used include: FastGWA v 1.93: GCTA | Yang Lab (westlake.edu.cn); Plink 1.9 and Plink 1.2: PLINK 1.9 (cog-genomics.org); genomic-SEM v 0.03d: GitHub - GenomicSEM/GenomicSEM: R-package for structural equation modeling based on GWAS summary data; LDSC v1.01: bulik/ldsc: LD Score Regression (LDSC) (github.com); Two sample MR (v 0.4.26): Two Sample MR functions and interface to MR Base database • TwoSampleMR (mrcieu.github.io); MR PResso: GitHub - rondolab/MR-PRESSO: Performs the Mendelian Randomization Pleiotropy RESidual Sum and Outlier (MR-PRESSO) method.; CAUSE (v 1.2): Introduction to CAUSE (jean997.github.io); MR power calculator: Online sample size and power calculator for Mendelian randomization with a binary outcome (shinyapps.io); MAGMA (v 1. 10): https://ctg.cncr.nl/software/magma; HMAGMA (using MAGMA v 1. 08): GitHub - thewonlab/H-MAGMA; GCTB (SBayesS v 2.0): GCTB (cnsgenomics.com); Polyfun (installed November 2021): GitHub - omerwe/polyfun: PolyFun (POLYgenic FUNctionally-informed fine-mapping); SMR: SMR | Yang Lab (westlake.edu.cn); Hyprcoloc: GitHub - jrs95/hyprcoloc: Hypothesis prioritisation in multi-trait colocalization; PRScs: GitHub - getian107/PRScs: Polygenic prediction via continuous shrinkage priors; R statistics packages: nlme, stats, Psychmeta, lmer, effsize, lsa, ape, factoExtra, igraph; R plotting packages: ggseg, ggplot2, tidyverse, ggpubr, ggstatsplot, ComplexHeatmap, scico, karyoploteR, tidySEM, circlize

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All summary statistics for the GWAS meta-analyses will be made available upon publication without restriction on https://portal.ide-cam.org.uk/overview/483. Data from the UK Biobank and ABCD can be applied for and accessed by approved researchers. GWAS summary statistics for other brain imaging phenotypes can be obtained from: The Oxford Brain Imaging Genetics PheWeb (PheWeb (ox.ac.uk)), GWAS catalog (GWAS Catalog (ebi.ac.uk)), GWAS ATLAS (Genome wide association study ATLAS (ctglab.nl)), and Brain Imaging Genetics Knowledge Portal Brain Imaging Genetics Summary Statistics. The SPARK dataset can be obtained by application to SFARIbase (SFARI | SFARI Base). The DDD dataset can be obtained via EGA (Deciphering Developmental Disorders (DDD) - EGA European Genome-Phenome Archive (ega-archive.org)).

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	We use the term sex throughout, and in our study this refers to genetic sex, determined by the composition of sex chromosomes.
Reporting on race, ethnicity, or other socially relevant groupings	We restricted our analyses to individuals of predominantly european ancestry as inferred by genetic principal components. This is also referred to as genetically referred ancestry.
Population characteristics	Population characteristics are mentioned in the Methods. We used participants from two widely used cohorts - UK Biobank and ABCD. UK Biobank participants were aged between 40 and 85 with a mean age of 62. ABCD participants were aged between 9 and 10.
Recruitment	Participants were recruited into ABCD and UK Biobank by other researchers. Details of ow participants have been recruited have been provided in papers referenced in the Methods section.
Ethics oversight	This research complies with all relevant ethical regulations. Ethical procedures for the UK Biobank are controlled by the Ethics and Guidance council (http://www.ukbiobank.ac.uk/ethics), and the study was conducted in accordance with the UK Biobank an Ethics and Governance Framework document (http://www.ukbiobank.ac.uk/wp-content/uploads/2011/05/EGF20082.pdf), with institutional review board approval by the North West Multi-center Research Ethics Committee. Ethical approval for ABCD was obtained from multiple institutional review boards.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No power calculations were conducted. We used all data where participants passed quality control, resulting in a final sample size of 36,663 individuals.
Data exclusions	We excluded individuals of non-European genetically inferred ancestries (beyond 5 standard deviations from the mean of the first two genetic principal components). These individuals were excluded due to the small sample size, the lack of methods that can run fast mixed-models based GWAS, and the poor registration with neuroimaging data. We further excluded people with sex mismatches, with low genotyping rate, and excess heterozygosity. We also excluded individuals who were more than 5 standard deviations of median absolute deviations from the mean or median of the imaging phenotypes. We included all genoytped and imputed SNPs with a minor allele frequency > 0l.1% and with an imputation R2 > 0.4. We excluded regions "52", "PI", and "PHA2" for ICI and FI because of no variance.
Replication	We validated the top findings from the UKB cohort in ABCD exactly once. We identified high genetic correlations between the cohorts, and consistent effect direction. For 237 GWAS loci in UKB for which data were available in ABCD, 204 SNPs {86%) had concordant sign of genetic

	association (p < 0.001, two-tailed binomial sign test) with modest positive correlation of effect size (r=0.54, 95% confidence interval [CI] 0.45-0.63. 75 {31%) of these SNPs had p-values (p)<0.05 in ABCD with concordant effect direction.
Randomization	No randomisation as all data were from observational studies.
Blinding	Blinding is not relevant as this is an association study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☐	☒ MRI-based neuroimaging

Magnetic resonance imaging

Experimental design

Design type	tructural minimally processed Tl and T2-FLAIR weighted data was obtained from UK BioBank (application 20904) and the ABCD study (via the NIH Data Archive Repository).
Design specifications	Details can be found here: https://www.nature.com/articles/nn.4393 and https://www.nature.com/articles/s41586-022-04554-y Further details are provided in the Methods section of the Study.
Behavioral performance measures	N/A

Acquisition

Imaging type(s)	Structural minimally processed Tl and T2-FLAIR weighted data was obtained from UK BioBank (application 20904) and the ABCD study (via the NIH Data Archive Repository).
Field strength	3T in UKB, at least 3T in ABCD.
Sequence & imaging parameters	Details can be found here: https://www.nature.com/articles/nn.4393 and https://www.nature.com/articles/s41586-022-04554-y Further details are provided in the Methods section of the Study.
Area of acquisition	Whole brain

Diffusion MRI ☐ Used ☐ Not used

Preprocessing

Preprocessing software

We used FreeSurfer workflows for structural image processing and AMICO for quantification of microstructural imaging features. Details of all processing are provided in the methods section: Data was acquired as part of the UK Biobank and ABCD cohort studies with the following protocols. For the UK Biobank (https://www.fmrib.ox.ac.uk/ukbiobank/protocol/V4_23092014.pdf), T1-weighted structural imaging was obtained using the following parameters: 1.0mm isotropic resolution, TR = 2000ms, TE = 2.01ms, TI = 880ms Flip angle 8 degrees. T2-weighted FLAIR structural imaging was obtained using the following parameters: 1.0x1.0x1.1mm resolution, TR = 5000ms, TE = 395.0 ms, TI = 1800 ms. Diffusion weighted imaging - 2.0mm isotropic resolution was obtained using the following parameters. MB = 3, R = 1, TE/TR = 92/3600 ms, PF 6/8, fat sat, b = 0 s/mm2 (5x + 3× phase-encoding reversed), b = 1 000 s/mm2 (50×), b = 2000 s/mm2 (50×), 105 + 6 time-points (PA-AP). For ABCD (https://github.com/nih-fmrif/abcd_protocols) and ref(Casey et al. 2018), T1-weighted - 1.0mm isotropic resolution: TR = 2500ms, TE = 2.88ms, TI = 1060ms, Flip angle 8 degrees; T2-weighted: 1.0mm isotropic resolution, TR = 3200ms, TE = 565ms, Flip angle variable; DWI: 1.7mm isotropic resolution, TR = 4100ms, TE = 88ms, Flip angle 90 degrees, 500 (6-dirs); 1000; (15-dirs) 2000; (15-dirs); 3000 (60-dirs)

While not processed as part of the present analysis we also obtained framewise displacement parameters from each individual's accompanying resting-state fMRI scan.

Structural minimally processed T1 and T2-FLAIR weighted data was obtained from UK BioBank (application 20904) and the ABCD study (via the NIH Data Archive Repository). These images were preprocessed with FreeSurfer (v6.0.1) 81 using the T2-FLAIR weighted image to improve pial surface reconstruction when available. Recon-all reconstruction included bias field correction, registration to stereotaxic space, intensity normalisation, skull-stripping, and white matter segmentation. A triangular surface tessellation fitted a deformable mesh model onto the white matter volume, providing grey-white and pial surfaces with >160,000 corresponding vertices registered to fsaverage standard space. When no T2-FLAIR image was available FreeSurfer reconstruction was done using the T1 weighted image only. Given systematic variation related to the inclusion of T2-FLAIR , this was included as a confound variable in downstream analyses. Cortical surfaces were reconstructed for each individual using FreeSurfer and registered using FreeSurfer's surface-based registration to fsaverage. The HCP's multi-modal parcellation v1.0 17 was resampled from fs_LR to fsaverage using existing transformations 82 and from there back to the individual's surface meshes based on the FreeSurfer folding-based surface registration. Reconstruction quality was assessed using the Euler index 83 and included as a covariate in subsequent analyses (Supplementary Note 4, SF 16).

Structural diffusion weighted imaging was obtained in processed form from UK BioBank and ABCD in similar fashion. As described in the UK Biobank Brain Imaging Documentation (v1.8) 84, UKBiobank diffusion images were corrected for eddy currents, head motion and outlier-slices using Eddy tool (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/EDDY). EPI correction was performed using a field map estimated from 3 b=0 images with standard (Anterior-Posterior) phase encoding and 3 b=0 images acquired with reversed phase encoding. Similarly, ABCD diffusion images were corrected for eddy currents (12 free parameters), head motion (rigid-body-registration) and EPI distortion using pairs of b=0 with opposite phase encoding polarities 85. Neurite Orientation Dispersion and Density Indices (NODDI) parameters were estimated using the Accelerated Microstructure Imaging via Convex Optimization (AMICO) 86 processing approach from the minimally processed diffusion images. The subject specific T1 aligned (based on surface alignment procedures) parcellation template was co-registered to the diffusion weighted image using fsl FLIRT and regional values for FA, MD and the three NODDI parameters were extracted using AFNI's 3dROIstats function for all of the 360 cortical regions included in the Human Connectome Parcellation and averaged across hemisphere to reduce the number of regions to 180 bilateral regions. We also evaluated a direct surface-based registration approach in line with HCP protocols for surface based registration (see Supplementary note 4b).

Normalization

Recon-all reconstruction included bias field correction, registration to stereotaxic space, intensity normalisation, skull-stripping, and white matter segmentation. A triangular surface tessellation fitted a deformable mesh model onto the white matter volume, providing grey-white and pial surfaces with >160,000 corresponding vertices registered to fsaverage standard space.

Normalization template

FSAverage

Noise and artifact removal

N/A

Volume censoring

N/A

Statistical modeling & inference

Model type and settings

N/A (see method on GWAS association)

Effect(s) tested

N/A (see method on GWAS association)

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☒ Both

Anatomical location(s)

All imaging derived phenotypes were extracted for 180 bilateral cortical regions based on the HCP parcellation scheme.

Statistic type for inference

(See Eklund et al. 2016)

N/A (see method on GWAS association)

Correction

N/A (see method on GWAS association)

Models & analysis

n/a	Involved in the study
☒	☐ Functional and/or effective connectivity
☒	☐ Graph analysis
☒	☐ Multivariate modeling or predictive analysis