nature portfolio

Corresponding author(s): David J. Glass

Last updated by author(s): Jan 14, 2026

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 was obtained from UK Biobank (UKB) via the UKB Research Analysis Platform (RAP), using the publicly available dxpy Python package (https://github.com/dnanexus/dx-toolkit). For other cohorts data were obtained as described in the Methods section.
Data analysis	Code for reproducing the genetic association analyses is publicly available at https://github.com/rgcgithub/regenie. REGENIE v4.1 was used for the analyses in this study. Full details of the data analysis pipeline and additional open-source tools and resources used for each step are provided in the Methods section. Base R packages were used for all other statistical calculations and linear model fits, and ggplot2 was used to generate plots.

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

UKB individual-level genotypic and phenotypic data may be accessed by approved investigators via a data access agreement with the UK Biobank (www.ukbiobank.ac.uk/). Additional information about registration for access to the data are available at www.ukbiobank.ac.uk/register-apply/. Regeneron can make individual-level genomic data from the GHS-RGC DiscovEHR collaboration available to qualified academic noncommercial researchers through the REGN pre-clinical Research portal at https://regeneron.envisionpharma.com/vt_regeneron/ under a data access agreement. MDCS data may be available to qualified academic non-commercial researchers to reproduce results reported in this manuscript through the portal at https://www.malmo-kohorter.lu.se/malmo-cohorts, following the principles outlined in this policy https://www.malmo-kohorter.lu.se/sites/malmo-kohorter.lu.se/files/mdcs_mpp_mos_request_form_vermar20.doc. For other cohorts/biobanks included in this manuscript, academic, non-commercial researchers interested in reproducing the results reported in this manuscript may request access to individual-level data via a data access agreement by reaching out to the associated biobank.

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	In this study we include sex and sex interaction terms as covariates in the models used to test for genetic association, such that the inferred coefficients represent the genetic effect common to men and women after removing differences due to sex. This is carried out in order to boost power to detect associations that are common to men and women. The data are normalized prior to analysis using a sex-stratified rank inverse normal transform (RINT), such that values for male and female study participants can be analyzed together in a joint analysis. In order to check for sex-specificity of the inferred effects, we also carried out a sex-stratified genetic association analysis for the key traits analyzed, and observed that the male- and female-specific effect sizes are not significantly different. These sex-stratified results are included in the manuscript.
Reporting on race, ethnicity, or other socially relevant groupings	Genetically estimated ancestry was included in the set of covariates for the genetic analysis, via genetic principal components as described in the Methods section.
Population characteristics	Variation in age at time of data acquisition was accounted for via covariates in the genetic analyses, as described in the Methods section.
Recruitment	Our study included participants from both population-based and health-system based cohorts as described in the Methods section.
Ethics oversight	All studies received approval from the relevant ethics committees, and participants gave their informed consent to take part in the research. More details are provided in the references for respective cohorts in the Methods section.

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	Sample size was determined based on the available set of participants with genetic and phenotypic data in each cohort after applying quality control filters as described in the Methods section.
Data exclusions	For imaging-derived phenotypes, individuals with low-quality images were excluded from the analyses using pre-established protocols as described in the Methods section. Standard quality filters were also applied to other quantitative traits, and pre-established exclusion criteria applied to binary traits as described.
Replication	Reproducibility of results in individual cohorts was assessed via meta-analysis across cohorts where possible.
Randomization	Covariates were controlled for in the genetic association analyses as described in the Methods section.

Blinding	Phenotypic data collection was carried out prior to genetic sequencing, hence all persons involved in sample processing and data collection were blinded to genetic status of the individuals involved.

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

Plants

Seed stocks	*Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.*
Novel plant genotypes	*Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.*
Authentication	*Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.*