nature portfolio

Corresponding author(s): Chloé Sarnowski

Last updated by author(s): Jul 18, 2025

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	Phenotype and covariate data were downloaded from dbGaP. Contributing studies have custom pipelines for data generation.
Data analysis	We used publicly-available software for all analyses in this manuscript. PRS-CSx (https://github.com/getian107/PRScsx) R version 4.4.2 GENESIS (https://bioconductor.org/packages/release/bioc/html/GENESIS.html) MetaSubtract (https://cran.r-project.org/web/packages/MetaSubtract/index.html)

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

Data for each participating study can be accessed through the database of Genotypes and Phenotypes (dbGaP) with the corresponding accession numbers (ARIC, phs001211; CARDIA, phs001612; CHS, phs001368; FHS, phs000974; GeneSTAR, phs001218; GENOA, phs001345; GOLDN, phs001359; HyperGEN, phs001293; HCHS/SOL, phs001395; JHS, phs000964; MESA, phs001416; SAFHS, phs001215; WHI, phs001237). GWAS summary data on glycemic traits have been contributed by MAGIC investigators and downloaded from www.magicinvestigators.org. Study-specific GWAS summary statistics for glycemic traits were provided by MAGIC consortium.

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	Self-reported sex was included as a covariate in all analyses performed. Both sexes were included in the analyses.
Reporting on race, ethnicity, or other socially relevant groupings	Race and ethnicity were self-reported by participants in the study. Participants included in the analyses represented four major races or ethnicities. Pooled analyses were conducted by combining data across studies and races and ethnicities. Race and ethnicity were not used in the primary pooled analyses that were adjusted for genetic principal components. In secondary analyses, participants were stratified by race and ethnicity, if sample sizes were large enough. These variables were used for stratification because genetic variants allele frequency can vary between populations.
Population characteristics	Descriptives tables are presented in the Supplementary Materials. We included ~17K participants from 5 studies representing 4 races or ethnicities in the validation set used to derive the cross-population polygenic score (mean age ranging from 44yrs to 67; %men ranging from 38% to 48%), and 21K participants from 8 studies representing 4 races or ethnicities in the testing set used for association analyses with the cross-population polygenic score (mean age ranging from 45yrs to 73; %men ranging from 31% to 47%). In association analyses with neurological outcomes, the number of participants ranged from 7.7K to 19.2K (mean age ranging from 61yrs to 75; %men ranging from 42% to 53%) from 9 studies representing 4 races or ethnicities.
Recruitment	Recruitment of participants varies per study and is described in the Supplementary Material with appropriate references to study-specific publication(s) for each study.
Ethics oversight	All study participants provided informed consent, and each study was approved by their respective 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	17K participants in the validation set, 21K participants in the testing set for polygenic scores analyses with HOMA-IR 9.4K, 19.2K, and 7.7K participants for Dementia/AD, general cognitive function, and volumes derived from MRI analyses respectively
Data exclusions	Exclusions have been described in the Methods section. Duplicates within and between studies were excluded. Participants with type 2 diabetes at the time of glycemic traits measurement (for HOMA-IR analyses) and at the time of neurological traits measurement (for Dementia, AD, general cognitive function, and volumes derived from MRI analyses) were excluded.
Replication	We divided our datasets into a validation and a testing set as described in the Methods section to assess the robustness of our polygenic scores. We performed association analyses in all TOPMed cohort studies with harmonized neurological outcomes available.
Randomization	NA - non interventional

Blinding | NA - non interventional

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: NA

Novel plant genotypes: NA

Authentication: NA

Magnetic resonance imaging

Experimental design

Design type: Association analyses of genetic scores related to insulin resistance with brain structural volumes (intracranial, total brain, hippocampal, and ventricular)

Design specifications: Used all TOPMed studies with brain MRI data available.

Behavioral performance measures: NA / not used.

Acquisition

Imaging type(s): Structural brain MRI.

Field strength: 1 T to 3 T

Sequence & imaging parameters: Varied, all had T1-weighted and T2 FLAIR sequences.

Area of acquisition: Whole brain including brainstem

Diffusion MRI ☐ Used ☒ Not used

Preprocessing

Preprocessing software: Varied by study. Freesurfer, HAMMER, multi-atlas based segmentation, and in-house pipelines were used. Described in Supplementary Material.

Normalization: Normalized via specific approaches described above.

Normalization template: Varied across cohorts based on programs above.

Noise and artifact removal	Scans where volumes could not be assessed due to artifacts were excluded.
Volume censoring	Varied across cohorts based on programs above.

Statistical modeling & inference

Model type and settings	Linear mixed-effects regression models with volume and covariates followed by score tests.
Effect(s) tested	Associations of genetic scores related to insulin resistance with various brain volumes

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

Anatomical location(s): Atlases as specified above

Statistic type for inference (See Eklund et al. 2016)	Simple association of genetic scores related to insulin resistance with brain structural volumes (intracranial, total brain, hippocampal, and lateral ventricular)
Correction	Bonferroni correction for multiple-testing

Models & analysis

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