nature portfolio

Corresponding author(s): Dr. Zdenka Pausova

Last updated by author(s): 09/15/2024

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	Participating study-specific data collection information is provided in the Supplementary Note 1 and Supplementary Table 1.
Data analysis	Data analysis and plotting was performed in R v4.x, unless otherwise stated. No previously unreported custom code was used to generate the findings, and methods have been described in previous publications. Cohort specific GWAS results were quality controlled using Easy QC v9.0, and results were meta-analyzed using METAL. Genetic correlations between traits was performed using LDSC v 1.0.1. GWAS enrichment of celltype specific signatures was performed using scDRS (v1.0.2), and summary statistics were processed using MAGMA (v1.09mac). Polygenic risk scores were calculated using using PRS-CS (v1.1.0) as implemented in the FinnGen PRS-CS pipeline (described at https://github.com/FINNGEN/CS-PRS-pipeline). Two sample MR was performed using the 'TwoSampleMR' R package. Cell type specific genes were derived from CELLEX package (v 1.2.2), used for cell-type enrichment with MRI-derived WMH-CT profiles.

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

The summary statistics from the meta-GWAS generated in this study have been deposited in the following database: https://figshare.com/articles/dataset/White_matter_hyperintensities_and_cortical_atrophy_genetic_risk_factors_and_underlying_neurobiology/27038320
Allen Human Brain Atlas gene expression, as parcellated in the Desikan-Killiany atlas is shared here: https://figshare.com/articles/dataset/Cell-specific_gene-expression_profiles_and_cortical_thickness_in_the_human_brain/4752955
Allen Human multiple cortical snRNAseq data was downloaded from here: https://portal.brain-map.org/atlases-and-data/rnaseq/human-multiple-cortical-areas-smart-seq
GWAS for plasma protein levels were downloaded from: https://www.synapse.org/Synapse:syn51364943/wiki/
Source Data are provided as a source data file.

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	Information on sex for each cohort is presented in Supplementary Table 1. Roughly, there were 37% to 53% of males across the individual studies. Sex was included as a covariate within the models.
Reporting on race, ethnicity, or other socially relevant groupings	The data comes from those of European ancestry, as recruited by individual cohorts. Genetic analysis within each cohort used the first 10 genetic principal components (to account for genetic relatedness).
Population characteristics	All the individuals, aged between 19 and 100 years, were stroke- and dementia-free and of European ancestry; brief description of cohorts and basic participant characteristics are provided in Supplement (Supplementary Table 1).
Recruitment	Data comes from population based cohort studies within the neuroCHARGE working group, and the UKBB.
Ethics oversight	Ethics oversight of this study was provided by the SickKids Research Ethics Board (#1000073323). Individual cohort protocols were approved by the research ethics boards at their respective institutions. Details can be found in the Supplementary Note 1.

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 formal power calculations were conducted. The sample size was dependent on the availability of both MRI (T1, and T2/FLAIR) and genetic data from the same individuals. There was a total of 51,065 participants from 10 cohort studies.
Data exclusions	Exclusion criteria included prior diagnosis of stroke or dementia, to only study those with 'normal' aging and WMH burden.
Replication	No attempts for replication, mainly due to limited availability of data.
Randomization	No randomization performed. Covariates were controlled by including them as variables within linear regression models.
Blinding	No blinding was performed.

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.