

Corresponding author(s): Maik Pietzner, Harry Hemingway, Claudia Langenberg

Last updated by author(s): Jan 23, 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 No software was used for data collection.

Data analysis R (v4.1.2); REGENIE (v2.2.4); LDSC (v1.0.1); LDstore (v1.0); R packages susieR (v0.11.92), coloc (v2.0)

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 individual level data is publicly available to bona fide researchers from the UK Biobank (<https://www.ukbiobank.ac.uk/>). This research has been conducted under the application 44448. GWAS summary statistics will be made available upon publication at <https://doi.org/10.7303/syn52716618>. Mapping of Read codes to phecodes can be downloaded from <https://github.com/spiros/ukbiobank-read-to-phecode>.

Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

Reporting on sex and gender	For survival models, we included 241,702 and 197,215 participants still alive by 01/01/2020 who self-identified as women and men, respectively, and had matching karyotypes. We systematically tested for sex-differential effects across all diseases investigated. For genetic analysis, we included a total of 239,758 and 201,913 participants who self-identified as women and men, respectively, and had matching karyotypes in our study for genetic analysis.
Population characteristics	UK Biobank is a national resource that has been described extensively elsewhere (https://www.ukbiobank.ac.uk). Individuals were not directly selected for inclusion in the study on the basis of any disease or health parameter. For survival analysis, people were on average 67.4 years old (s.d.=8.1), 55.1% were female, and 95.4% were of white European ancestry. Characteristics for white European participants used in the genetic analysis were similar.
Recruitment	All people aged 40–69 years (men and women) who were registered with the National Health Service and living up to 25 miles from one of the 22 study assessment centers were invited to participate in 2006–2010. Overall, about 9.2 million invitations were mailed to recruit 503,325 participants (a response rate of 5.47%).
Ethics oversight	All UK Biobank participants provided written informed consent, the study was approved by the National Research Ethics Service Committee North West–Haydock and all study procedures were performed in accordance with the World Medical Association Declaration of Helsinki ethical principles for medical research.

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](https://www.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	We used the full available sample in UK Biobank (with the exception of exclusions below) for discovery analyses.
Data exclusions	Individuals failing standard genotyping quality control parameters defined initially by the UK Biobank study or individuals of non-European ancestry (only genetic analysis) were excluded from analysis. These decisions were made prior to performing any downstream analysis. Survival analysis were further limited to participant still alive at 01/01/2020.
Replication	No replication was done, as comparable genetic analysis were not available. Findings, however, agreed well with the previous literature.
Randomization	N/A - randomization occurred naturally as genetic variants were the exposure.
Blinding	N/A - genetic association testing does not require blinding.

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

Methods

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