nature portfolio

Corresponding author(s): Seung Hoan Choi; Quanshun Mei

Last updated by author(s): Jun 25, 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	The authors did not perform primary data collection. Phenotypic and genomic data were obtained from the All of Us Research Program Controlled Tier Dataset v7 (Curated Data Repository release C2022Q4R9) and accessed by authorized users through the All of Us Researcher Workbench. Short-read whole-genome sequencing data were generated by the All of Us Genome Centers using standardized protocols. Full details are provided in the All of Us Research Program Genomic Research Data Quality Report: https://support.researchallofus.org/hc/en-us/article_attachments/19370367115796
Data analysis	VEP (v109); dbNSFP (v4.4); GENESIS (v2.18) adapted version (https://github.com/seanjosephjurgens/UKBB_200KWES_CVD) in R; Custom transcript-aware association test code (https://github.com/seuchoi/meta_transcript).

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

This study used data from the All of Us Research Program's Controlled Tier Dataset v7 (C2022Q4R9), available to authorized users on the Researcher Workbench (https://www.researchallofus.org/). Access is restricted due to participant privacy protections and program access policies.
Short-read whole-genome sequencing data were generated by the All of Us Genome Centers using standardized protocols. Full details are provided in the All of Us Research Program Genomic Research Data Quality Report: https://support.researchallofus.org/hc/en-us/article_attachments/19370367115796
Source data used for plotting and summary-level figure generation are provided in the 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	This study adjusted for genetic sex as a covariate in the association tests. Genetic sex was not used as an inclusion or exclusion criterion. Sex- or gender-specific effects were not the primary focus of this study, and sex- or gender-stratified analyses were not performed. Therefore, findings apply to the general population.
Reporting on race, ethnicity, or other socially relevant groupings	The main analyses in this study did not stratify by race, ethnicity or other socially relevant groupings in order to enhance generalizability. Population stratification was controlled by adjusting for genetic principal components (pre-calculated by All of Us) in the association tests. Genetically predicted ancestries (also pre-calculated by All of Us) were used in sensitivity analyses to assess whether estimated effect sizes were broadly consistent across genetic ancestry groups. These analyses were interpreted as genetic-ancestry-based sensitivity analyses and not as analyses of race, ethnicity, or other socially defined groupings.
Population characteristics	The analyses were conducted for 129 cardiopulmonary traits, with sample sizes and participant characteristics varying by trait. Trait-specific sample sizes, case/control counts, covariate distributions, and summary statistics are provided in the supplementary files.
Recruitment	Participants were recruited and enrolled by the All of Us Research Program. Participants could enroll through JoinAllofUs.org and through All of Us recruitment sites. Additional recruitment details are available from the All of Us Research Program. Details are available at https://www.researchallofus.org/faq/how-are-participants-recruited-and-what-does-participation-entail/. Because All of Us is a volunteer-based cohort, potential self-selection and differential participation may affect representativeness and should be considered when interpreting the results.
Ethics oversight	This study used de-identified data from the NIH All of Us Research Program accessed through the All of Us Researcher Workbench. The All of Us Research Program has received approval from the All of Us Research Program Institutional Review Board, and all participants provided informed consent. This study did not involve new participant recruitment or direct participant contact.

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 sample-size calculation was performed. Sample size was determined by the availability of participants with the required phenotype, covariate, and genomic data in the All of Us Controlled Tier Dataset v7. Because the analyses were performed across 129 cardiopulmonary traits, sample size varied by outcome. Trait-specific sample sizes are reported in the manuscript and supplementary tables.
Data exclusions	Participants were included based on the availability of required demographic, phenotype, covariate, and genomic data. Exclusion criteria were defined before association testing and were based on data availability, phenotype definitions, and genomic quality-control procedures. For binary phenotypes, case status was defined using corresponding ICD-9 and ICD-10 codes from electronic health records, and phenotypes with fewer than 200 cases were excluded to improve statistical robustness. Highly correlated or redundant binary phenotypes were identified by hierarchical clustering, and one representative phenotype with better clinical interpretability was selected from each cluster. For continuous traits, invalid values, including physiologically implausible values or measurements with missing units, were excluded during data curation, and

	the most recent non-missing measurement was used for each individual. For sample-level genomic quality control, flagged population outliers, potential duplicates, samples with call rate below 90%, and individuals lacking genetically determined sex information were excluded. For genotype- and variant-level quality control, only genotypes passing central quality-control procedures with Genotype Quality score above 20 were retained, and variants with missing rate greater than 10% were excluded. These exclusions resulted in a final analysis dataset of 242,881 participants. Full descriptions are provided in the manuscript and supplementary files.
Replication	External replication analyses were conducted using publicly available UK Biobank results from Genebass. Two of the eight evaluated findings were successfully replicated. Findings that did not replicate may reflect differences in phenotype definitions, ancestry composition, sample size, variant/transcript availability, allele frequency, or statistical power between All of Us and UK Biobank. Additional details are provided in the manuscript and supplementary files.
Randomization	Random allocation was not applicable because participants were not assigned to experimental groups. This study used de-identified observational data from the All of Us Research Program.
Blinding	Blinding was not relevant because this was a secondary analysis of de-identified observational data and did not involve experimental interventions, treatment assignment, group allocation, or subjective outcome assessment by investigators.

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.*