nature portfolio

Corresponding author(s): Fuu-Jen Tsai and Chin-Chi Kuo

Last updated by author(s): Sep 27, 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	We used SAS 9.4 software for data cleaning and management.
Data analysis	We used publicly available software for the data analysis (plink 1.9 and 2.0, liftOver, SHAPEIT2, IMPUTE2, Regenie v3.1, METAL, LDSC v1.0.1, R v4.2.3, FUMA v1.5.2., MAGMA v1.08, GCTA v1.94.1, R package "gtx" v2.1.6 [https://github.com/tobyjohnson/gtx],R package "coloc" v5, PRSice2)

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 GWAS summary data supporting our findings are available on Figshare (doi:10.6084/m9.figshare.24356587), and the statistical data are available in the

Supplementary Data. The data from Taiwan Biobank (TWB; Application No. TWBR11111-02) and UK Biobank (UKB; Application No. 81803) were obtained through approved applications and they are publicly available to approved researchers for health-related research (TWB: https://www.biobank.org.tw/english.php; UKB: https://www.ukbiobank.ac.uk/enable-your-research/apply-for-access). The GWAS summary data from Biobank Japan (BBJ) are publicly available without permission at http://jenger.riken.jp/en/result. The individual-level raw data from CMUH are unavailable because they contain information that may compromise participant privacy.

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	The proportion of male participants varied across the different biobanks used in this study. Biobank Japan (BBJ) had the highest proportion of males (54.9%), followed by Taiwan Biobank (TWB) replication dataset (49.9%), UK Biobank (UKB) (45.7%), China Medical University Hospital-Clinical Research Data Repository (CMUH-CRDR) (44.43%) and TWB discovery dataset (32.4%). We performed our analyses using data from all genders, and the gender distribution of each biobank is provided in Supplementary Table 1.
Reporting on race, ethnicity, or other socially relevant groupings	We used the data from BBJ, TWB, and CMUH-CRDR, which mostly come from Asian population. UKB primarily includes European population.
Population characteristics	BBJ is a prospective biobank that recruited approximately 200,000 participants, mainly of Japanese ancestry. Mean age of participants in this study at recruitment was 62.9 years old, and 54.9% were male. TWB is a prospective biobank that collected samples from about 160,000 Taiwanese individuals. The discovery dataset of this study had a mean age of 50.4 years and 32.4% male participants. The replication dataset of this study had a mean age of 49.26 years and 49.9% male participants. CMUH-CRDR is a retrospective cohort from a hospital in Taiwan, with an average age of 56.72 years and a male proportion of 44.43%. UKB is a prospective cohort that enrolled about 500,000 people from the UK, with an average age of 54.98 years and a male proportion of 45.7%.
Recruitment	BBJ is a hospital-based registry of collected DNA samples, serum samples, and clinical information from approximately 200,000 patients at 66 hospitals affiliated with 12 medical institutions between 2003 and 2007. TWB is an ongoing project that contains data and samples from a national prospective cohort of the Taiwanese population; its aim is to longitudinally collect a wide range of phenotypic measurements and genomic data from Han Chinese people aged 20–70 years without cancer history. The CMUH-CRDR is a database that contains clinical information of over 3 million patients from 2003 to 2020. It covers various aspects of health care, such as demographics, diagnoses, surgeries, lab tests, and death records. The data source is the Health and Welfare Data Science Center of the Ministry of Health and Welfare. The UKB project recruited approximately 500,000 people aged between 40–69 years from 2006 to 2010 from across the United Kingdom.
Ethics oversight	This study has been approved by the Research Ethical Committee/Institutional Review Board (REC/IRB) of CMUH (IRB No. CMUH105-REC3-068, CMUH110-REC2-145, CMUH111-REC3-138, CMUH112-REC2-036), Taiwan Biobank (Application No. TWBR11111-02), and UK Biobank (Application No. 81803).

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	For GWAS, we included a total of 244,952 as discovery cohort and 27,058 as replication cohort. For validating the polygenic risk score of CKD, we included 25,345 patients from CMUH and 260,245 from UKB.
Data exclusions	This study excluded individuals who were younger than 18 years old and had no serum creatinine measurement.
Replication	We conducted a replication data set (n = 27,058) from Taiwan (Taiwan Biobank). In addition, the polygenic risk score was validated in both CMUH data and UKB data.
Randomization	Not applicable.
Blinding	Not applicable.

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