nature research

Corresponding author(s): Jianfeng Feng
Wei Cheng

Last updated by author(s): Feb 12, 2022

Reporting Summary

Nature Research 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 Research 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 involved in data collection (data used is all directly available from UK Biobank, as described in detail in the paper)
Data analysis	Matlab 2018b was used to perform nonlinear association analysis. Freesurfer v6.0 was used to process the imaging data. PLINK 1.90 and PRSice version 1.25 (www.PRSice.info) were used to perform genome-wide association analysis and calculate the polygenic risk score respectively. R version 3.6.0 package: lavaan 0.8 was used to perform longitudinal analysis, mediation analysis and structural equation model, AER 1.2-9 was used to perform interaction test, rms 6.2-0 was utilized to conduct restricted cubic spine analysis, GenomicSEM version 0.0.3 was utilized to calculate the heritability and genetic correlation, two-lines test version 0.52 was utilized to identify the breakpoints of the nonlinear model. Scripts used to perform the analyses are available at https://github.com/yuzhulineu/UKB_sleep

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 Research 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 list of figures that have associated raw data
- A description of any restrictions on data availability

This project corresponds to UK Biobank application ID 19542. Neuroimaging, genotype and behavioral data from UK Biobank dataset are available from https://biobank.ndph.ox.ac.uk/ by application. The variables utilized here are detailed in Supplemental Table 1. Previous published GWAS of sleep duration was downloaded via http://www.t2diabetesgenes.org/data/. European ancestral background LD scores from the 1000 Genomes Project were downloaded from https://alkesgroup.broadinstitute.org/LDSCORE/.

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 statistical methods were used to predetermine sample sizes and all currently available sample in the UK Biobank were used including 498,277 subjects with behavioral data, 156,884 subjects with follow-up mental health data and 39,692 subjects with imaging and genotype data. The sample sizes were provided in the method section.
Data exclusions	For calculation of polygenetic risk score, we excluded single-nucleotide polymorphisms (SNPs) with call rates < 95%, minor allele frequency < 0.1%, deviation from the Hardy–Weinberg equilibrium with p < 1E-10 and selected subjects that were estimated to have recent British ancestry and have no more than ten putative third-degree relatives in the kinship table, consistent with the previous study.
Replication	All available data were used to maximize statistical power of the analysis therefore we did not repeat the analysis.
Randomization	In our nonlinear association analysis, we regressed out covariates including age, sex, body mass index, scanning site of imaging, Townsend deprivation index measuring socioeconomic status, qualifications, smoking status and drinking status.
Blinding	Blinding was not applicable to this study as this study is observational.

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
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

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

Human research participants

Policy information about studies involving human research participants

Population characteristics	A total of 498,277 participants aged between 38 and 73 (54% females) was obtained from the UK biobank study, which is a large-scale database containing cognitive assessment, mental health questionnaires, brain imaging and in-depth genetic information from UK participants. Neuroimaging data of 39,694 participants were available and utilized in the current analyses. See Method section for further details. Population characteristics of the participants are listed in Table 1.

Recruitment	The UK Biobank recruited over 500,000 people aged 40-69 since 2006 at 22 recruitment centers across the UK. Previous investigation showed UK biobank subject to a healthy sample bias.
Ethics oversight	UK Biobank has approval from the North West Multi-centre Research Ethics Committee (https://www.ukbiobank.ac.uk/learn-more-about-uk-biobank/about-us/ethics) as a Research Tissue Bank approval and provides oversight for this study. Written informed consent was obtained from all participants.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Magnetic resonance imaging

Experimental design

Design type	Structural MRI
Design specifications	UK Biobank designed the imaging acquisition protocals including 6 modalities, covering structural, diffusion and functional imaging. The collection order is T1-weighted structural image, resting-state functional MRI, task functional MRI, T2-weighted FLAIR structural image, Diffution MRI and susceptibility-weighted imaging. T1-weighted structural image was acquired using straight sagittal orientation for 5 minutes.
Behavioral performance measures	Cognitive tests were first administered via touchscreen interface in the UK biobank assessment center at the baseline visit and repeated at the neuroimaging visit. Six cognitive tests including reaction time, numeric memory, fluid intelligence, trail making, prospective memory and pair matching test were utilized in the current study. Measurement of depressive symptoms via 4-item Patient Health Questionnaire-4 (PHQ-4) was first assessed in the UK biobank assessment center (2006-2010, n=499,585) and then repeated at the neuroimaging visit (2014-2017, n=48,571).

Acquisition

Imaging type(s)	T1-weighted structural imaging
Field strength	3T
Sequence & imaging parameters	The EPI-based acquisitions utilize simultaneous multi-slice (multiband) acceleration. Biobank uses pulse sequences and reconstruction code from the Center for Magnetic Resonance Research (CMRR), University of Minnesota https://www.cmrr.umn.edu/multiband. The resolution is 1x1x1 mm and field of view is 208x256x256 matrix. Straight sagittal orientation is used. TR and TE are 2000ms and 2.01ms respectively. The flip angle is 8 deg. Detailed sequence and imaging parameters are openly available here: https://biobank.ndph.ox.ac.uk/showcase/showcase/docs/ brain_mri.pdf
Area of acquisition	Whole brain

Diffusion MRI ☐ Used ☒ Not used

Preprocessing

Preprocessing software	Imaging derived phenotypes (IDP) generated by an imaging-processing pipeline developed and run on behalf of UK Biobank were used in the study. T1 images were processed with Freesurfer, surface templates were utilized to extract imaging derived phenotypes (IDP) referring to atlas regions' surface area, volume and mean cortical thickness42. Subcortical regions were extracted via FreeSurfer's aseg tool. The full processing pipeline is openly available here: http://doi.org/10.1016/j.neuroimage.2017.10.034
Normalization	see above
Normalization template	fsaverage
Noise and artifact removal	see above
Volume censoring	see above

Statistical modeling & inference

Model type and settings	mass univariate, nonlinear regression model, cross-lagged panel model, structural equation modeling
Effect(s) tested	For nonlinear regression model, F-tests were utilized to access statistical significance and derive F-statistics and corresponding one-sided p values adjusted for multiple comparisons. For cross-lagged panel model and structural equation modeling, Wald tests were utilized to derive the two-sided p value.

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

Statistic type for inference (See Eklund et al. 2016)	voxel-wise association, voxel-wise Bonferroni correction
Correction	Bonferroni

Models & analysis

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