nature portfolio

Corresponding author(s): Jie Zhang and Lena Palaniyappan

Last updated by author(s): Oct 4, 2025

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	This study uses a publicly available dataset, and the method of data collection and the software used in the process can be found on the dataset official website (https://abcdstudy.org/) and related github pages (https://github.com/ABCD-STUDY).
Data analysis	The main codes used for data analysis in this study were all based on R (version 4.4.2), plink 2.0.1 was used for genetic association analysis and FUMA (https://fuma.ctglab.nl) was utilized to perform genetic enrichment analysis. This study used all code can be found in the author's making home page (https://github.com/XinRanWu/ABCD_YoungAsym).

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

Neuroimaging and behavioural data from ABCD dataset is obtained from https://nda.nih.gov/abcd with the approval of the ABCD consortium.

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	We take sex into considerations in our study and sex was utilized as a covariate in all the analysis.
Reporting on race, ethnicity, or other socially relevant groupings	In the ABCD (Adolescent Brain Cognitive Development) study, racial and ethnic information is primarily collected through self-reporting by participants and/or their parents. Typically, participants or their guardians complete a demographic questionnaire that includes questions about race and ethnicity. This data is collected following guidelines similar to those used in population studies like the U.S. Census, providing standardized categories for race and ethnicity. Race and ethnicity information was included as covariates in this study to prevent them from interfering with the main statistical results.
Population characteristics	A total of 11,878 participants aged between 10 and 14 (52% males) was obtained from the ABCD study, which is a large-scale database containing cognitive assessment, mental health questionnaires, brain imaging and in-depth genetic information from US participants. This study used age, sex, race and ethnicity as population characteristics.
Recruitment	Participants were registered according to the sex, ethnicity, and race distribution of the U.S. population. For more information, please refer to https://abcdstudy.org/.
Ethics oversight	Data collection, processing, and publication for the ABCD study were approved by the Institutional Review Boards (IRBs) at the individual study sites and coordinating institutions.

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	11,878 subjects in ABCD dataset. All currently available sample size were used in the study.
Data exclusions	Of the 11,878 participants, 11,771 were successfully pretreated with Freesurfer at baseline, 8,123 at 2-year longitudinal follow-up, and 3,045 at 4-year longitudinal follow-up. After excluding Freesurfer quality control failure subjects based on ABCD's official recommended quality control documents, there were 11,368 subjects at baseline and 7,695 subjects at 2-year longitudinal follow-up. In 4 years, due to the lack of Freesurfer quality control files, quality control was carried out only based on T1 images quality, leaving 3,034 available subjects. Of the data available at baseline, a total of 4,236 individuals with European ancestry were used for the GWAS analysis. In calculating the longitudinal rate of change from baseline to 2-year follow-up, the intersection of data available at baseline and 2-year follow-up was taken, with a total of 7,411 image-available data and 3,029 European ancestry data available at GWAS.
Replication	N/A. All available data are used to maximize statistical power of the analysis.
Randomization	N/A. This is an observational study.
Blinding	N/A. This is an observational study.

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.*

Magnetic resonance imaging

Experimental design

Design type	Structural MRI and diffusion MRI.
Design specifications	N/A (recruitment not part of this study).
Behavioral performance measures	The study used psychiatric questionnaires and cognitive tests to quantify behavioral performance, details of which are listed in the methods section of the text and in supplemental materials.

Acquisition

Imaging type(s)	Structural MRI and diffusion MRI.
Field strength	3 Tesla.
Sequence & imaging parameters	High-resolution T1- and T2-weighted structural Magnetic Resonance Images (1mm isotropic, prospective motion correction) were collected.
Area of acquisition	Whole brain.

Diffusion MRI ☒ Used ☐ Not used

Parameters: High angular resolution diffusion imaging (1.7mm isotropic) was collected using multiband acquisition (b-values=4, factors=4, directions=96).

Preprocessing

Preprocessing software	Freesurfer, AFNI, FSL, AutoTrack
Normalization	The data in the ABCD study undergo both linear and non-linear transformations. T1-weighted (T1w) and T2-weighted (T2w) images are registered to standard templates such as the ICBM152 atlas. This involves pre-registration to individual subject images and fine-tuning registration with tools like FSL's brain extraction tool to remove noise and improve alignment. Additionally, for fMRI, slice timing correction and field distortion corrections are applied, but no non-linear normalization is typically applied to raw BOLD signals unless noted in specific analyses.
Normalization template	The preprocessing pipeline registers anatomical images to group-standardized spaces like the ICBM152 and fsLR-32k for cortical surfaces. These templates ensure that data from different subjects are aligned to a common anatomical framework, which facilitates group-level analyses.
Noise and artifact removal	Trained technicians review cortical surface reconstructions, assessing five artifact categories: motion, intensity inhomogeneity, white matter underestimation, pial overestimation, and magnetic susceptibility artifacts. Each category is rated 0-3 (absent to severe). Reconstructions are recommended for use (1) or exclusion (0), with exclusion if any severe

	artifact is present. Automated quality control measures include metrics like signal-to-noise ratio (SNR) and brain intensity statistics.
Volume censoring	No fMRI information was applied in this study, so volume censoring was not performed.

Statistical modeling & inference

Model type and settings	In this study, a linear mixed model with random intercepts (sites) was used for the primary statistical analysis, and a structural equation model was used to estimate the relationship between the multiple variables.
Effect(s) tested	T-value of the one-sample T-test corresponding to the beta value in the linear mixed model was used as a measure of the effect size.

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

Anatomical location(s)	Labels for cortical grey matter were assigned based on the Desikan-Killiany parcellation, which divides the cortex into 68 regions of interest (ROIs). Subcortical structures are segmented using the automatic subcortical segmentation (ASEG) procedure of Freesurfer 5 and contained 14 ROIs. Carpus callosum are labeled based on white matter tracks atlas of AtlasTrack.
Statistic type for inference (See Eklund et al. 2016)	This study only used the mean within ROI for analysis, and did not involve voxel-wise and cluster-wise inference.
Correction	FDR correction was used in this study.

Models & analysis

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