nature research

Corresponding author(s): Thomas Werge

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 Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. 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
☐	☒	An indication of 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 statistics 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
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	R v3.3.1 was used to organize subject records into appropriate input files.
Data analysis	R v3.3.1, python v2.7, Plink v1.90 beta3, GCTA v1.25.2, LDSC v1.0.0, FineMap v1.1, GOElite, Eigensoft 6.0.1, KING v1.9, SHAPEIT3, Impute2

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/reviewers upon request. 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

In accordance with the consent structure of iPSYCH and Danish law, individual level genotype and phenotype data are not able to be shared publicly. Cross-disorder (XDX) GWAS summary statistics are available for download (https://ipsych.au.dk/downloads/). Summary statistics from secondary GWAS of single disorders are

available upon request from the corresponding author.

Field-specific reporting

Please select 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/authors/policies/ReportingSummary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	A genome-wide associate study of lifetime psychiatric diagnoses as of 2012 in a 1981-2005 birth cohort .
Research sample	We used data from the iPSYCH case-cohort study: A random, representative sample of the Danish neonatal biobank for individuals born between 1981 and 2005, with a known mother, alive on their first birthday, who have not opted out of the Danish national register based research program and the complete population of individuals from the Danish population birth cohort meeting the same inclusion criteria and with a major psychiatric diagnosis in the national registers as of 2012. The case sample was chosen to be complete and the random population sample to be representative and of comparable size to the largest individual case cohort. Further details can be found in PMID: 28924187.
Sampling strategy	Data was ascertained according to a case-cohort design where the size of the cohort (random sample) was chosen to represent 2% of the broader population it was sampled from (population of Denmark born between 1981 and 2005). Cases were all cases in the population at the time of ascertainment. These samples are state-of-the-field for single cohort GWAS studies, representing the largest of its kind, and consistent with previous reports adequately powered for discovery.
Data collection	We used data from the iPSYCH case-cohort study: DNA was extracted from dried neonatal blood spots and amplified before genotyping on the Infinium PsychChip v1.0. Blood was collected between 4-7 days after birth and stored at -20 C until the time of ascertainment. Psychiatric diagnoses were aggregated from national registers. Demographic and social variables were aggregated from national civil registers. Further details can be found in PMID: 28924187.
Timing	All data was initially collected in 2012 and psychiatric diagnoses were later updated, complete through 2014.
Data exclusions	Among the 78,050 samples with genotype data available in the iPSYCH case-cohort study, 5,353 were excluded according to genotype and imputation quality control procedures described in great detail in Supplementary Note 2. These data exclusion criteria were determined before the study was designed or conducted.
Non-participation	All subjects in the iPSYCH case-cohort data resource were initially included. This sample was drawn only from those meeting consent requirements.
Randomization	When control samples were split into random sub-cohorts for analysis the sample() function in R was used to select subsamples of subject unique IDs.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants

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	See above
Recruitment	Patients were recruited according to consent by non-opt out from a national research program in which very few individuals opt out. Our control sample was drawn nearly randomly from the national biobank making the potential for ascertainment effect negligible. Our case cohort is the complete population as of 2012, which is young for psychiatric disorders meaning there may be some enrichment for early onset and some yet-to-convert cases in the controls for anyone disorder. These are unlikely to affect most of our results.