nature portfolio

Corresponding author(s): Yong Chen, PhD

Last updated by author(s): Mar 10, 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 data is collected from electronic health data from RECOVER Initiative.
Data analysis	All analyses were performed using R version 4.4.0. The code used for the analysis in this study is available and can be accessed in the code ocean repository at https://doi.org/10.24433/CO.4519602.v1.

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 results reported in this study are based on detailed individual-level patient data compiled as part of the RECOVER Initiative. Due to the high risk of reidentification based on the number of unique patterns in the data, patient privacy regulations prohibit us from releasing the data publicly. The data are maintained in a secure enclave, with access managed by the program coordinating center to remain compliant with regulatory and program requirements. Please

direct requests to access the data, either for reproduction of the work reported here or for other purposes, to the RECOVER EHR Pediatric Coordinating Center (recover@chop.edu).

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	A total of 110,955 pediatric population aged 5-20 years with a documented SARS-CoV-2 infection and preexisting neuropsychiatric conditions met the inclusion criteria within the RECOVER Database were included. Of these, 12,416 (11.2%) were prescribed an SSRI/SNRI before infection, and 98,539 (88.8%) were not. The SSRI/SNRI group had a higher proportion of female patients (68.7% vs. 50.6%) compared to the non-SSRI/SNRI group. The sex variable used in our analysis is based on the person domain in PEDSnet CDM table. A gender_concept_id of 8532 indicates "Female", and a gender_concept_id of 8507 indicates "Male". All other values are categorized as "Other/Unknown", and we only included Male and Female in our analysis. More details can be found here: https://github.com/PEDSnet/Data_Models_Public/blob/master/PEDSnet/docs/Conventions%20Docs/v5.5_PEDSnet_CDM_ETL_Conventions.md.
Reporting on race, ethnicity, or other socially relevant groupings	Across the cohorts, non-Hispanic White patients comprised a greater share of the SSRI/SNRI group (68.4% vs. 47.3%) compared to the non-SSRI/SNRI group, while non-Hispanic Black patients (8.6% vs. 18.1%) and Hispanic patients (13.2% vs. 22.8%) were underrepresented. The raw data for all ethnic/racial groups are from the "person" CDM table from the PEDSnet CDM in the EHR data. The ethnic and racial information is identified from columns "ethnicity_concept_id" and "race_concept_id". More details can be found here: https://github.com/PEDSnet/Data_Models_Public/blob/master/PEDSnet/docs/Conventions%20Docs/v5.5_PEDSnet_CDM_ETL_Conventions.md.
Population characteristics	The following patient characteristics were considered: demographic factors, including age at index date, sex (female, male), and race/ethnicity (Asian American/Pacific Islander, Hispanic, non-Hispanic Black, non-Hispanic White, Multiple, or Other/Unknown); clinical variables, including obesity status, a chronic condition indicator defined by the Pediatric Medical Complexity Algorithm (PMCA, no chronic condition, non-complex chronic condition, or complex chronic condition), and a list of pre-existing chronic conditions; health care utilization in the baseline period, including the number of inpatient, outpatient, emergency department visits, the number of unique medications, and the number of negative SARS-CoV-2 tests (0, 1, 2, or ≥3); vaccine status, including dosage of COVID-19 vaccine received prior to infection (0, 1, or ≥2) and the time since most recent vaccination (no vaccine, <4 months, or ≥ 4 months); calendar year-month of index infection; severity of acute COVID-19 (categorized as asymptomatic, mild, moderate, or severe) based on COVID-19-related hospitalization, intensive care unit (ICU) admission, respiratory support procedures, and diagnosis codes indicating respiratory acute COVID-19 complications within 14 days of the index date; and indicators for the 37 data-contributing sites.
Recruitment	The National Institutions of Health (NIH) launched the new RECOVER Initiative in 2021 to leverage electronic health record (EHR) data to better identify and characterize patients with post-acute sequelae of SARS-CoV-2 infection (PASC). RECOVER obtains EHRs from three large national healthcare networks within the United States, covering regional catchment areas across 41 states. These networks collectively hold the EHRs of over 60 million patients, including records from more than 7 million individuals who have been affected by COVID-19.
Ethics oversight	This study constitutes human subject research. Institute Review Board (IRB) approval was obtained under Biomedical Research Alliance of New York (BRANY) protocol #21-08-508. As part of the Biomedical Research Alliance of New York (BRANY IRB) process, the protocol has been reviewed in accordance with the institutional guidelines. The Biomedical Research Alliance of New York (BRANY) waived the need for consent and HIPAA authorization.

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

Behavioural & social sciences study design

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

Study description	Quantitative observational study
Research sample	*State the research sample (e.g. Harvard university undergraduates, villagers in rural India) and provide relevant demographic information (e.g. age, sex) and indicate whether the sample is representative. Provide a rationale for the study sample chosen. For studies involving existing datasets, please describe the dataset and source.*
Sampling strategy	This study used a retrospective cohort design based on existing electronic health record (EHR) data compiled through the NIH RECOVER Initiative. The sampling procedure was based on non-random, convenience sampling of all eligible pediatric patients (ages 5–20) from participating health systems between March 2020 and March 2024. Eligible participants were older than 5 years and

younger than 21 years, had a documented SARS-CoV-2 infection during the cohort entry period, and had a preexisting neuropsychiatric condition, and were required to have at least one healthcare encounter during a baseline period (24 months to 7 days before SARS-CoV-2 infection) and at least one encounter during the standard follow-up period (28-179 days after SARS-CoV-2 infection). Individuals younger than 5 or older than 21 years, had no preexisting neuropsychiatric condition, or had no follow-up visit during the follow-up period were excluded. Patients were classified as SSRI/SNRI users if they received SSRI/SNRI treatment (SSRIs included citalopram, escitalopram, fluoxetine, fluvoxamine, paroxetine, and sertraline; SNRIs included desvenlafaxine, duloxetine, and venlafaxine) during the period spanning 28 days before to 28 days after the index date, based on established pediatric prescribing practices and published research. All others were classified as non-SSRI/SNRI users.

Data collection

This study used de-identified electronic health record (EHR) data compiled through the NIH RECOVER Initiative. Data were extracted from clinical data warehouses of 37 health systems across the United States and aggregated in a secure research enclave.
Information was obtained from routine clinical encounters and included demographics, diagnostic codes (ICD-10), laboratory results, procedures, medication records, and encounter-level data.
No direct interaction occurred between researchers and participants. All data were passively collected as part of clinical care. No instruments or surveys were administered by the study team. Because the data originated from existing EHRs, no one was present with the participants beyond standard clinical personnel during medical visits.The research team did not have access to identifiable information, and the analyses were conducted on de-identified data within a secure computing environment. Blinding to experimental condition was not applicable, as this was an observational study using routinely collected data.

Timing

From March 1, 2020, to March 19, 2024

Data exclusions

Eligible participants were older than 5 years and younger than 21 years, had a documented SARS-CoV-2 infection during the cohort entry period, and had a preexisting neuropsychiatric condition, and were required to have at least one healthcare encounter during a baseline period (24 months to 7 days before SARS-CoV-2 infection) and at least one encounter during the standard follow-up period (28-179 days after SARS-CoV-2 infection). Individuals younger than 5 or older than 21 years, had no preexisting neuropsychiatric condition, or had no follow-up visit during the follow-up period were excluded.

Non-participation

No participants dropped out.

Randomization

Non-random. We employed the propensity score model approach to adjust for covariates stemming from various patient characteristics.

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.