nature portfolio

Corresponding author(s): Yuan Yuan

Last updated by author(s): Nov 10, 2021

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 directly download publicly available data from SafeGraph COVID-19 Data Consortium, CDC weisbite and The Texas Department of State Health Services. No codes involved at this stage.
Data analysis	We use python to perform all data analysis and machine learning algorithms. Gephi is used to generate Fig 1.

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 U.S. Mobile phone mobility data and census data are available through the SafeGraph COVID-19 Data Consortium https://www.safegraph.com/covid-19-data-consortium.
Vaccination data are available on CDC website https://covid.cdc.gov/covid-data-tracker/#datatracker-home. Texas vaccination data are available on https://github.com/shiruken/covid-texas-data/. Synthetic data and the intermediate data generated from the aforementioned public will be publicly available on GitHub: https://github.com/yuany94/covid-vaccine

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	This is a computational epidemiological study.
Research sample	We employ publicly available data on the mobility patterns in U.S. The network is constructed using the U.S. mobility from SafeGraph, a company that provides aggregated data collected from mobile applications. All data is anonymized and aggregated by the company so that individual information is not re-identifiable.
Sampling strategy	We do not perform sampling but replied on previously collected datasets. All datasets were chose as they were comprehensive. The representativeness of the data has been discussed at https://safegraph.com/blog/what-about-bias-in-the-safegraph-dataset .
Data collection	We use publicly available data. The U.S. Mobile phone mobility data and census data are available through the SafeGraph COVID-19 Data Consortium https://www.safegraph.com/covid-19-data-consortium. Vaccination data are available on CDC website https://covid.cdc.gov/covid-data-tracker/#datatracker-home. Texas vaccination data are available on https://github.com/shiruken/covid-texas-data/. Synthetic data and the intermediate data generated from the aforementioned public will be publicly available on GitHub: https://github.com/yuany94/covid-vaccine
Timing	Safegraph data: Jan-Dec 2019 Vaccination rate: July 1 Census: from 2016 to 2019
Data exclusions	Haiwaii is removed from our analysis as the vaccination rates are not publicly available, but as its population takes a very small proportion of the U.S. population, we believe it does not influence our main conclusions.
Non-participation	Because we relied on publicly available, anonymize, aggregated data from cell phone mobility tracking, we had no access to individual-level data and do not know how many participants drop-out or declined participation.
Randomization	This is not a randomized controlled trial so not applied.

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	See above.
Recruitment	See above.
Ethics oversight	We obtained IRB exemption from MIT IRB office (COUHES). All data are aggregated, anonymized, and publicly available.

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