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. t, F, 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 are from the UnitedHealth Group claims database. Medical and pharmacy claims data are captured, predominantly electronically, from sites of care seeking third-party reimbursement for both Medicare and commercial plans using the industry standard data collection forms HICFA/CMS-1500 for facility claims, UB04/CMS-1450 for professional services and outpatient claims, and NCPDP for pharmacy claims or their electronic equivalents. Structured data from these standardized forms are coded using the International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM), National Drug Codes (NDC), Current Procedural Terminology (CPT) codes, and Logical Observation Identifiers Names and Codes (LOINC) codes, and Diagnosis Related Groups (DRG). This nomenclature ensures consistency of data collection across geographic regions, health systems, and payers throughout the United States.

Data analysis

The R and C codes are available at https://github.com/Penncl/dCLR and through our R package (version 1.0-2): "pda" at https://cran.r-project.org/web/packages/pda/index.html.

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

Data for this study are unavailable due to patient privacy concerns. Data may be made available to the editors upon reasonable request.

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**: Describe how sample size was determined, detailing any statistical methods used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.

- **Data exclusions**: Describe any data exclusions. If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.

- **Replication**: Describe the measures taken to verify the reproducibility of the experimental findings. If all attempts at replication were successful, confirm this OR if there are any findings that were not replicated or cannot be reproduced, note this and describe why.

- **Randomization**: Describe how samples/organisms/participants were allocated into experimental groups. If allocation was not random, describe how covariates were controlled OR if this is not relevant to your study, explain why.

- **Blinding**: Describe whether the investigators were blinded to group allocation during data collection and/or analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.

Behavioural & social sciences study design

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

- **Study description**: Our objective is to develop an association model between clinical-and-demographic covariates (i.e., age, sex, line of business, and Charlson comorbidity index, which is the sum of the individual’s comorbidities weights developed by Charlson et al) and therapeutic patient outcomes.

- **Research sample**: The data are recorded from 230 sites with 14,215 insured (Commercial and Medicare) patients obtained from the UnitedHealth Group (UHG) Clinical Discovery Database. Our analytical dataset is composed of hospitalized patients who were diagnosed with COVID-19 prior to June 29, 2020 from a single large national health insurer, which covers a broad swath of the population.

- **Sampling strategy**: N/A

- **Data collection**: Medical and pharmacy claims data are captured, predominantly electronically, from sites of care seeking third-party reimbursement for both Medicare and commercial plans using the industry standard data collection forms HCFA/CMS-1500 for facility claims, UB04/ CMS-1450 for professional services and outpatient claims, and NCPDP for pharmacy claims or their electronic equivalents. Structured data from these standardized forms are coded using the International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM), National Drug Codes (NDC), Current Procedural Terminology (CPT) codes, and Logical Observation Identifiers Names and Codes (LOINC) codes, and Diagnosis Related Groups (DRG). This nomenclature ensures consistency of data collection across geographic regions, health systems, and payers throughout the United States.

- **Timing**: Our analytical dataset is composed of hospitalized patients who were diagnosed with COVID-19 prior to June 29, 2020 from a single large national health insurer, which covers a broad swath of the population.

- **Data exclusions**: Please see Figure 3 "Diagram of the patient inclusion-exclusion criteria" in the manuscript.

- **Non-participation**: N/A
Randomization: N/A

Ecological, evolutionary & environmental sciences study design

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

Study description: Briefly describe the study. For quantitative data include treatment factors and interactions, design structure (e.g. factorial, nested, hierarchical), nature and number of experimental units and replicates.

Research sample: Describe the research sample (e.g. a group of tagged Passer domesticus, all Stenocereus thurberi within Organ Pipe Cactus National Monument), and provide a rationale for the sample choice. When relevant, describe the organism taxa, source, sex, age range and any manipulations. State what population the sample is meant to represent when applicable. For studies involving existing datasets, describe the data and its source.

Sampling strategy: Note the sampling procedure. Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.

Data collection: Describe the data collection procedure, including who recorded the data and how.

Timing and spatial scale: Indicate the start and stop dates of data collection, noting the frequency and periodicity of sampling and providing a rationale for these choices. If there is a gap between collection periods, state the dates for each sample cohort. Specify the spatial scale from which the data are taken.

Data exclusions: If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.

Reproducibility: Describe the measures taken to verify the reproducibility of experimental findings. For each experiment, note whether any attempts to repeat the experiment failed OR state that all attempts to repeat the experiment were successful.

Randomization: Describe how samples/organisms/participants were allocated into groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.

Blinding: Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.

Field work, collection and transport

Field conditions: Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).

Location: State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).

Access & import/export: Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).

Disturbance: Describe any disturbance caused by the study and how it was minimized.

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 |