

Supplementary Materials for
**Cannabidiol inhibits SARS-CoV-2 replication through induction of the host
ER stress and innate immune responses**

Long Chi Nguyen, Dongbo Yang, Vlad Nicolaescu, Thomas J. Best, Haley Gula,
Divyasha Saxena, Jon D. Gabbard, Shao-Nong Chen, Takashi Ohtsuki, John Brent Friesen,
Nir Drayman, Adil Mohamed, Christopher Dann, Diane Silva, Lydia Robinson-Mailman,
Andrea Valdespino, Leticia Stock, Eva Suárez, Krysten A. Jones, Saara-Anne Azizi,
Jennifer K. Demarco, William E. Severson, Charles D. Anderson, James Michael Millis,
Bryan C. Dickinson, Savaş Tay, Scott A. Oakes, Guido F. Pauli, Kenneth E. Palmer,
The National COVID Cohort Collaborative Consortium, David O. Meltzer,
Glenn Randall*, Marsha Rich Rosner*

*Corresponding author. Email: grandall@bsd.uchicago.edu (G.R.); mrosner@uchicago.edu (M.R.R.)

Published 23 February 2022, *Sci. Adv.* **8**, eabi6110 (2022)
DOI: 10.1126/sciadv.abi6110

The PDF file includes:

Figs. S1 to S24
Patient Data Analysis Supplement
Tables S1 to S6
References

Other Supplementary Material for this manuscript includes the following:

Data S1

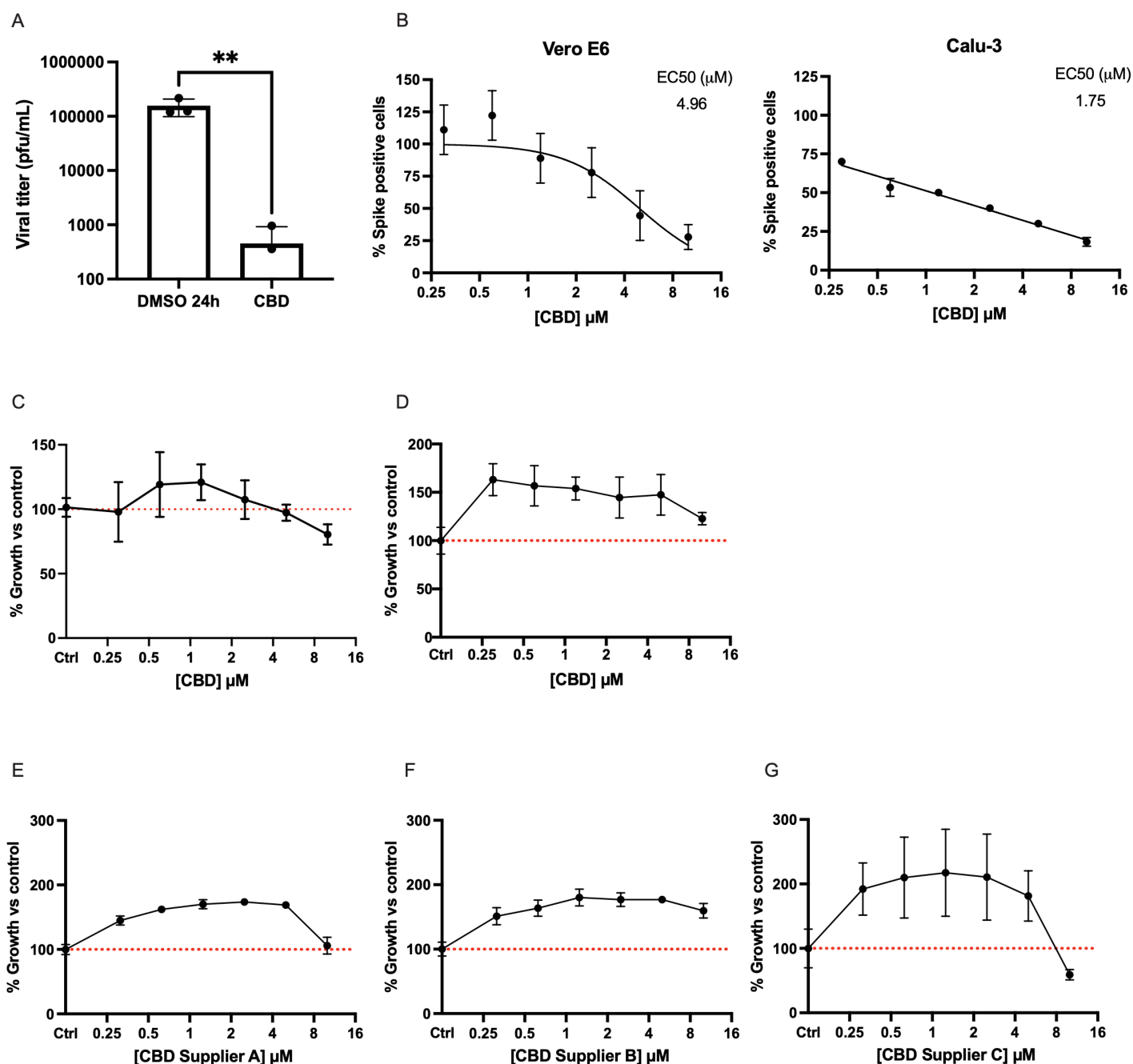


Fig. S1. CBD inhibits SARS-CoV-2 replication in different cell lines *in vitro*. (A) A549-ACE2 cells were treated with 10 μ M of CBD or DMSO followed by infection with SARS-CoV-2 at an MOI of 0.5 for 48 hours. Viral titers were determined from the supernatant. (B) Vero E6 cells (left) or Calu-3 cells (right) were treated with indicated doses of CBD followed by infection with SARS-CoV-2 at an MOI of 0.1 and .05 respectively for 24 or 48 hours, respectively. The cells were stained for spike protein and the percentage of cells expressing the spike protein in each condition was plotted. The EC50 value is indicated. (C) Vero E6 cells were treated with indicated doses of CBD for 48 hours. The cells were stained by Crystal Violet as described in Methods and viability levels were plotted as cell growth relative to untreated controls. (D) Calu-3 cells were treated with indicated doses of CBD for 48 hours. The cells were stained by Crystal Violet as described in Methods and viability levels were

plotted as cell growth relative to untreated controls. (E-G) A549-ACE2 cells were treated with indicated doses of CBD from three different suppliers A, B or C for 48 hours. Viability levels were plotted as described in (C).

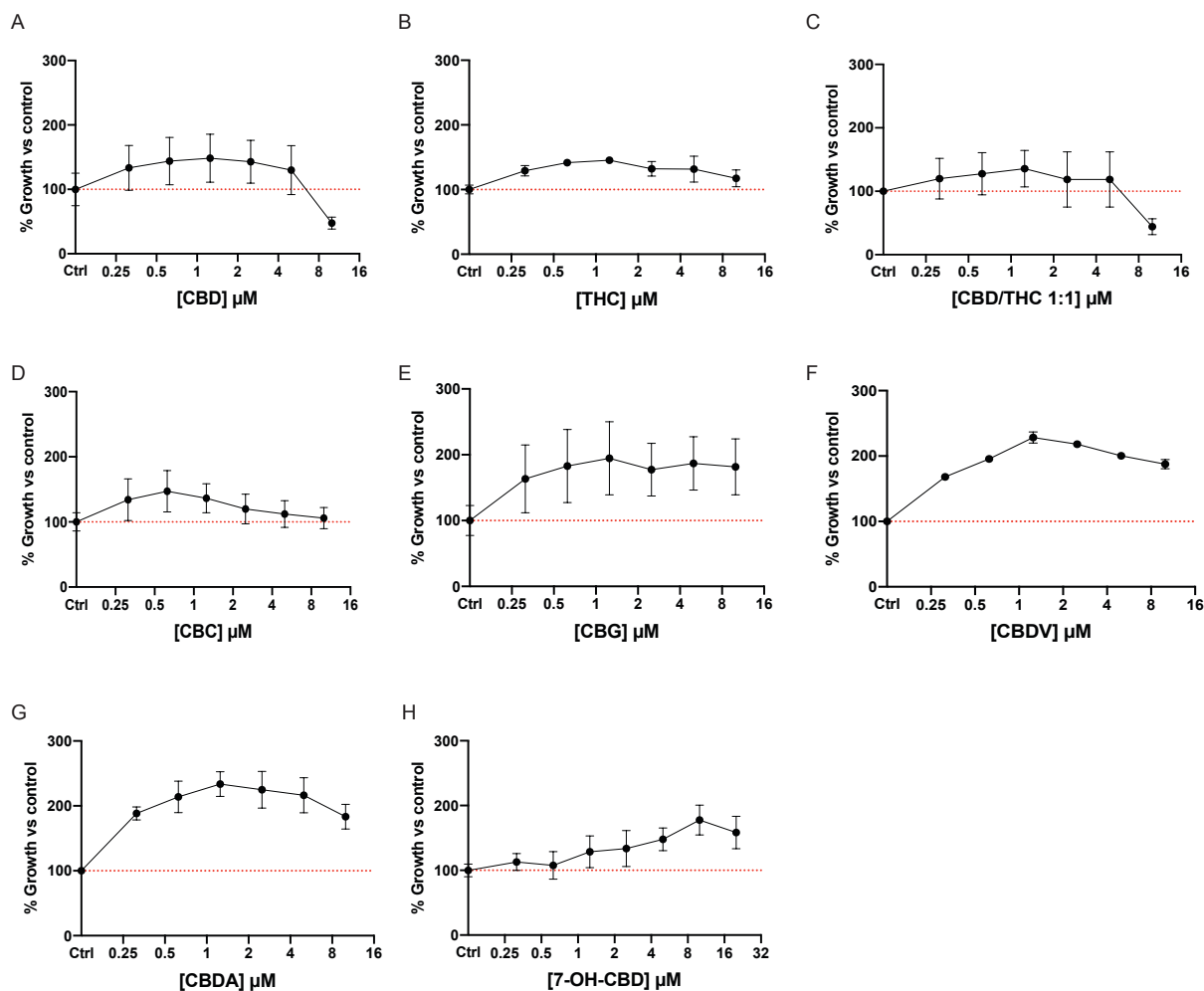


Fig. S2. Viability of A549-ACE2 cells for different cannabinoids. A549-ACE2 cells were treated with indicated doses of each cannabinoid for 48 hours. Cell viability was obtained as described in fig. S1C. (A-G) Viability data for cannabinoids used in Fig. 2A are shown. (H) Viability data for 7-OH-CBD used in Fig. 2C are shown.

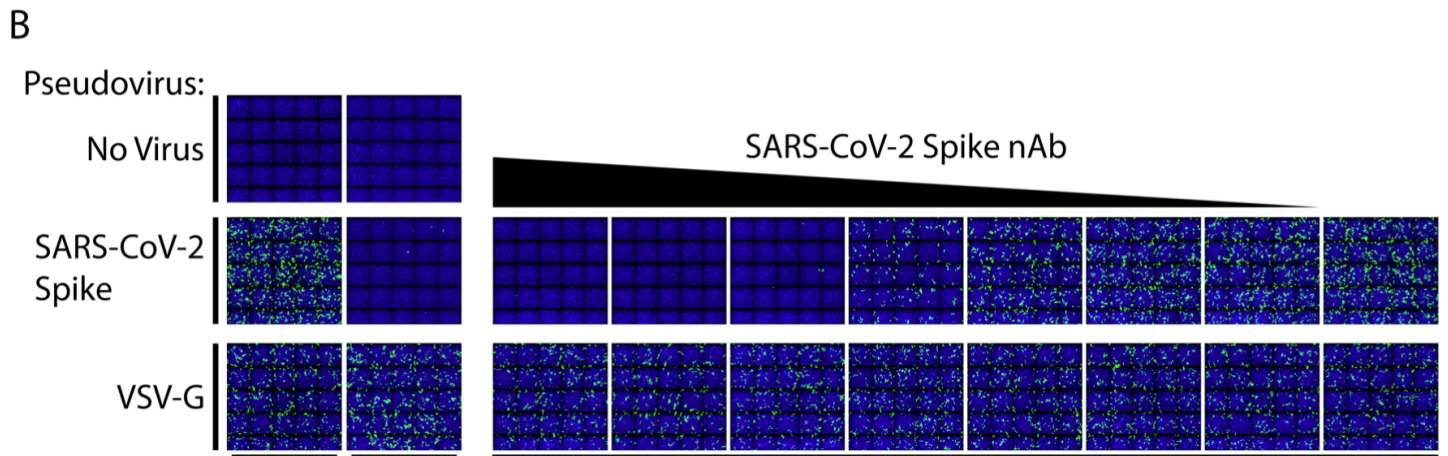
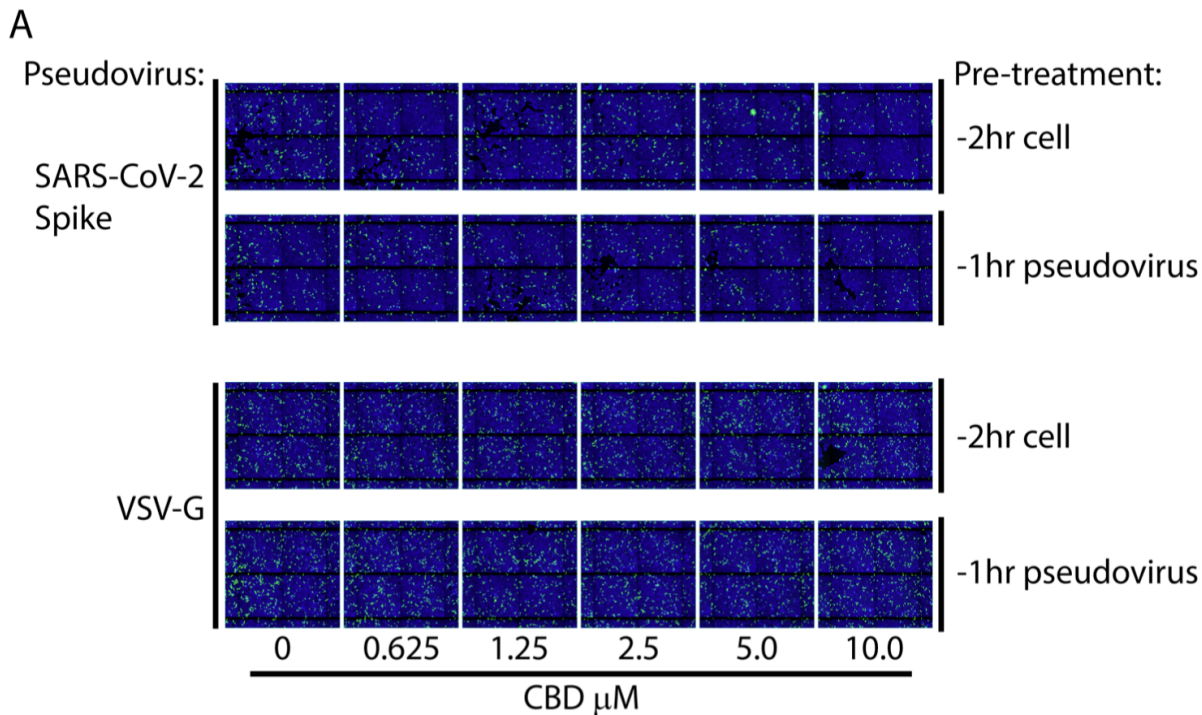


Fig. S3. CBD's inhibition of SARS-CoV-2 infection occurs after viral entry. (A) 293T+ACE2 cells were transduced with pseudovirus expressing either SARS-CoV-2 spike or VSV-G at 72hr post-infection following cell or pseudovirus treatment with CBD. Green = Pseudovirus infected, Blue = Nuclear marker Hoechst 33342. (B) 293T or 293T-ACE2 cells were infected for 72hrs with GFP-expressing lentiviruses pseudotyped with either SARS-CoV-2 spike or VSV-G. Pseudoviruses were incubated with 3-fold serial dilutions of SARS-CoV-2 spike neutralizing antibody (highest dose 300ng/100ul/well) for 1hr at 37C prior to infection. Green = Pseudovirus infected, Blue = Nuclear marker Hoechst 33342.

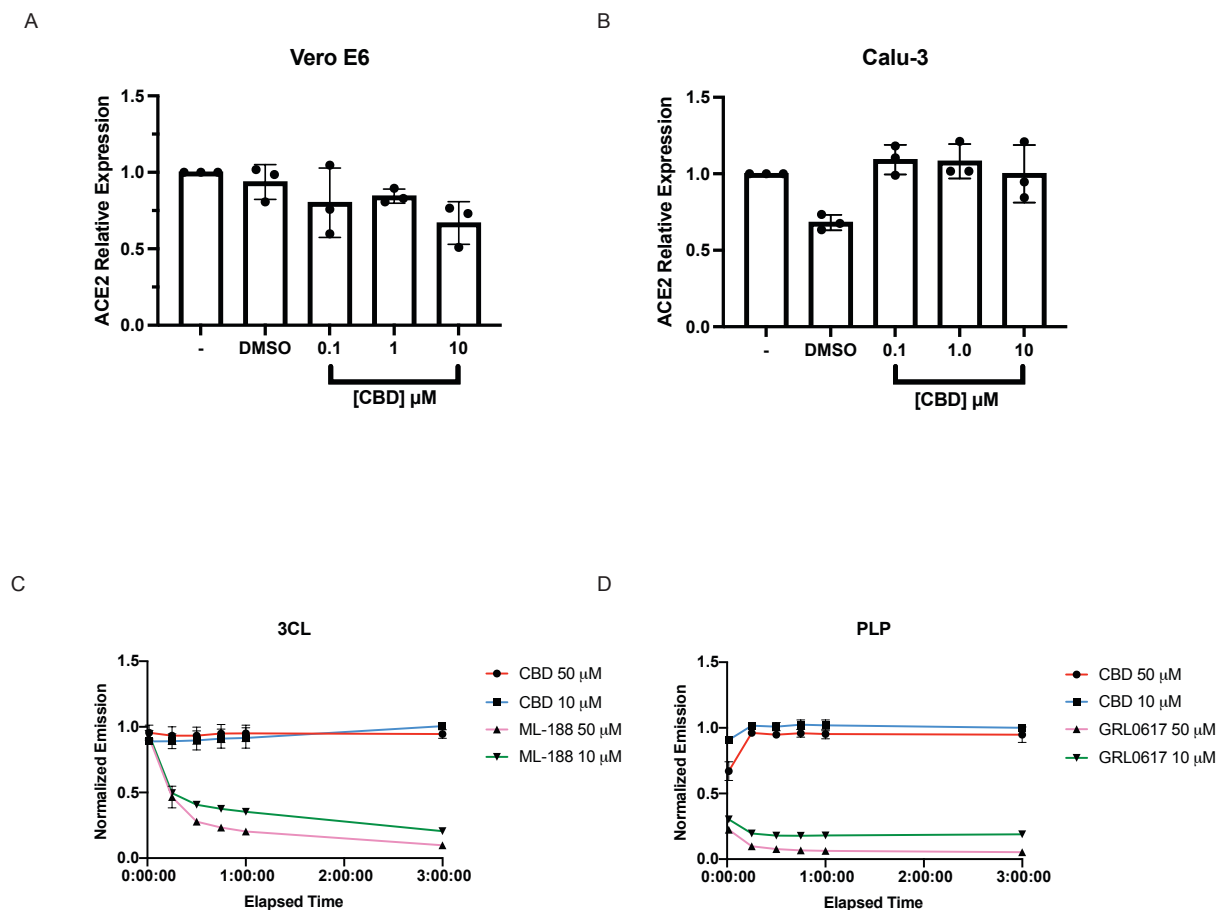
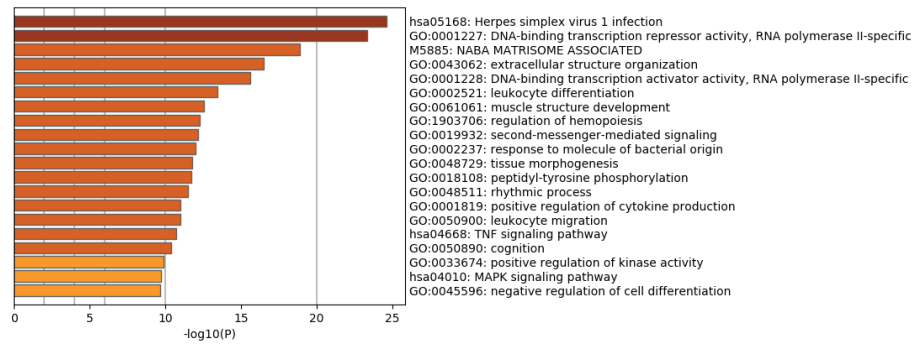


Fig. S4. CBD does not inhibit ACE2 expression or the two main viral proteases. Immunoblots of ACE2 protein expression from Vero E6 (A) or Calu-3 (B) cell lysates either untreated or treated with vehicle or CBD at indicated doses ($n=3$). Blots were probed with antibodies against ACE2 and tubulin. ACE2 protein expression levels were normalized to the tubulin signal within each sample. ACE2 expression levels were plotted relative to untreated samples. Enzyme assays were performed in duplicate containing indicated concentrations of CBD or positive control inhibitors and 3CLpro enzyme (0.4 μM , C) or PLpro enzyme (0.3 μM , D) as described in Methods. Data at each timepoint were normalized to the negative control (vehicle) wells and the normalized emission levels were plotted.

A



B

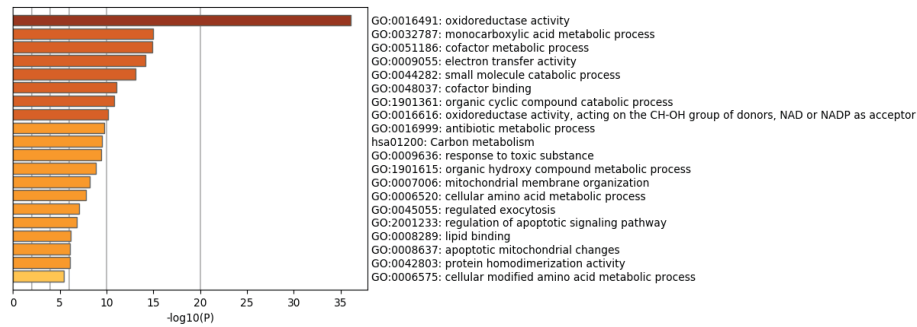


Fig. S5. Metascape analyses of genes with expression significantly altered by SARS-CoV-2 infection. Enriched sets of genes with expression significantly up-regulated (A) or down-regulated (B) by SARS-CoV-2 infection from the RNA-seq data were plotted. Differential expression for each gene was determined by the $\log_2(\text{fold change})$ of at least 1 in either direction with $q\text{-value} < 0.01$.

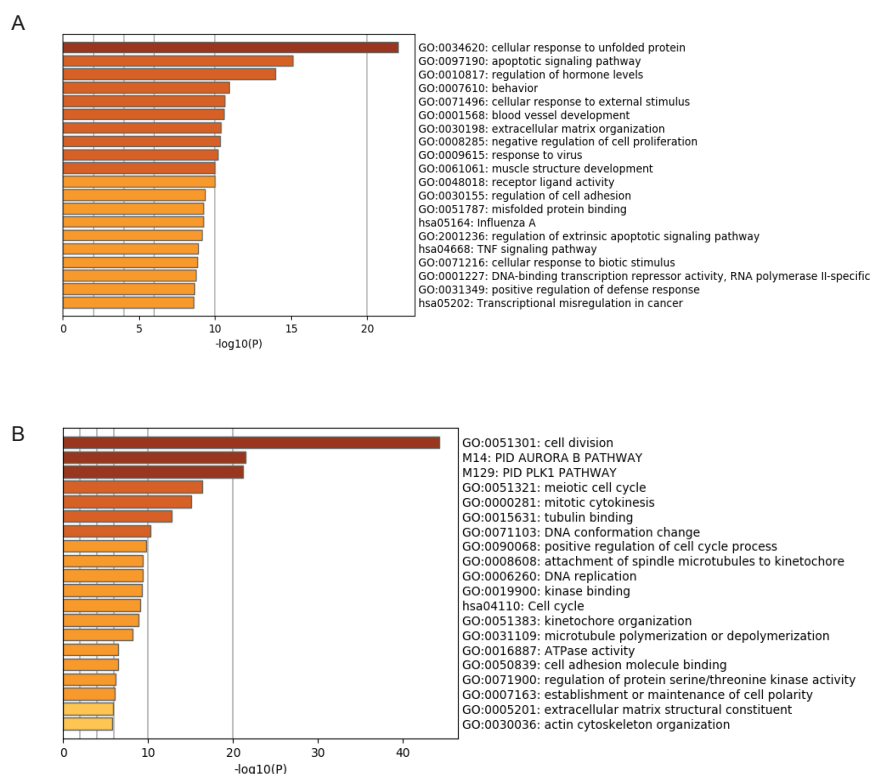


Fig. S6. Metascape analyses of genes with expression significantly altered by CBD treatment. Enriched sets of genes with expression significantly up-regulated (A) or down-regulated (B) by CBD treatment from the RNA-seq data were plotted. Differential expression for each gene was determined by the \log_2 (fold change) of at least 1 in either direction with q-value < 0.01.

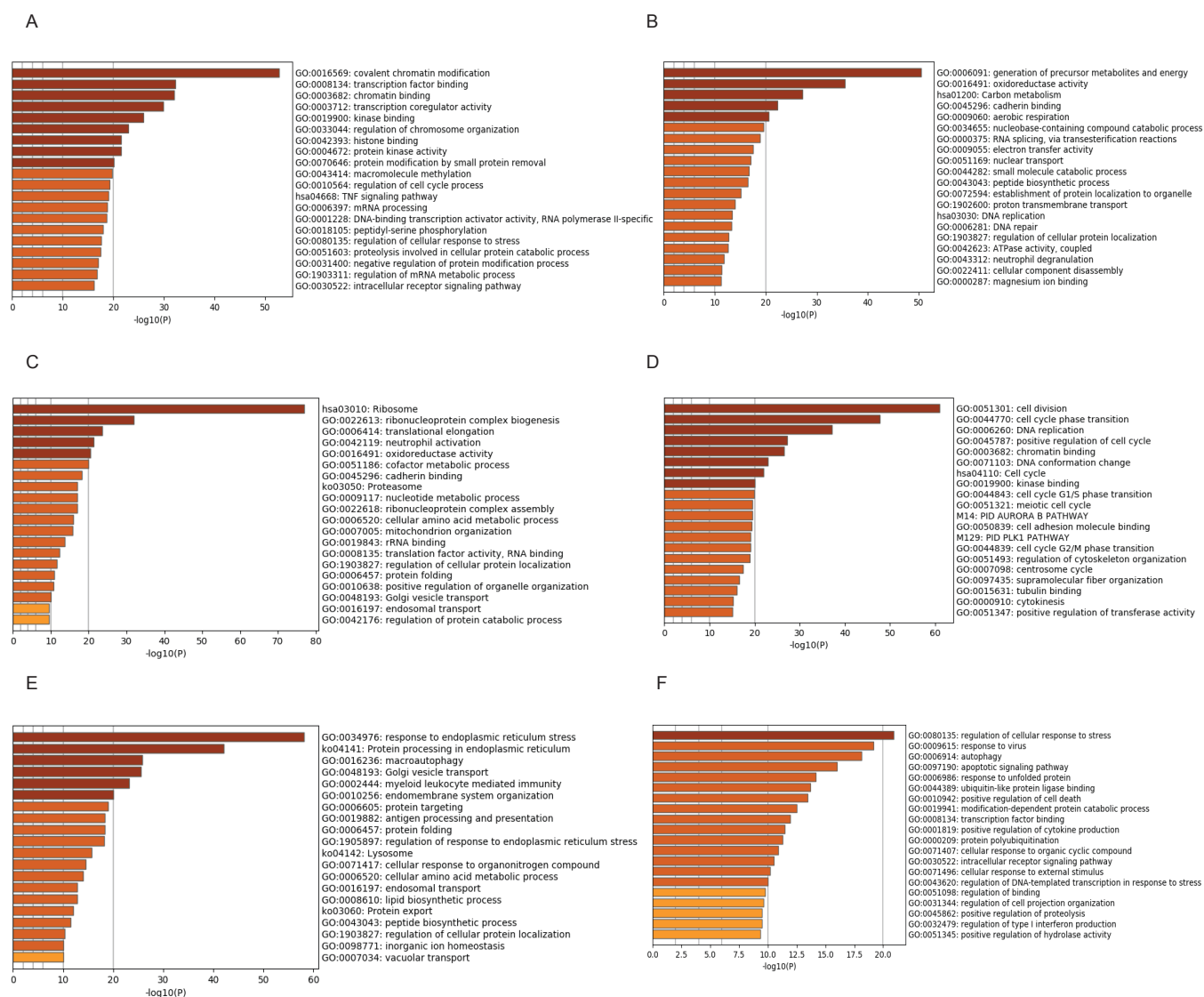


Fig. S7. Metascape analyses of the most variable genes among the RNA-seq samples. (A-F) Enriched sets of genes from each of the six clusters as described in Fig. 4D were plotted.

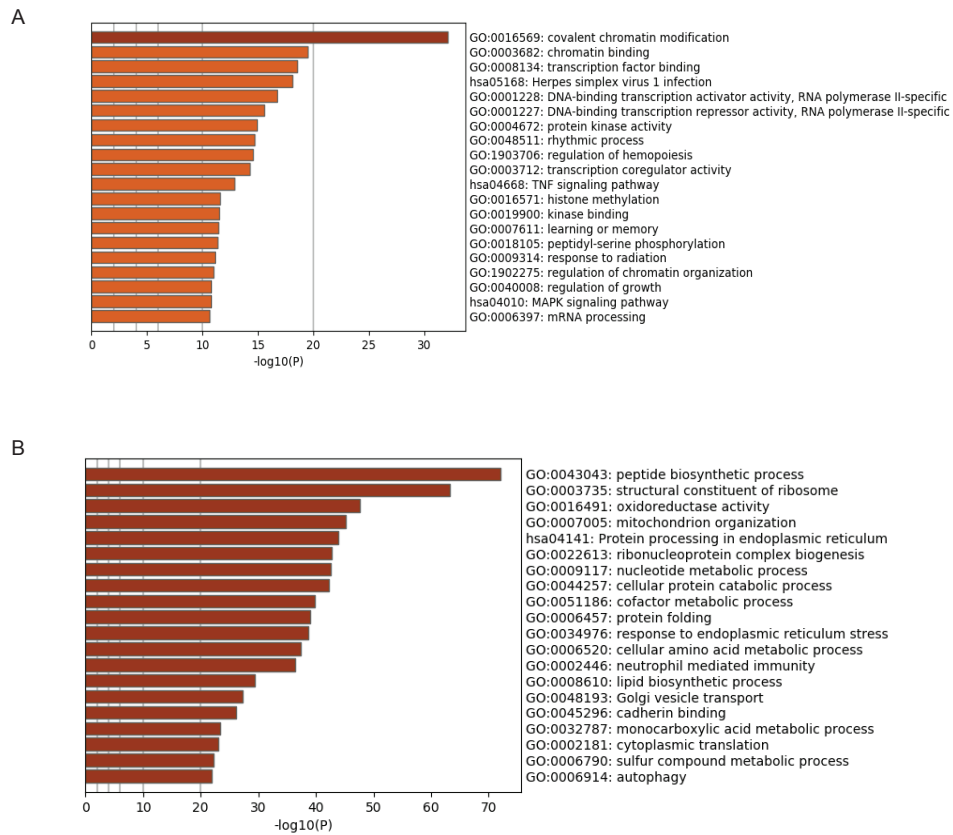
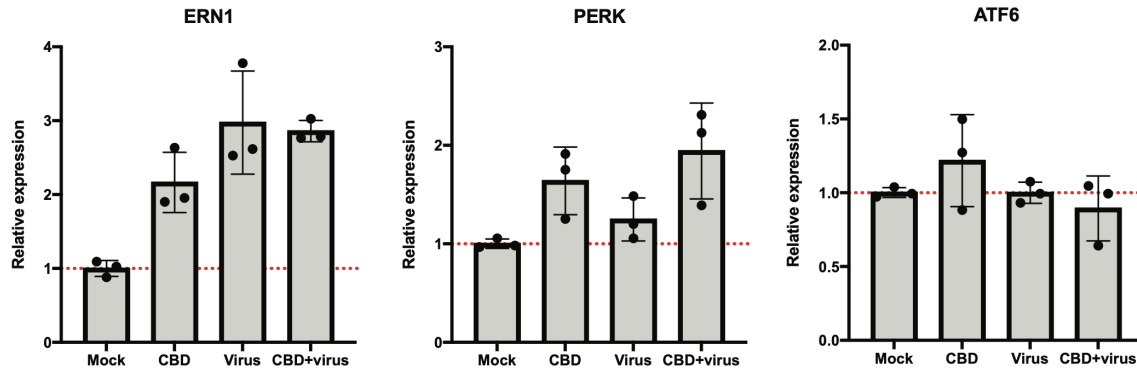


Fig. S8. Metascape analyses of genes with transcriptional changes caused by SARS-CoV-2 infection and reversed by CBD treatment. Enriched sets of genes with expression significantly up-regulated by viral infection but down-regulated by CBD treatment (A) or vice versa (B) were plotted. Differential expression for each gene was determined by the \log_2 (fold change) of at least 1 in either direction with q -value < 0.01 .

A



B

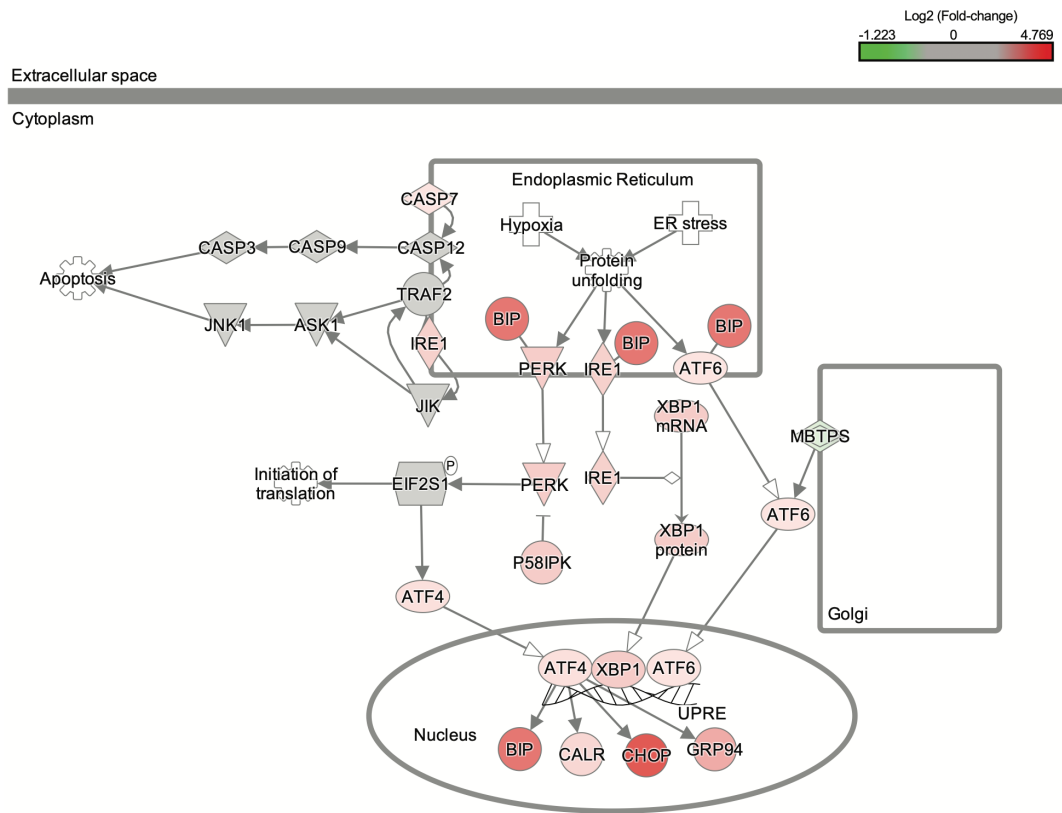


Fig. S9. Changes in the host ER stress canonical pathway following CBD treatment compared to untreated cells. (A) qRT-PCR was performed on samples from the RNA-seq experiment in Fig. 4 to determine expression levels of ER stress genes. (B) Ingenuity-generated ER stress pathways. Up-regulated genes are colored red and down-regulated genes are colored green. Genes without significant differential expression are colored gray. Color intensity corresponds to log₂(fold-change) from RNA-seq.

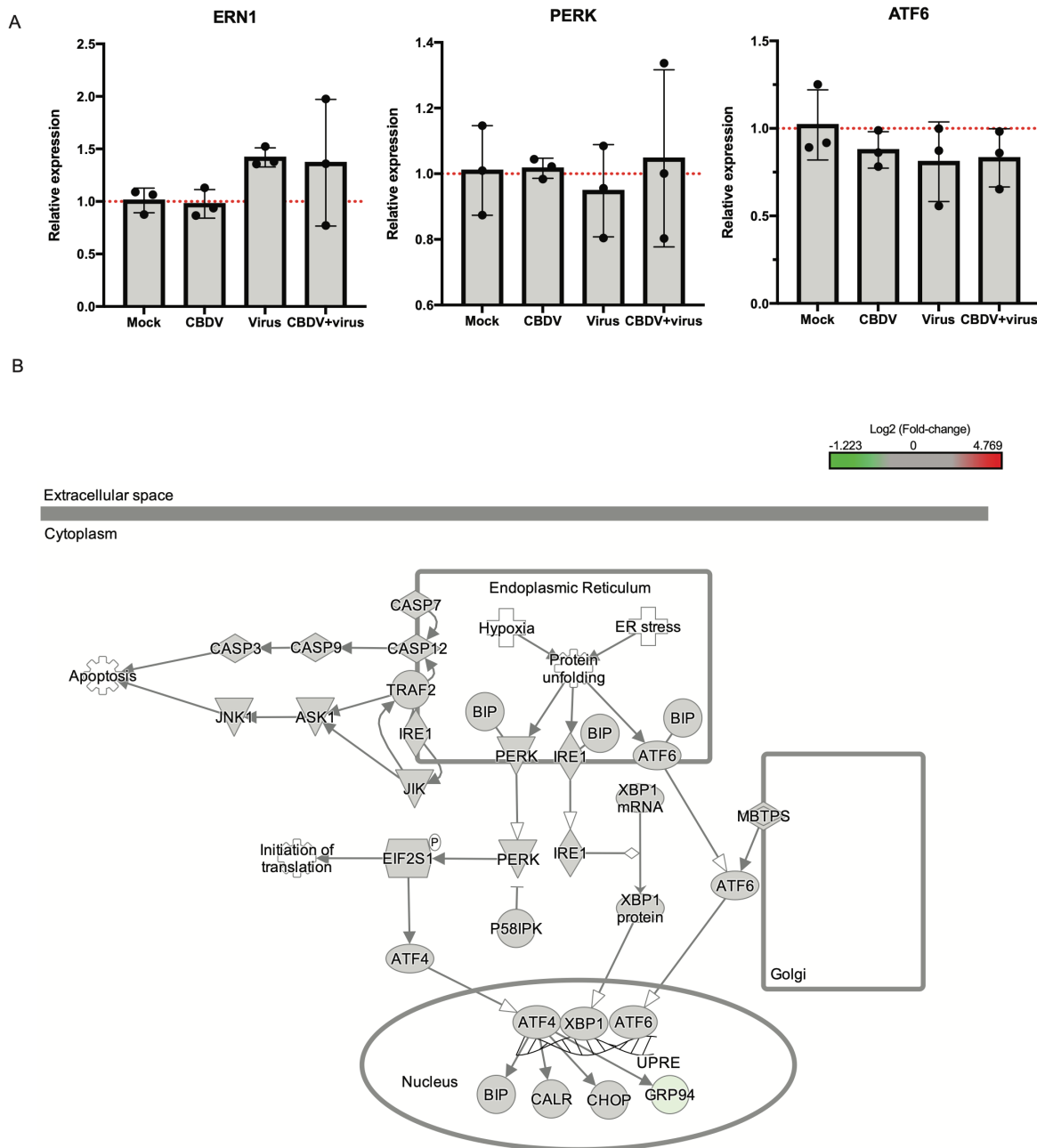


Fig. S10. Changes in the host ER stress canonical pathway following CBDV treatment compared to untreated cells. (A) qRT-PCR was performed on samples from the RNA-seq experiment in Fig. 5 to determine expression levels of ER stress genes. (B) Ingenuity-generated ER stress pathways. Up-regulated genes are colored red and down-regulated genes are colored green. Genes without significant differential expression are colored gray. Color intensity corresponds to $\log_2(\text{fold-change})$ from RNA-seq.

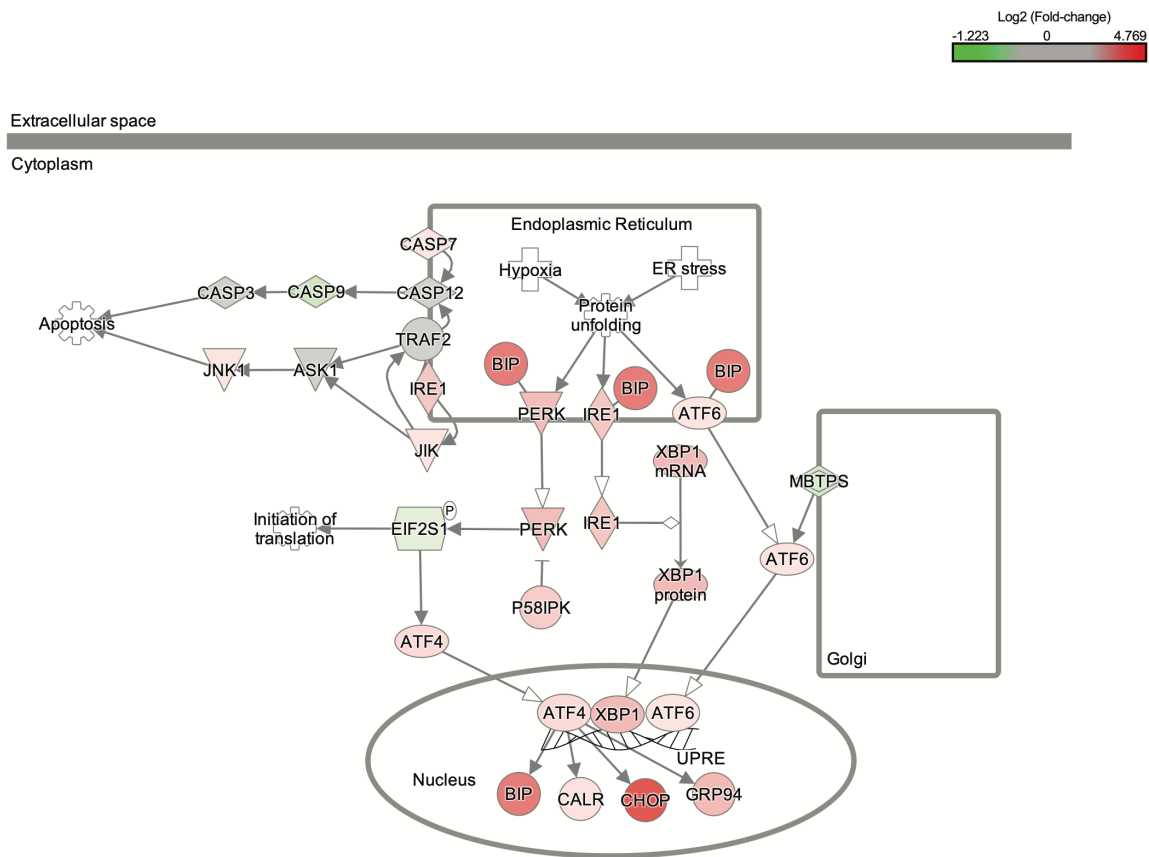


Fig. S12. Changes in the host ER stress canonical pathway following SARS-CoV-2 infection and CBD treatment compared to untreated cells. Ingenuity-generated ER stress pathways. Up-regulated genes are colored red and down-regulated genes are colored green. Genes without significant differential expression are colored gray. Color intensity corresponds to $\log_2(\text{fold-change})$ from RNA-seq.

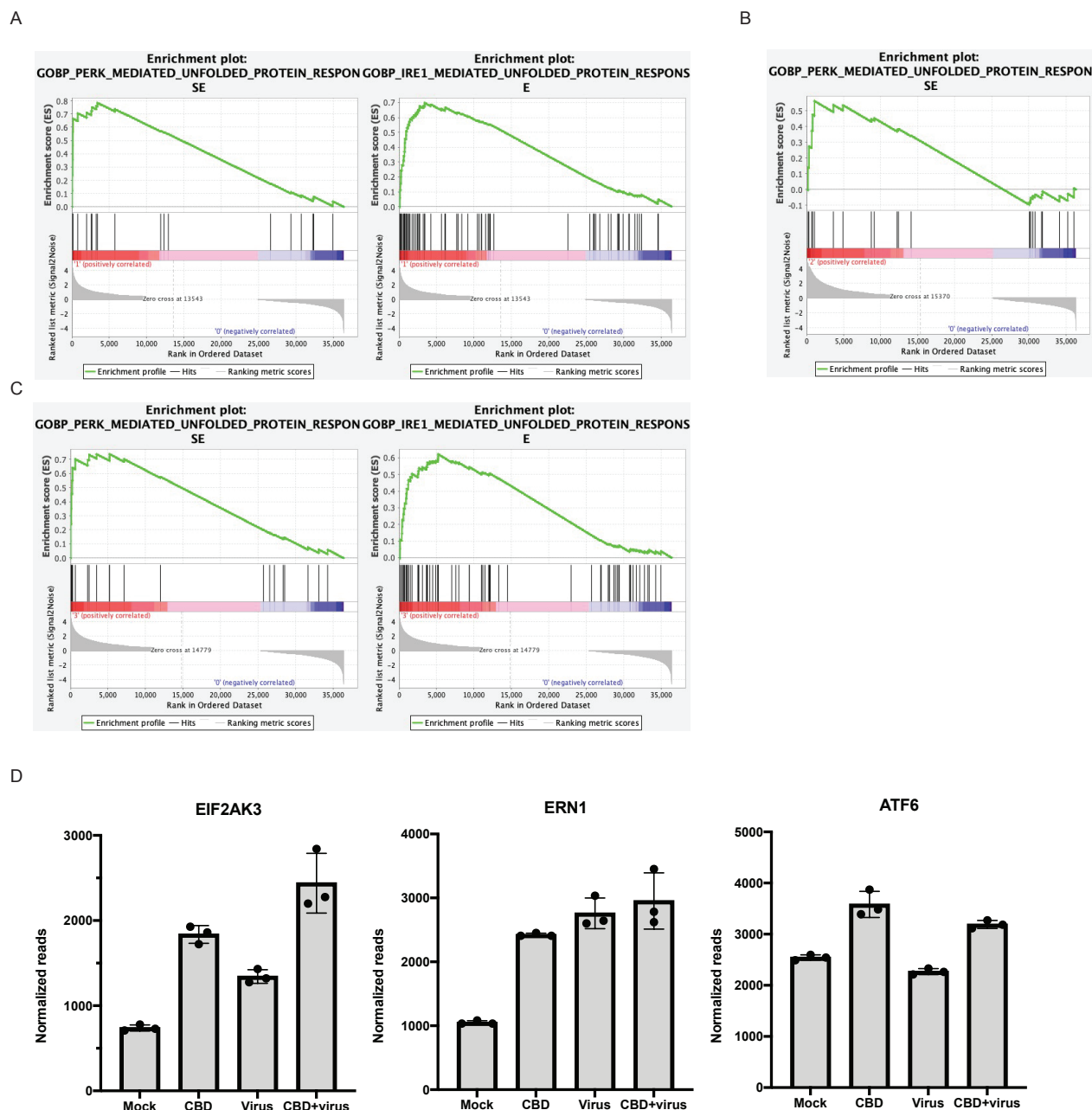


Fig. S14. Summary of changes in each branch of the unfolded protein response following CBD treatment and/or SARS-CoV-2 infection. Gene set enrichment plots of specified pathways are shown for expression changes in CBD vs mock (A), SARS-CoV-2 vs mock (B) and CBD+virus vs mock (C). Normalized enrichment score of each analysis is listed in Table 1. DESeq2 normalized reads from RNA-seq data for genes PERK (EIF2AK3), IRE1 (ERN1), and ATF6 are shown (D). Fold-changes of each comparison is listed in Table 1.

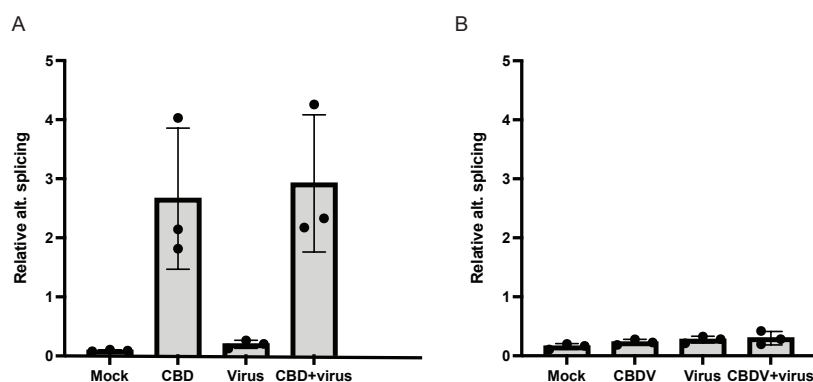


Fig. S15. CBD and CBDV differentially induce XBP1 alternative splicing as assessed by RT-PCR using RNA-seq samples. qRT-PCR was performed on samples from the CBD RNA-seq experiment in Fig. 4 (A) or the CBDV RNA-seq experiment in Fig. 5 (B) for spliced XBP1 and total XBP1. Relative XBP1 alternative splicing was determined as described in the Methods.

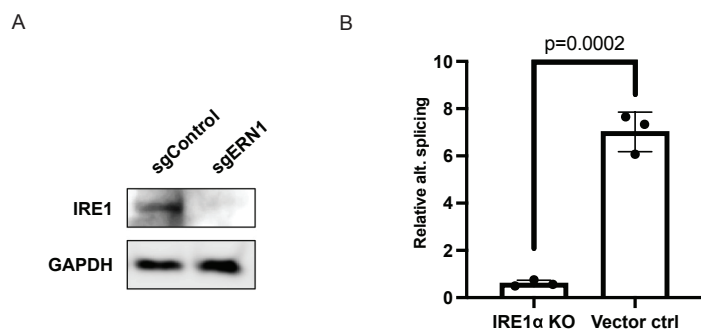


Fig. S16. Validations of IRE1 α KO in A549-ACE2 cells. (A) Immunoblot of IRE1 α and GAPDH protein expression from A549-ACE2 sgControl and sgERN1 cell lysates. Blots were probed with antibodies against IRE1 α and GAPDH. (B) Percentage of relative XBP1 alternative splicing in A549-ACE2 sgControl and sgERN1 cells determined by qRT-PCR as described in the Methods.

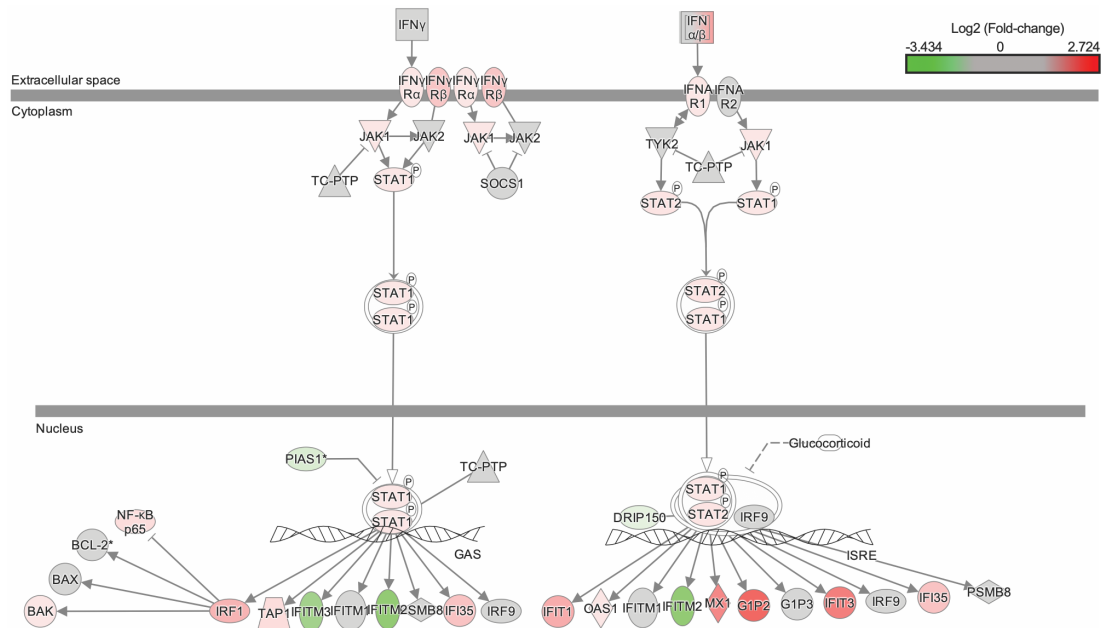


Fig. S18. Changes in the host interferon signaling canonical pathway following CBD treatment compared to untreated cells. Ingenuity-generated interferon signaling pathways. Up-regulated genes are colored red and down-regulated genes are colored green. Genes without significant differential expression are colored gray. Color intensity corresponds to $\log_2(\text{fold-change})$ from RNA-seq.

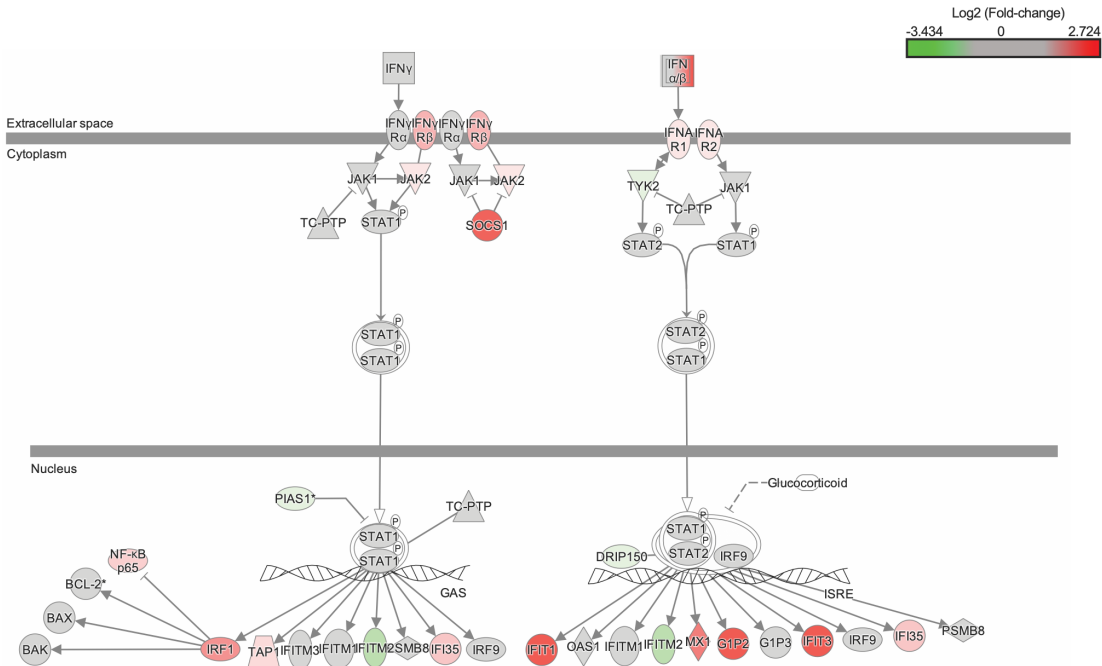


Fig. S19. Changes in the host interferon signaling canonical pathway following SARS-CoV-2 infection and CBD treatment compared to untreated cells. Ingenuity-generated interferon signaling pathways. Up-regulated genes are colored red and down-regulated genes are colored green. Genes without significant differential expression are colored gray. Color intensity corresponds to log₂(fold-change) from RNA-seq.

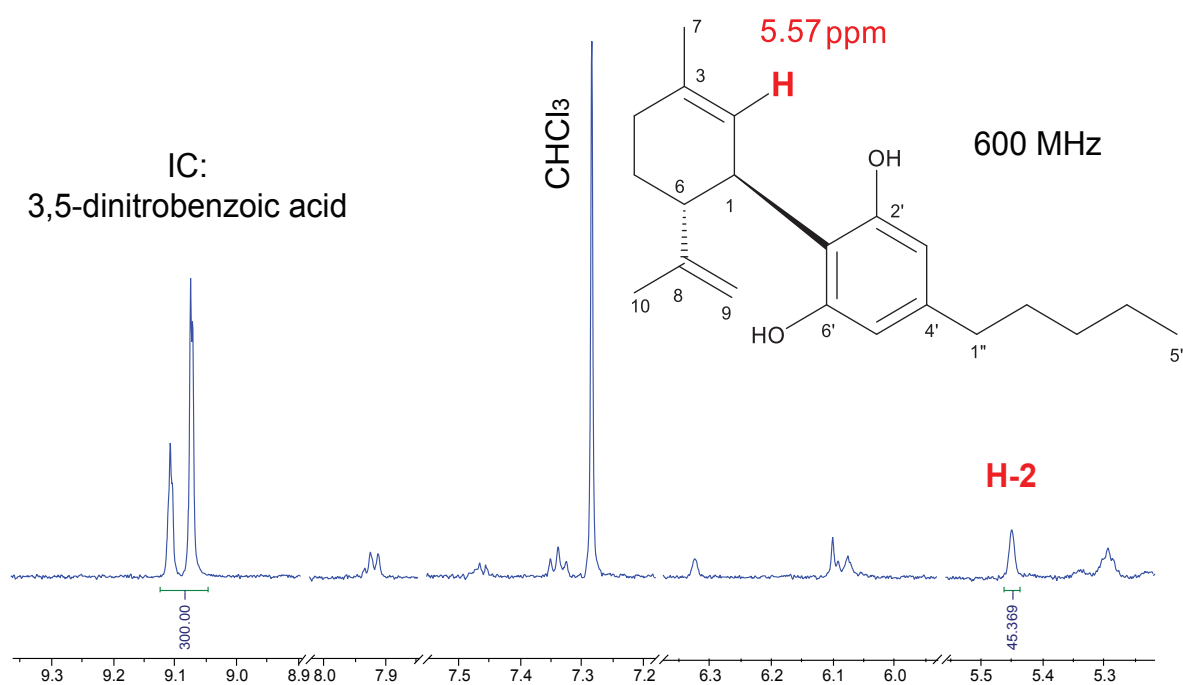


Fig. S22. ¹H NMR spectrum of hemp oil. The absolute qNMR method with internal calibration (3,5-dinitrobenzoic acid as the calibrant) was used to determine the content of CBD in hemp oil as 0.30%. Data was acquired on 600 MHz NMR instrument as described in Methods. Signal at 5.57 ppm (H-2 of CBD) was used for determining CBD quantity.

ANALYSIS OF PATIENT DATA

Institutional review board determination

The N3C data transfer to NCATS is performed under a Johns Hopkins University Reliance Protocol # IRB00249128 or individual site agreements with NIH. The N3C Data Enclave is managed under the authority of the NIH; information can be found at <https://ncats.nih.gov/n3c/resources>. The patient data analysis was determined to be exempt from further review by the University of Chicago Biological Sciences Division institutional review board (IRB21-0591), under the Federal Regulations (45 CFR 46) category 104(d)(4)(ii): secondary research for which consent is not required because the identities of the study participants cannot readily be ascertained by the investigators, the investigators do not contact the participants, and the investigators will not reidentify participants.

Methods

In essence, our methods were to identify patients who were likely using cannabidiol on the date of their COVID-19 testing and to compare their positive test rates with patients who were likely not using cannabidiol on the date of testing but were otherwise similar to the cannabidiol patients in terms of demographics, medical conditions, medication use, and other factors. This section defines our measurements of COVID-19 tests and COVID-19 test outcomes and of similarity between the cannabidiol patients and non-cannabidiol patients, our formation of a “Main Analysis Sample” that is a subset of the entire N3C patient population, and our statistical procedures to match cannabidiol and non-cannabidiol patients within the Main Analysis Sample and to compare their COVID-19 positive test rates.

Measurements

COVID-19 tests. We aimed to analyze only non-antibody COVID-19 tests and measurements, which were identified by manually reviewing the list of all COVID-19 test concepts within related N3C concept sets and excluding concepts with concept name substrings of “Ab” or “antibody” or similar. Our manual review resulted in the inclusion of all concepts in Table S1 below.

Table S1. N3C measurement concepts used to identify COVID-19 tests in our patient data analysis.

concept_id	concept_name
586516	SARS-CoV-2 (COVID-19) [Presence] in Specimen by Organism specific culture
586517	SARS-CoV-2 (COVID-19) whole genome [Nucleotide sequence] in Isolate or Specimen by Sequencing
586524	SARS-CoV-2 (COVID-19) N gene [Presence] in Respiratory specimen by Nucleic acid amplification using CDC primer-probe set N1
586526	SARS-CoV-2 (COVID-19) RNA [Presence] in Nasopharynx by NAA with probe detection
586529	SARS-CoV-2 (COVID-19) RNA [Cycle Threshold #] in Specimen by NAA with probe detection
700360	Infectious agent detection by nucleic acid (DNA or RNA); severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (Coronavirus disease [COVID-19]), amplified probe technique
704058	2019-nCoV Coronavirus, SARS-CoV-2/2019-nCoV (COVID-19), any technique, multiple types or subtypes (includes all targets), non-CDC, making use of high throughput technologies as described by CMS-2020-01-R
704059	Infectious agent detection by nucleic acid (DNA or RNA); severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (Coronavirus disease [COVID-19]), amplified probe technique, making use of high throughput technologies as described by CMS-2020-01-R
706154	SARS-CoV-2 (COVID-19) N gene [Presence] in Specimen by Nucleic acid amplification using CDC primer-probe set N2
706155	SARS-CoV-2 (COVID-19) N gene [Cycle Threshold #] in Specimen by Nucleic acid amplification using CDC primer-probe set N1
706156	SARS-CoV-2 (COVID-19) N gene [Presence] in Specimen by Nucleic acid amplification using CDC primer-probe set N1
706157	SARS-CoV-2 (COVID-19) N gene [Cycle Threshold #] in Specimen by Nucleic acid amplification using CDC primer-probe set N2
706158	SARS-CoV-2 (COVID-19) RNA panel - Respiratory specimen by NAA with probe detection
706160	SARS-CoV-2 (COVID-19) RdRp gene [Presence] in Respiratory specimen by NAA with probe detection
706161	SARS-CoV-2 (COVID-19) N gene [Presence] in Respiratory specimen by NAA with probe detection
706163	SARS-CoV-2 (COVID-19) RNA [Presence] in Respiratory specimen by NAA with probe detection
706169	SARS-CoV-2 (COVID-19) RNA panel - Specimen by NAA with probe detection
706170	SARS-CoV-2 (COVID-19) RNA [Presence] in Specimen by NAA with probe detection

706175	SARS-CoV-2 (COVID-19) N gene [Presence] in Specimen by NAA with probe detection
713859	Contact with and (suspected) exposure to COVID-19
715260	SARS-CoV-2 (COVID-19) RNA [Presence] in Saliva (oral fluid) by NAA with probe detection
715272	SARS-CoV-2 (COVID-19) N gene [Presence] in Nasopharynx by NAA with probe detection
723468	SARS-CoV-2 (COVID-19) S gene [Cycle Threshold #] in Specimen by NAA with probe detection
723476	SARS-CoV-2 (COVID-19) RNA [Presence] in Nasopharynx by NAA with non-probe detection
723477	SARS-CoV-2 (COVID-19) Ag [Presence] in Respiratory specimen by Rapid immunoassay
723478	SARS-CoV-2 (COVID-19) ORF1ab region [Presence] in Respiratory specimen by NAA with probe detection
742218	Infectious disease (bacterial or viral respiratory tract infection), pathogen-specific nucleic acid (DNA or RNA), 22 targets including severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), qualitative RT-PCR, nasopharyngeal swab, each pathogen reported as detected or not detected
742224	Infectious agent antigen detection by immunoassay technique, (eg, enzyme immunoassay [EIA], enzyme-linked immunosorbent assay [ELISA], immunochemiluminometric assay [IMCA]) qualitative or semiquantitative, multiple-step method; severe acute respiratory syndrome coronavirus (eg, SARS-CoV, SARS-CoV-2 [COVID-19])
756055	Measurement of Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)
757677	SARS-CoV-2 (COVID-19) RNA [Presence] in Nose by NAA with probe detection
757678	SARS-CoV-2 (COVID-19) N gene [Presence] in Nose by NAA with probe detection
757685	SARS-CoV+SARS-CoV-2 (COVID-19) Ag [Presence] in Respiratory specimen by Rapid immunoassay
4034850	Laboratory test
36661375	Influenza virus A and B and SARS-CoV-2 (COVID-19) identified in Respiratory specimen by NAA with probe detection
36661377	SARS-CoV-2 (COVID-19) RNA [Presence] in Respiratory specimen by Sequencing
40218804	2019-ncov coronavirus, sars-cov-2/2019-ncov (covid-19), any technique, multiple types or subtypes (includes all targets), non-cdc
40218805	Cdc 2019 novel coronavirus (2019-ncov) real-time rt-pcr diagnostic panel

COVID-19 antibody tests were excluded from all analysis because they may indicate COVID-19 disease at a less-certain and/or more-distant time in the past than other COVID-19 tests (e.g., sometime within the past few weeks), and thus these tests are less useful than non-antibody tests for our purpose of assessing associations between use of cannabidiol and COVID-19 test results at the time of potential COVID-19 exposure. The most common Table S1 concepts in the N3C data were 700360, 706160, 706163, 706170, and 723477, whereas all others were used relatively rarely. We included concept 4034850 because it was marked as “Related” in ATLAS (45) to many COVID-19 test concepts, but also considered rationale for excluding it. Upon examination of the effects of including 4034850, less than 20 of the cannabidiol patients in our main analysis (i.e., too small to specify precisely and adhere to N3C guidelines) would have their First COVID-19 Test or Dx Date (described below) shifted by removing this concept from the list and only by at most 2 days. Also, none of the cannabidiol patients would have their COVID-19 status (positive or negative) changed by removing this concept. Furthermore, no new cannabidiol patients would be added to the main analysis by removing this concept.

First COVID-19 Test or Dx Date. To mitigate potential confounding from intra-patient dynamics, such as changing COVID-19 risk over time or over multiple SARS-CoV-2 tests, we defined each patient’s First COVID-19 Test or Dx Date as their earliest non-antibody SARS-CoV-2 test measurement date (from the N3C measurement table) or their first COVID-19 condition start date (from the condition_occurrence table) if that date was earlier than the first COVID-19 test date. Patients first COVID-19 condition start dates were infrequently earlier than their first COVID-19 test date. COVID-19 conditions were identified using condition concept 37311061 (“COVID-19”, or “Disease caused by 2019-nCoV” in ATLAS (45). We used this First COVID-19 Test or Dx Date as a reference point for additional components of our analysis, describe below.

COVID-19 Status. A patient with a positive non-antibody SARS-CoV-2 test result or a COVID-19 diagnosis within 14 days before or after their First COVID-19 Test or Dx Date was assigned a COVID-19 Status of Positive, otherwise, if the patient had only negative COVID-19 tests within those 29 days (-14 through +14), they were assigned a COVID-19 Status of Negative. A particular test result was defined as positive if the result was labeled with the words “positive” and/or “detected”, and negative if the result was labeled with the words “negative” and/or “not detected”.

Cannabidiol medications. As an initial, formative step in our analysis, we manually reviewed records in the N3C drug_exposure and drug_era tables for medication that may contain cannabidiol, to generate a cohort of patients who may have taken medication containing primarily cannabidiol around the time of their First COVID-19 Test or Dx Date. We were not sure initially how such medication records would be labeled in the N3C data, so we searched all records in the drug_exposure table for drug_concept_names that contained any of the following substrings (case insensitive): *canemes, cannabi, cannibi, cbd, cesamet, dextrabinol, dronabinol, epidiolex, hemp, marijuana, marinol, nabilone, nabiximol, sativex, syndros, thc*. Over 99.98% of the resulting drug_concept_names contained either “dronabinol” or “cannabidiol” (never both), and most of the remainder contained either the word “nabilone” or “hemp” (Ns too small to report). Dronabinol and nabilone contain synthetic tetrahydrocannabinol (THC) and do not regularly contain cannabidiol and therefore we did not include them in our medications of interest. The hemp medications were very infrequent and likely different in composition than the remaining cannabidiols, so we focused only on 3 of the concepts in the search results, shown in Table S2.

Table S2. Concepts used to define cannabidiol medications, in decreasing order of frequency.

drug_concept_id	drug_concept_name
35200836	cannabidiol 100 MG/ML Oral Solution
35200842	cannabidiol 100 MG/ML Oral Solution [Epidiolex]
1510417*	cannabidiol

* Less than 10% of the records of drug_concept_id 1510417 were removed from further analysis because their drug_source_concept_name was “*bamifylline Oral Tablet [BAMIDIL]*”, which suggested that the medication could have been miscoded by N3C data contributors, because the trade name is incorrect (correct is Bamifex) and bamifylline is not known to contain cannabidiol. All other source names were “*cannabidiol*”.

Epidiolex^R is reportedly the first and only FDA-approved prescription cannabidiol (<https://www.epidiolex.com/>). It is an oral solution available in a concentration of 100 mg per mL, and there are no known FDA-approved generic alternatives to date. Hence, it is plausible (but unverifiable) that records of drug_concept_id 35200836 were actually Epidiolex^R coded without the trade name. Concept 1510417, cannabidiol without a corresponding concentration or route, could be a wide variety of medications and products, including self-reported Epidiolex^R medications for which the patient did not know or clearly explain the concentration, or multiple over-the-counter products such as oral solutions, tablets, topical creams, or other products with variable and uncertain cannabidiol concentrations and purities. Under the assumption that the heterogeneity in COVID-19 outcomes would be less if we included these cannabidiol medications of unknown concentration and route, we chose to focus primarily on concepts 35200836 and 35200842 for our analysis: cannabidiol 100 mg/mL records, and to conduct some sensitivity analysis with also including concept 1510417.

Cannabidiol 100 mg/mL patients. We defined 2 dichotomous indicator variables to classify whether a patient used cannabidiol 100 mg/mL or not (setting the variable equal to 1 or 0, respectively). They were calculated based on the presence and type of cannabidiol 100 mg/mL oral solution drug records and their proximity to patients’ First COVID-19 Test or Dx Dates. Below we describe how each was calculated and the associated rationale.

1. *cannabidiol100_ever patients.* Patients with at least 1 record in the drug_exposure or drug_era tables with a drug_concept_id of either 35200836 (“cannabidiol 100 MG/ML Oral Solution”) or 35200842 (“cannabidiol 100 MG/ML Oral Solution [Epidiolex]”) had their value of cannabidiol100_ever set equal to 1, and all other patients had it set equal to 0.
2. *cannabidiol100_dur_c19 patients.* This variable identified the subset of cannabidiol100_ever patients who we hypothesized were likely to have had *active* cannabidiol 100 mg/mL use on the date of their First COVID-19 Test or Dx Date. Multiple definitions of *active* were considered plausible given the extent of missing information related to the medication records. For example, the records were often void of information such as the days-supply of a cannabidiol 100 mg/mL prescription, the number of refills prescribed, the dose per use (which is known to vary according to patient weight), or the total quantity of oral solution per refill. Importantly, the end date of the record was often missing, which could indicate the patient was still taking the

cannabidiol 100 mg/mL, but could also indicate an old, short-term prescription that ran out but was not documented as such, for example. Furthermore, some patients who only had start dates after their First COVID-19 Test or Dx Date could have, in fact, been taking cannabidiol 100 mg/mL before their COVID-19 test but the medication reconciliation process administered around the time of their COVID-19 test did not accurately indicate this prior use. We settled upon a conservative definition: we defined a cannabidiol 100 mg/mL record as “active” during their First COVID-19 Test or Dx Date only if either (the drug had a recorded start date before the First COVID-19 Test or Dx Date and the drug had a recorded end date on or after the First COVID-19 Test or Dx Date) or (the drug had a recorded start date equal to the First COVID-19 Test or Dx Date). The recorded start and end dates were taken from the drug_exposure and drug_era N3C tables; according to OMOP documentation (46), drug era table entries should be constructed by applying rules to the drug_exposure records to generate continuous periods of active medication records; however, we did not want to assume that the drug_exposure records were used correctly and consistently for each data contributing partner/site to generate drug eras, so we used both tables to construct our definition of an active cannabidiol 100 mg/mL medication. This definition is conservative in the sense that it does not assume a cannabidiol 100 mg/mL record was active if it had a start date strictly before the First COVID-19 Test or Dx Date but no end date; the only records with no end date that were assumed to be active were those with a start date equal to the First COVID-19 Test or Dx Date. In our statistical analysis, when matching controls to patients with an active cannabidiol 100 mg/mL record, we allowed the matching algorithm to select patients with non-active cannabidiol 100 mg/mL records as controls. If these patients were chosen as matches and were, in fact, actively taking cannabidiol 100 mg/mL on their First COVID-19 Test or Dx Date, our conservative definition would reduce our potential to detect a significant negative association between cannabidiol 100 mg/mL use and COVID-19 Positive status, because some of the matched controls could have been actively taking cannabidiol 100 mg/mL. In the Results section below, we report how many matched controls may have been taking cannabidiol 100 mg/mL during their First COVID-19 Test or Dx Date and discuss this further.

Using a lookback of at least 15 days to identify conditions and non-cannabidiol medications of interest. We set all indicator variables for conditions and non-cannabidiol medications of interest equal to 1 if the patient had an associated record in the N3C condition_occurrence or drug_exposure tables with a start date of at least 15 days before the patient’s First COVID-19 Test or Dx Date. We did not count conditions or medications with a start date within 14 days before their First COVID-19 Test or Dx Date, on that date, or any number of days after that date to reduce potential confounding of condition or medication prevalence by early manifestations of COVID-19. We aimed to avoid counting conditions or medications that occurred, were used, or were otherwise documented only because the patient had COVID-19 or its symptoms. In other words, by only considering conditions and medications with a start date at least 15 days before their First COVID-19 Test or Dx Date, we aimed to only consider conditions that were present before the patient was first tested for COVID-19 and/or contracted the SARS-CoV-2 virus. The threshold of 15 days was used because of general recommendations by the CDC and other public health groups that persons should watch for COVID-19 symptoms until 14 days after exposure (47).

Conditions. We included two groups of condition-related dichotomous indicator variables as covariates in our statistical matching procedures and comparison of COVID-19 positive test rates: (1) indicators for conditions associated with increased COVID-19 risk and/or severity according to the Centers for Disease Control and Prevention (CDC), and (2) indicators for conditions common among cannabidiol patients. For both groups, the indicator variable was set equal to 0 unless the person_id had at least one record in the condition_occurrence table with a start date at least 15 days before their First COVID-19 Test or Dx Date and a condition_concept_id that was within a defined concept set for that indicator, in which case the indicator variable was set equal to 1. Below explains how we define the concept sets for each group.

1. *Conditions associated with high COVID-19 risk and/or severity according to the CDC.* To decide which conditions to included, we consulted a list of at-risk conditions maintained by the CDC (29). For all conditions on that list, we aimed to form one or more related concept sets of condition_concept_ids. As a basis for the conditions included in each concept set, we were often able to rely upon externally tested and published sets of ICD-10-CM diagnoses codes that constitute the condition according to the HCUP Elixhauser Comorbidity Software (Agency for Healthcare Research and Quality). Condition concept IDs were added to the concept set for all ICD-10-CM codes listed under categories in the Elixhauser file named *Comorb_ICD10CM_Format_v2021-1.sas*, titled “Creation of Format Library for Comorbidity Groups, ICD-10-CM Comorbidity Software, Version 2021.1.” The ICD-10-CM codes included in each category can be found by looking up the category names from the middle column of Table S3 below in this SAS file on the AHRQ HCUP webpage.

Table S3. At-risk medical conditions[§] associated with severe COVID-19 outcomes according to the CDC (29) and details of how we define them for our statistical analysis.

Condition indicator variable name in our statistical analysis	AHRQ HCUP Elixhauser ICD-10-CM code category name(s) used to identify the initial set of conditions included in this condition indicator variable*	N3C concept set ID number(s)
Arthritis	ARTH	380906623
Cancer	CANCER_LEUK, CANCER_LYMPH, CANCER_METS, CANCER_SOLID	721859940
Cerebrovascular disease or stroke	CBVD_POA, CBVD_SQLA, CBVD_SQLAPARALYSIS	905306000
Chronic kidney disease	CHFHTN_CXRENFL_SEV, HTN_CXRENFL_SEV, RENLFL_MOD, RENLFL_SEV	423682967
Chronic lung disease	LUNG_CHRONIC	299455728 plus concept 4177862
Dementia	DEMENTIA	110790697
Depression	DEPRESS	743223706
Developmental disabilities	No Elixhauser category existed, so we created a concept set based on all descendants of ATLAS condition concept ID 435244, “Developmental Disorder” plus a few concepts that were Related in ATLAS to non-standard developmental disabilities that were common among cannabidiol100_ever patients. We then imported this concept set into N3C per standard N3C import methods, which caused the concepts to narrow to only those in N3C.	257316444 (plus concepts 4031877, 4173811, and 4098458)
Diabetes	DIAB_CX, DIAB_UNCX	25948849 and 314686290
HIV or AIDS	AIDS	210615227
Hypertension	CHFHTN_CX, CHFHTN_CXRENFL_SEV, HTN_CX, HTN_CXRENFL_SEV, HTN_UNCX	187975095 and 779214702
Liver disease	ALCOHOLLIVER_MLD, LIVER_MLD, LIVER_SEV	771461026
Neurologic, movement disorders	NEURO_MOVT	377876370
Neurologic, other disorders	NEURO_OTH	271697283
Pregnancy	No Elixhauser category existed, so we used a concept set created by another N3C user: Andrew Girvin.	326616819
Psychoses	DRUG_ABUSEPSYCHOSES, PSYCHOSES	352933736
Pulmonary circulatory disorders	PULMCIRC	945923578
Sickle cell disease	No Elixhauser category existed, so we used a concept set created by another N3C user: Joy Alamgir.	662958201
Smoker, former or current	No Elixhauser category existed, so we used a concept set created by another N3C user: Andrew Girvin.	506560962
Substance use disorder	ALCOHOL, ALCOHOLLIVER_MLD, DRUG_ABUSE, DRUG_ABUSEPSYCHOSES	144104309
Transplant recipient status	No Elixhauser category existed, so we created a concept set by searching ATLAS for “transplant” and narrowing the results to all standard condition concepts with non-zero records in ATLAS, plus their direct descendants. We then imported this concept set into N3C per standard N3C import methods, which caused the concepts to narrow to only those in N3C.	275526159
Trisomy 21	No Elixhauser category existed, so we used a Down Syndrome Trisomy 21 concept set created by another N3C user: Seth Russell.	933343675

* Some conditions did not have related AHRQ HCUP Elixhauser categories. For these, we either used a concept set produced by another user of N3C and gave recognition to that user in our Acknowledgments, or created our own concept set as described in the table.

§ Overweight and obese conditions are included on the CDC website as at-risk conditions for severe COVID-19 outcomes, but we do not include them in this table, because we instead control for being overweight or obese through a BMI categorical variable described below.

Beyond the concepts related to the ICD-10-CM codes, we also expanded each concept set to include condition_concept_ids found using the COVID-19 ATLAS search engine and concept set creation tool by the Observational Health Data Sciences Institute (45). In detail, we searched ATLAS for other similar condition concept IDs and added to our concept set those that were similar to the concepts associated with its ICD-10-CM codes, by typically added all Observational Medical Outcomes Partnership (OMOP) standard condition concept IDs with records that were “Related” to the ICD-19-CM non-standard concepts according to ATLAS and/or had keywords in their concept name that were similar and consistent with keywords in the non-standard concept names identified through the AHRQ Elixhauser codes. Most of the resulting concept sets contain hundreds or thousands of condition_concept_ids, and therefore we do not list the concepts here. Instead, any user with N3C access can observe the concepts within these sets by searching for the concept set IDs in the rightmost column of Table S3 with the “Browse Versions” button of the “Concept set browser” linked on the N3C homepage.

2. *Conditions common among cannabidiol patients.* We generated a list of all condition_concept_ids from the condition_occurrence table that were linked (through N3C’s person_id identifier) to more than 5% of all cannabidiol100_ever patients. The resulting list had 184 conditions, which are shown in Table S4 in the Results section below. In our main statistical analysis, we did not aggregate any of these conditions nor did we remove conditions from the list that may have been partially correlated with the prior group of conditions that may cause increased COVID-19 risk according to the CDC (e.g., we kept concept 40486120 “Delay in physiological development” on the list even though it was also included in our “developmental disabilities” concept set). However, we did use concept 433736 “Obesity” from the list in our main statistical analysis, because we presumed it would be very highly correlated with another variable that we did use in the statistical analysis to indicate whether a patient’s Body Mass Index (BMI) was greater than or equal to 30 kg/m². This process of list generation was repeated for cannabidiol_dur_c19 patients to form their own list of most-frequent condition_concept_ids.

Non-cannabidiol medications. To identify non-cannabidiol medication records to use in our statistical analysis, we generated a list of drug_concept_ids in the drug_exposure table that were linked (through N3C’s person_id identifier) to more than 5% of all cannabidiol100_ever patients in our Main Analysis Sample (defined below). As described above, only medication records with a start date at least 15 days before their First COVID-19 Test or Dx Date were considered. The result was a list of 202 distinct drug_concept_ids. Examination of the drug_concept_names on this list revealed that some of the drug_concept_ids had the same ingredient composition with a different concentration or route. For example, two of the most common drug_concept_names on the list were the seizure medication clobazam in two different forms: “clobazam 10 MG Oral Tablet” and “clobazam 2.5 MG/ML Oral Suspension”. Another observation was that some drug_concept_names were identical except for the addition of the trade name at the end of the drug_concept_name, such as “ondansetron 2 MG/ML Injectable Solution” and “ondansetron 2 MG/ML Injectable Solution [Zofran]”. We hypothesized that there were likely many medication concepts that did not appear on this list but had either the same exact ingredient composition or a very similar ingredient composition as an entry on this list. Therefore, we decided to group all medications with the same first-listed ingredient, and create a dichotomous indicator variable for each group for use in our statistical analysis. We only looked at the first listed ingredient of multi-ingredient medications, because over 90% of the common medications among the cannabidiol patients were single-ingredient medications and to reduce the number of covariates to be used in our statistical analysis. The medications were grouped using a series of Python regular expressions (48) that are viewable in the N3C data lineage for our work (linked above). In summary, the series of regular expressions removed the concentrations, routes, trade names, and other ingredients from the 202 drug_concept_names, leaving only the first word (e.g., “clobazam”) or continuous sequence of words (e.g., “sodium chloride” by itself was maintained as-is, but “glucose / sodium chloride” was reduced to “glucose”). The resulting 81 first-listed ingredients can be found in Table S4. This entire process was repeated for cannabidiol_dur_c19 patients to form their own list of 92 most-frequent first-listed medication ingredients. After

the regular expressions was applied, we manually combined 1 pair of potentially distinct categories, “ketorolac tromethamine” and “ketorolac”, because these two pain medication ingredient terms appear to be used synonymously in the literature (49) and some of our preliminary statistical analysis gave undefined confidence intervals around the estimated odds ratio association between ketorolac (by itself) and COVID-19 positive test rates, presumably due to small sample size issues.

Prior Emergency Room (ER) encounter, or prior hospital encounter. We hypothesized that patients who had one or more visits to an ER or hospital at least 15 days before their First COVID-19 Test or Dx Date could potentially have higher overall illness acuity or SARS-Cov-2 exposure risk prior to possible actual exposure, potentially putting them at higher risk of developing COVID-19 than patients without prior ER or hospital visits. Similarly, patients who were actively in the ER or hospital during their COVID-19 test could have had different probabilities of being tested and/or of testing positive. We therefore calculated 4 indicator variables: whether a patient had an ER encounter in the N3C data with a start date at least 15 days before their First COVID-19 Test or Dx Date, whether a patient had an ER encounter with a start date on or before their First COVID-19 Test or Dx Date and either no end date or an end date on/after their First COVID-19 Test or Dx Date, and 2 more indicators that used the same logic except for hospital encounters instead of ER encounters. We describe how we use these 4 indicator variables in our statistical analysis subsection below.

Demographics.

- *Age.* Patient dates of birth were not available in the de-identified N3C data used for this project, but years of birth were available. Age in years was approximated by subtracting each patient’s year of birth from 2020. In addition to this continuous approximation, a categorical age variable was also used in our cohort matching procedures (described below). The categories we used were those age categories compared according to their COVID-19 severity risk by the CDC (50) that were also representative of more than 5% of the cannabidiol patients in our main analysis. Older age categories with frequencies below this threshold were combined with the next closest younger age category until all categories had at least 5% of cannabidiol patients. This procedure resulted in the same final age categories for our main analysis of cannabidiol100_ever patients as for our analysis of cannabidiol100_dur_c19 patients.
- *Gender.* All the cannabidiol100_ever patients in our Main Analysis Sample had either male or female gender; no other gender values were present in this group’s data. Thus, we created a female indicator variable for use in our statistical analysis that was set equal to 1 if the person had a female gender_concept_id (ID 8532) in the person N3C table, 0 if the person had a male gender_concept_id (ID 8507), and an arbitrary third value otherwise. According to the OMOP Common Data Model documentation (51), gender_concept_id is meant to capture the biological sex at birth of the person.
- *Race.* We categorized race by grouping race_concept_id values into 4 categories that each comprised more than 5% of all cannabidiol100_ever and cannabidiol100_dur_c19 patients. Category “White” was race_concept_id 8527 (“White”); category “Black/African-American” was race_concept_ids 8516 (“Black or African American”) and 38003598 (“Black”); category “Other Race” was race_concept_ids 44814649 and 45878142 (both “Other”), and race_concept_id 8515 (“Asian”); and category “Race Imprecise” was race_concept_ids 0 (“No matching concept”), 8552 and 44814653 (both “Unknown”), 44814659 (“Multiple race”), and 44814650 and 46237210 (both “No information”). These same categorizations were used for the cannabidiol100_ever and cannabidiol100_dur_c19 analyses.
- *Hispanic/Latino ethnicity.* We constructed a dichotomous indicator variable for Hispanic/Latino ethnicity that was set equal to 1 if the patient had an ethnicity_concept_id of (“Hispanic/Latino ethnicity”) and 0 otherwise. Responses of unknown ethnicity were less than 5% of both the cannabidiol100_ever and cannabidiol100_dur_c19 main analysis patient groups, so these patients were grouped with the larger of the two known categories: Not Hispanic/Latino.

- *Data partner/contributor.* We hypothesized that patients from different data partners/contributors could plausibly have different rates of COVID-19 testing and positive test results that were caused in part by differences in geography and site-specific testing practices and patterns, for example. There were 33 data partners that contributed all of the cannabidiol100_ever patients in our main analysis, the majority of which contributed too few patients to estimate statistical associations between data partner membership and cannabidiol use or COVID-19 positive test rates in multivariable models. Therefore, we aggregated all data partners into a categorical variable we called Data Partners and created a separate category for patients of each data partner whose patients were more than 5% of the cannabidiol patients in our main analysis, and an Other Data Partners category for all other patients. This procedure resulted in 7 categories for the cannabidiol100_ever analysis and 6 categories for the cannabidiol100_dur_c19 analysis (see Table S4). We masked the N3C data partner identifiers with an arbitrary identifier for this manuscript.

We are not aware of a publicly available description of how N3C data contributors/partners collected their submitted information about date of birth, gender, race, ethnicity, and other demographic-related variables. The authors are aware of their own institutional policy, which requires that such data are self-reported by patients by selecting from a pre-defined list of categories in the electronic health record, often during patient registration. In some circumstances at the authors' institution, nurses or other care providers will select demographics for patients in consultation with others when possible; for example, a nurse will occasionally select a patient's demographics based on consultation of available documentation or a patient's family members when the patient is unable to make their own demographic selections directly.

Body Mass Index (BMI). The BMI value used for a given patient was the most recent exact, calculated, or approximate value at least 15 days before their First COVID-19 Test or Dx Date. This 15-day qualification window was imposed to exclude BMI values that may have been observed while the study participant had COVID-19 disease or related symptoms. For 62.1% of the Main Analysis Sample (defined below), an exact BMI value was available in N3C (already calculated) at least 15 days before their First COVID-19 Test or Dx Date in the N3C measurement table. Another 8.7% of the Main Analysis Sample did not have an exact BMI value available, but we were able to fill in approximations by estimating BMI using the most recent weight and height at least 15 days before their First COVID-19 Test or Dx Date. To do this, we modified the "Calculate BMI Template" N3C Knowledge Object template originally written by Steve Johnson. The original knowledge object was only applied to patients age 18 or older, had conservative limitations on the amount of time that could elapse between weight and height measurements (i.e., 7 days), and had minimum and maximum weight and height validation limits that could have excluded some young patients in particular. Given that many patients in our Main Analysis Sample were under 18 years of age, we modified the template to remove the age and date proximity restrictions other than requiring the measurements to be at least 15 days before their First COVID-19 Test or Dx Date, and we also broadened the height and weight validation restrictions, so that a maximal number of measurements could be included in our search results and therefore potentially used to approximate patient BMIs approximately 15 days before their First COVID-19 Test or Dx Date. After these steps, 29.2% of patients in the Main Analysis Sample still did not have populated BMI values for our analysis. We were able to approximate another 9.0% of the Main Analysis Sample by using the most recent BMI-related observation or condition concept 15 days before their First COVID-19 Test or Dx Date. For one example, we approximated a BMI of 21.5 for patients with a BMI concept that had an observation_source_concept_name with the text "19-24". For another example, we approximated a BMI of 30.1 if the patient had the condition concept 433736 "Obesity" 15 days before their First COVID-19 Test or Dx Date. Lastly, we were able to manually fill in less than 20 BMI values that had a clear mismatch in units of measure for weight and/or height, but once the units were converted, gave valid BMI estimates (BMI is calculated as kg/m^2 , but weight was sometimes reported in ounces or pounds, and height was sometimes reported in feet or inches or centimeters.). For the remaining patients in the Main Analysis Sample who still lacked a qualifying BMI value, we labeled the patient with a factor variable level of "No

qualifying BMI”. Given the approximate nature of many of the BMI estimates used in this analysis along with the likely correlation between BMI and COVID-19 severity (52), we did not exclude BMI from our analysis but we also did not attempt to statistically control for continuous BMI values in our matching or regression models. Rather, we aggregated them into common categories (53) and controlled for BMI categorically. We required each category to contain more than 5% of the cannabidiol patients in our Main Analysis Sample. When a category included fewer patients, it was combined with the next closest category until all categories had more than 5% of the cannabidiol patients. The resulting unordered factorial variable categories are shown in Table S4. Also in the Results Sensitivity Analysis subsection, we describe the results of analysis in which we modified our BMI inclusion criteria.

Payer category indicator variables. A note provided by N3C administrators about the payer_plan_period table indicates that a minority of data partners/contributors provide data about the payer of their patients’ associated healthcare encounters. However, for those patients that do have payer data available, we wanted to allow our statistical analyses to consider incorporate this data because it could be an indicator of patient heterogeneity such as differences in socioeconomic status and other determinants of health. We therefore created 5 dichotomous indicator variables for whether the patient had any encounters with a start date at least 15 days before their First COVID-19 Test or Dx Date that had a payer of (1) Medicaid, charity, or Hill-Burton, (2) Medicare, (3) private insurance, (4) self-pay, or (5) any other payer. A given patient could have multiple payers (i.e., multiple of these indicator variables could be equal to 1 for a given patient).

Study participants

Main Analysis Sample. We defined the Main Analysis Sample as all patients (i.e., non-null person_ids) that met all the following eligibility criteria.

- *Had a positive or negative non-antibody COVID-19 test:* had at least one positive or negative non-antibody SARS-CoV-2 or COVID-19 measurement or test result (defined above) in the N3C measurement table with a non-null date as of June 18th, 2021. The earliest potential date of COVID-19 testing was February 1, 2020 and the latest was June 18, 2021, although all real dates were replaced with shifted dates in the de-identified data used for this analysis.
- *No prior COVID-19 vaccine:* Did not have any COVID-19 vaccine dose on or before their First COVID-19 Test or Dx Date (defined above).
- *Had a seizure-type condition documented at least 15 days before COVID-19 testing:* Epidiolex^R is FDA-approved only for the treatment of seizures associated with three disorders: Lennox-Gastaut syndrome (LGS), Dravet syndrome (DS), and tuberous sclerosis complex (TSC) (54). The majority of cannabidiol 100 mg/mL patients (defined above) also had records of seizure-related conditions; however, most of their seizure-related conditions were not named LGS, DS, or TSC and were instead named more generically (e.g., the most common standardized concept name was simply “Seizure”). A minority of patients who had cannabidiol 100 mg/mL records did not have seizure-related records. Other cannabidiol solutions and oils are available over-the-counter with different concentrations and purities and are reportedly used for many conditions unrelated to seizures (55). Therefore, to reduce patient and cannabidiol heterogeneity in our analysis, we excluded patients without seizure-related conditions or medications from our analysis, to reduce the chances of including patients whose medication records were miscoded as cannabidiol 100 mg/mL when in fact they were other forms of cannabidiol. Patients with seizure-related conditions were defined by the “neurologic, seizures: NEURO_SEIZ” Elixhauser AHRQ condition indicator variable described above. Furthermore, we excluded patients who only had seizure records within 14 days before, on, or any number of days after their First COVID-19 Test or Dx Date, because seizures are a potential complication of COVID-19 (56), and therefore we aimed to further reduce patient heterogeneity by eliminating patients who only had seizures because of COVID-19.

Statistical analysis

Descriptive statistics. Categorical frequencies and continuous variable statistics were compared using two-sided Fisher exact tests or t-tests, respectively. Categorical characteristics were included as unordered factor variables in all models. Age was the only continuous/integer-valued measure was included in the models described below (and it was also included as a categorical measure in the matching models). R (<https://www.r-project.org/>) packages loaded by default to the high-memory N3C computational environment were used to generate all statistical results except for some p-values from bivariate analysis that were calculated with the *tabi* command of Stata (57) after permission was obtained to download the N3C data in this manuscript. Unless otherwise stated, the threshold used for “statistical significance” was $p < 0.05$.

Matching cannabidiol patients to controls. We matched the `cannabidiol100_ever` and `cannabidiol100_dur_c19` patients with other patients from the Main Analysis Sample using 1-to-n nearest neighbor propensity score matching, where the number of controls matched to 1 cannabidiol patient (n) was set equal to 1 in our main analysis and also 2 or 3 in sensitivity analysis. The propensity scores used for matching were the estimated probabilities of taking cannabidiol 100 mg/mL that were calculated from a multivariable logistic regression model applied to all patients in the Main Analysis Sample. Two of these matching logit models were fit; one for an outcome of “cannabidiol 100 mg/mL ever” and another for an outcome of “cannabidiol 100 mg/mL during COVID-19” for the `cannabidiol100_ever` and `cannabidiol100_dur_c19` analyses, respectively. Patients who had records of cannabidiol 100 mg/mL use but did not qualify for our definition of having active cannabidiol 100 mg/mL records during their First COVID-19 Test or Dx Date (i.e., patients in the `cannabidiol100_ever` group but not in the `cannabidiol100_dur_c19` group) were allowed as feasible matches in our matching model for the `cannabidiol100_dur_c19` main analysis. We discuss results of sensitivity analysis of this approach in the Results section below. The covariates submitted to the propensity scoring model are described above and listed in Table S4 of the Results section. For the nearest neighbor matching algorithm, the R package `matchit` version 4.1.0 was used (58), by submitting the estimated probabilities for the logit model as the measurement of distance. Age matching was accomplished by including the continuous age variable and the categorical age variable as matching covariates because our initial attempts to match on only continuous or only categorical ages resulted in dissimilarities between the age distributions of cannabidiol and non-cannabidiol patients. In the results section below, if we refer to a covariate distribution as “mismatched” between cannabidiol patients and matched controls, the associated standardized mean difference (SMD) is greater than 0.10 and the associated two-sided Fisher exact test p-value is less than 0.05. Literature suggests that neither SMD or p-values are independently sufficient for detecting mismatched covariates in small samples (59).

Analyses of associations with COVID-19 test results. We compared rates of COVID-19 Positive Status among cannabidiol patients and their matched controls with two-sided Fisher’s exact tests. We also used multivariable logistic regression models with COVID-19 Status as the outcome variable and studied the odds ratios (ORs) and their 95% confidence intervals (CIs) of COVID-19 Positive Status for the cannabidiol use term in the models. Covariates included the same variables as the matching analysis except for the indicators of common conditions among cannabidiol patients and common non-cannabidiol medications, which were used in the matching models but not in the outcome models because they had much smaller sample sizes. We also included covariates for ER or hospital use during their COVID-19 test in our main analysis, which were not included in our main matching models, and checked the effects of this inclusion in our sensitivity analysis (below). The only continuous/integer-valued covariate included was age; all other covariates were unordered and categorical. Model fit was assessed with information criteria (Akaike or AIC, and Bayesian or BIC), area under the receiver operating characteristic curve (AUC), pseudo- R^2 , goodness-of-link tests, and Hosmer-Lemeshow decile tests (60, 61).

Link to full data lineage including measurement construction and outcomes assessment

The following links to the N3C data lineage that produced the patient data analysis below: <https://unite.nih.gov/workspace/data-integration/monocle/graph/ri.monocle.main.quick-graph.6a5e6832-f7d1-4f55-bb04-b2fa54acfd0a>. The link will only allow access for N3C accounts with the required permissions, but N3C publication standards recommend including such a link in all publications that use N3C data, to maximize transparency.

Results

This section presents main results of our N3C patient data analysis and also presents results of analysis of the sensitivity of the main results to potential variations in our methods. To reduce the chances of re-identification of patients and in concordance with N3C requirements for publications, all sample sizes and cell sizes discussed in this section that were between 1 and 19 (inclusive) are masked with the label “<20” or “less than 20”.

Overall Analytic Population and Main Analysis Sample

We defined an Overall Analytic Population of 5,681,382 patients with a non-antibody COVID-19 test according to our selection criteria out of over 6.3 million total patients in the N3C person table as of the June 18, 2021 data release (v34), and 25.9% had a COVID-19 Positive Status according to our definition. The rate is intentionally higher than the approximately 7% reported in more comprehensive sources of U.S. COVID-19 test data (62). The N3C data ingestion phenotype team has been asking N3C data contributors since December 3, 2020 to form their data ingestion batches by first identifying cases of COVID-19 and then demographically matching those cases to controls who have tested negative for COVID-19, at a ratio of 1:2 (cases:controls) (63).

Over 99.7% of the Overall Analytic Population had no COVID-19 vaccination record before their First COVID-19 Test or Dx Date. Only 302,460 (5.3%) of 5,681,382 had no vaccination and a seizure-type condition at least 15 days before their First COVID-19 Test or Dx Date, and 14.0% of the 302,460 had a COVID-19 Positive Status. This group of 302,460 patients constituted our Main Analysis Sample.

Of the Overall Analytic Population, 1,517 (0.03%) patients had at least one cannabidiol 100 mg/mL medication record linked to their N3C unique person_id, and 1,212 (0.4%) of the Main Analysis Sample. The 1,212 patients constitute the cannabidiol100_ever group described above, of which 75 (6.2%) had a COVID-19 Positive Status. Of the 1,212 patients, 531 (44%) had a cannabidiol record that we labeled as “active” during their First COVID-19 Test or Dx Date according to our definition above and 26 (4.9%) of these had a COVID-19 Positive Status. Figure S23 summarizes the derivation of our Main Analysis Sample and the COVID-19 Positive Status rates at each step. In general, we observe that the COVID-19 Positive Status rates are decreasing with each narrowing of our population, which suggests that there are potential substantial differences in COVID-19 infection rates, COVID-19 exposure rates, patient characteristics, testing frequency, and/or other factors across these subgroups.

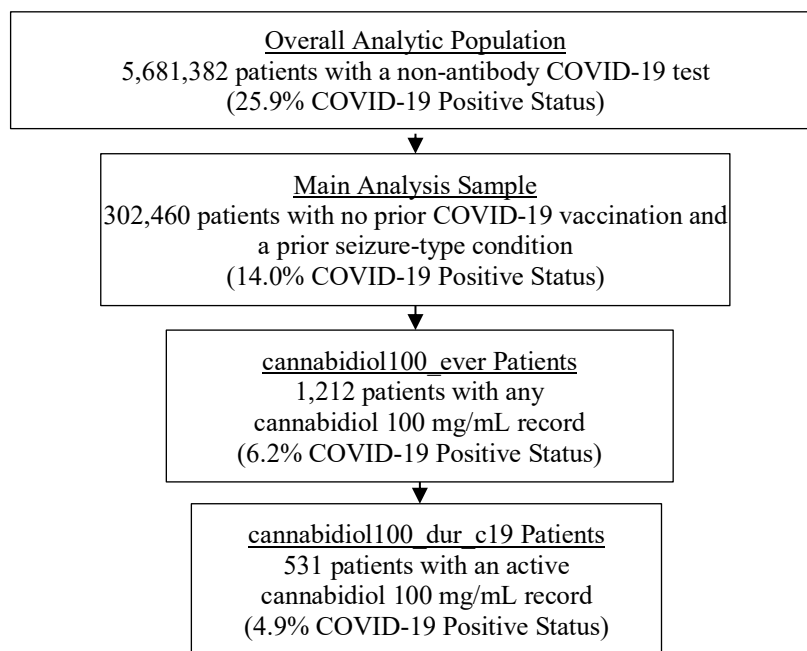


Figure S23. Derivation of our Main Analysis Sample and cannabidiol patient groups.

Main Analysis Results

Characteristics of the Main Analysis Sample and the matched subgroups. Table S4 compares the Main Analysis Sample and our subgroups of focus: the 1,212 cannabidiol100_ever patients and their 1,212 matched controls and the 531 cannabidiol100_dur_c19 patients and their 531 matched controls.

[Table S4 goes here- see below]

For the Main Analysis Sample, we see that cannabidiol 100 mg/mL patients had many covariate frequency differences from the rest of the Main Analysis Sample, even though they all had a seizure-related condition at least 15 days before their First COVID-19 Test or Dx Date. For example, cannabidiol 100 mg/mL patients were more likely to be younger, male, White, and from particular N3C data partners, and to have lower BMI, more prior ER and hospital use, Medicaid or similar insurance, developmental disabilities, neurologic movement disorders, and most of the conditions and medications that were (by definition) common in cannabidiol 100 mg/mL patients. Cannabidiol 100 mg/mL patients were less likely to have conditions of depression, diabetes, HIV or AIDS, hypertension, liver disease, other neurologic disorders, pregnancy, psychoses, and substance use disorders. The predicted cannabidiol group membership from our propensity score models had high correlation with actual cannabidiol group membership, with AUC=0.983 for the cannabidiol100_ever propensity score model and AUC=0.988 for the cannabidiol100_dur_c19 model. These high AUCs suggest that the cannabidiol 100 mg/mL patients were well-distinguished from the majority of the Main Analysis Sample by the covariates we chose to include in our propensity score models, and the multitude of dissimilar covariates provides support for our use of matching in our subsequent analysis.

Within the matched subgroups, we see that cannabidiol 100 mg/mL patients had similar covariate frequencies as their matched controls. The only mismatched covariates in the cannabidiol100_ever analysis was condition concept 4046213: Lennox-Gastaut Syndrome (LGS). Epidiolex^R is FDA-approved to treat seizures from LGS, so this mismatch implies that the matching model had difficulty finding enough “on-label candidates” for

cannabidiol 100 mg/mL who had no records of cannabidiol 100 mg/mL on any date in their N3C data timeline. The only mismatched covariate in the cannabidiol100_dur_c19 analysis was condition concept 40484102: abnormal finding on evaluation procedure, a concept that is Related in ATLAS to many imaging standard concepts and other testing concepts. There were other covariates that had either $p < 0.05$ or $SMD > 0.10$ (but not both) in the cannabidiol100_ever and cannabidiol100_dur_c19 analyses, which we studied in our sensitivity analysis described below. COVID-19 Positive Status was significantly lower for cannabidiol 100 mg/mL patients compared to their matched controls: 6.2% versus 8.9% ($p = 0.014$) for the cannabidiol100_ever analysis, and 4.9% versus 9.0% ($p = 0.011$) for the cannabidiol100_dur_c19 analysis. This negative bivariate association between cannabidiol 100 mg/mL medication records and COVID-19 Positive Status became more statistically significant when we analyzed the association in multivariable regression models. Table S5 shows odds ratios for COVID-19 Positive Status of 0.65 ($p = 0.009$, 95% C.I. [0.47,0.90]) for cannabidiol100_ever patients and 0.48 ($p = 0.006$, 95% C.I. [0.29,0.81]) for cannabidiol100_dur_c19 patients. The multivariable models passed our goodness-of-fit tests (Table S5, bottom).

Table S5. Multivariable regression models for COVID-19 Positive Status

	Any cannabidiol 100 mg/mL model (N=2424)	Active cannabidiol 100 mg/mL model (N=1062)
	OR (p; 95% CI)	OR (p; 95% CI)
Had cannabidiol 100 mg/mL? (reference = 0, i.e. "no")	0.65 (0.009; 0.47-0.90)	0.48 (0.006; 0.29-0.81)
Age in years	1.02 (0.014; 1.00-1.03)	1.01 (0.292; 0.99-1.04)
Female (reference = Male)	0.79 (0.155; 0.57-1.09)	1.01 (0.977; 0.60-1.69)
Race (reference = White)		
Black/African American	1.09 (0.752; 0.63-1.89)	1.18 (0.720; 0.48-2.86)
Other	2.21 (0.011; 1.20-4.07)	2.66 (0.042; 1.03-6.84)
Imprecise	2.03 (0.004; 1.25-3.30)	1.88 (0.146; 0.80-4.41)
Hispanic or Latino (reference = Not Hisp./Lat. or unknown)	1.47 (0.091; 0.94-2.30)	1.47 (0.310; 0.70-3.11)
BMI category (reference = [18.5,25))		
Less than 16.5	0.79 (0.370; 0.48-1.32)	0.85 (0.655; 0.40-1.77)
[16.5,18.5)	1.10 (0.727; 0.65-1.85)	0.46 (0.136; 0.16-1.28)
[25,30)	1.23 (0.438; 0.73-2.05)	0.80 (0.587; 0.36-1.79)
Greater than or equal to 30	1.52 (0.152; 0.86-2.68)	1.61 (0.280; 0.68-3.79)
No qualifying BMI	2.09 (0.007; 1.23-3.55)	1.28 (0.632; 0.47-3.52)
ER, Hospital Utilization (references = No such encounter)		
Prior ER encounter	1.23 (0.256; 0.86-1.76)	1.07 (0.815; 0.61-1.88)
Prior hospital encounter	0.70 (0.068; 0.48-1.03)	1.01 (0.985; 0.53-1.89)
ER encounter during First COVID-19 Test or Dx Date	1.17 (0.468; 0.77-1.77)	0.84 (0.616; 0.44-1.63)
Hospital enc. during First COVID-19 Test or Dx Date	0.79 (0.229; 0.54-1.16)	0.94 (0.834; 0.53-1.67)
Payer (references = No such payer)		
Medicaid, charity, or Hill-Burton	0.91 (0.713; 0.55-1.50)	0.71 (0.406; 0.32-1.58)
Medicare	0.54 (0.222; 0.20-1.46)	0.72 (0.696; 0.14-3.70)
Private insurance	0.92 (0.770; 0.52-1.62)	0.75 (0.571; 0.28-2.03)
Any other payer	1.29 (0.645; 0.44-3.75)	2.75 (0.216; 0.55-13.66)
Self-pay	0.62 (0.296; 0.25-1.53)	1.42 (0.598; 0.39-5.16)
Condition indicators (references = No such condition)		
Arthritis	1.10 (0.702; 0.68-1.79)	1.40 (0.382; 0.66-3.01)
Cancer	0.84 (0.714; 0.32-2.18)	0.95 (0.951; 0.21-4.41)
Cerebrovascular disease or stroke	1.75 (0.041; 1.02-2.99)	1.20 (0.757; 0.38-3.72)
Chronic kidney disease	1.26 (0.562; 0.57-2.78)	0.58 (0.514; 0.11-2.95)
Chronic lung disease	1.21 (0.379; 0.79-1.84)	0.80 (0.499; 0.41-1.54)
Dementia	0.71 (0.294; 0.37-1.35)	0.85 (0.737; 0.32-2.22)
Depression	0.30 (0.005; 0.13-0.70)	0.20 (0.029; 0.05-0.84)
Developmental disabilities	1.25 (0.353; 0.78-1.98)	1.38 (0.435; 0.62-3.08)
Diabetes	0.85 (0.572; 0.48-1.49)	1.18 (0.700; 0.51-2.76)
Smoker, former or current	1.42 (0.169; 0.86-2.33)	1.55 (0.293; 0.69-3.49)
HIV or AIDS	0.81 (0.527; 0.41-1.57)	1.33 (0.569; 0.50-3.54)
Hypertension	0.89 (0.713; 0.49-1.63)	0.68 (0.481; 0.24-1.97)
Liver disease	1.13 (0.771; 0.50-2.54)	0.64 (0.586; 0.12-3.24)
Neurologic, movement disorders	1.16 (0.422; 0.81-1.68)	0.94 (0.827; 0.52-1.68)
Neurologic, other disorders	0.59 (0.008; 0.40-0.87)	0.51 (0.037; 0.27-0.96)
Pregnancy	2.60 (0.004; 1.35-5.01)	2.05 (0.208; 0.67-6.26)
Psychoses	1.29 (0.463; 0.66-2.52)	2.03 (0.128; 0.82-5.06)

Pulmonary circulatory disorders	1.92 (0.216; 0.68-5.42)	3.08 (0.151; 0.66-14.27)
Sickle cell disease	1.53 (0.708; 0.17-13.96)	Excluded due to small N
Substance use disorder	0.91 (0.822; 0.40-2.07)	0.79 (0.744; 0.19-3.25)
Transplant recipient status	1.09 (0.903; 0.29-4.12)	2.10 (0.536; 0.20-22.10)
Data partners (reference = other data partners)		
1	0.50 (0.063; 0.24-1.04)	0.51 (0.152; 0.21-1.28)
2	0.77 (0.553; 0.33-1.82)	0.09 (0.035; 0.01-0.85)
3	0.25 (0.010; 0.09-0.72)	0.39 (0.140; 0.11-1.36)
4	1.49 (0.319; 0.68-3.27)	Excluded due to small N
5	1.66 (0.099; 0.91-3.03)	1.75 (0.217; 0.72-4.24)
6	0.30 (0.060; 0.09-1.05)	0.31 (0.195; 0.05-1.83)
Model intercept term	0.06 (0.000; 0.03-0.13)	0.08 (0.000; 0.02-0.27)
AUC; Pseudo R ²	0.734; 0.094	0.755; 0.112
Goodness-of-link test p; Hosmer-Lemeshow decile test p	0.311; 0.743	0.785; 0.904

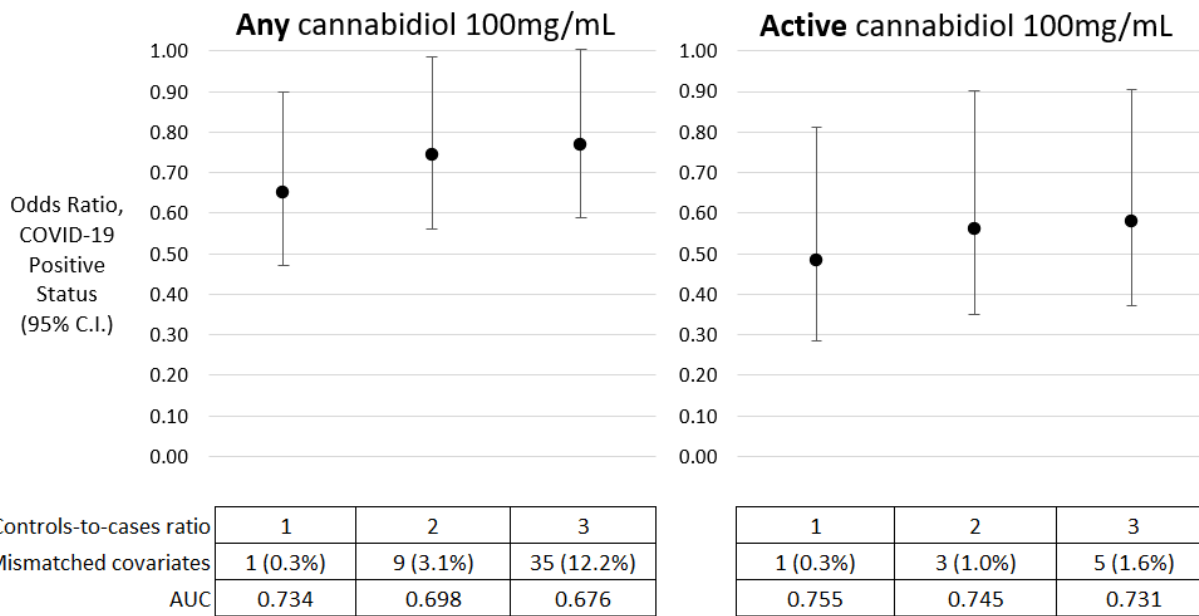
Table S5 footnotes. For the “Any cannabidiol 100 mg/mL model” column, an answer of “yes” to “Had cannabidiol 100 mg/mL?” is equivalent to cannabidiol100_ever=1. For the “Active cannabidiol 100 mg/mL model” column, an answer of “yes” to “Had cannabidiol 100 mg/mL?” is equivalent to cannabidiol100_dur_c19=1. Trisomy 21 / Down syndrome was not included in either regression model because there were very few patients (<20) with this condition and none of them had a COVID-19 Positive Status, so the R package could not calculate a valid OR estimate.

Sensitivity Analyses

This subsection describes the results of adjustments to our main analysis methods and other results that help establish the robustness of our methods to variations in our approach. Each sensitivity analysis is described below next to an italicized summary statement.

Adding potentially mismatched covariates to the multivariable models for COVID-19 Positive Status. To identify any substantial changes in the odds ratios from Table S5 after adding covariates which were on the verge of being mismatched (see the definition of “mismatched” in our Methods section), we sequentially added common condition/medication covariates that had $p < 0.05$ and/or $SMD > 0.10$ in Table S4. We observed no substantial change in the cannabidiol 100 mg/mL odds ratios or their statistical significance by adding these covariates to the model; all of the ORs stayed approximately the same and their p-values stayed below 0.011 for the cannabidiol100_ever model and below 0.008 for the cannabidiol100_dur_c19 model. Also, the AIC and BIC model fit statistics did not change substantially from adding these covariates.

Matching to multiple controls per cannabidiol 100 mg/mL patient. By design, 1-to-1 nearest-neighbor matching is heuristically more likely to statistically balance covariates than 1-to-n ($n > 1$) matching and thereby yield matched controls who are more similar to the cannabidiol patients than control groups formed by values of n above 1. However, two considerations prompt us to repeat our analysis with n above 1: (1) important other data/covariates may be missing from the matching models such that “close-but-not-matched” patients from our Main Analysis Sample are actually more similar to our cannabidiol patients and should thus be in our comparator groups, and (2) our Main Analysis Sample sizes could benefit from being increased, to assess the robustness of our findings to broadening of the control groups in terms of size and heterogeneity. Figure S2 presents the results of doubling or tripling the number of controls matched to the cannabidiol100_ever and cannabidiol100_dur_c19 groups (Figures S24.a and S24.b, respectively).



Figs S24.a (left), S24.b (right). Association between cannabidiol 100 mg/mL records and COVID-19 Positive Status.

As the number of matched controls increases, the point estimate of the magnitude of the negative association between cannabidiol use and COVID-19 positive test results is attenuated but is still statistically significant in all but the $n=3$ model for cannabidiol100_ever patients which is at the margin of significance. We also see that the match quality deteriorates as n increases, albeit at most 2% of the covariates become mismatched for the cannabidiol100_dur_c19 analysis. Our definitions of statistical significance and “mismatched” are given in the Methods section above.

Cannabidiol 100 mg/mL and COVID-19 associations among adults within our matched cohorts. The majority of the cannabidiol 100 mg/mL patients and matched controls were persons under age 18. In this sensitivity analysis, we studied whether significant negative bivariate associations existed between cannabidiol use and COVID-19 test results among adults, because older age is frequently reported elsewhere as a risk factor for COVID-19. Table S6 shows a statistically significant negative bivariate association for the subsample of our cannabidiol100_ever and cannabidiol100_dur_c19 analyses that were of age 18 or older. These findings support future research of COVID-19 risk and cannabidiol use among adults.

Table S6. Bivariate associations between cannabidiol 100 mg/mL records and COVID-19 Status among persons age 18 or older within the matched cohorts from Table S4.

COVID-19 Status	Adults in cannabidiol100_ever group	Adults in cannabidiol100_ever matched controls	
Negative	454	418	
Positive	33	50	$p = 0.038$
COVID-19 Status	Adults in cannabidiol100_dur_c19 group	Adults in cannabidiol100_dur_c19 matched controls	
Negative	>190	181	
Positive	<20	26	$p < 0.005$

Other definitions of active cannabidiol 100 mg/mL use during COVID-19 tests. As described in the Methods section, some of the 531 matched controls in Table S4 may have been actively taking cannabidiol 100 mg/mL during their First COVID-19 Test or Dx Date, but we did not assume as much because of, for example, missing medication end dates. For example, 100 of the 531 matched controls had a cannabidiol 100 mg/mL record with a start date strictly before their First COVID-19 Test or Dx Date but no end date. These 100 patients had very

similar COVID-19 positivity rates to the 531 patients that we defined as having active cannabidiol 100 mg/mL ($p > 0.70$), but the precise number of positive cases is too small to report. This supports the possibility that the significant negative association we report between cannabidiol 100 mg/mL use and COVID-19 Status could be even stronger if we had more precise data about active cannabidiol 100 mg/mL medication use during COVID-19 testing.

Excluding cannabidiol100_ever patients from the feasible control group in the cannabidiol100_dur_c19 analysis. In our main analysis of patients that were determined to have active cannabidiol 100 mg/mL records, 116 of the matched controls to the 531-patient group had a record of cannabidiol 100 mg/mL use that was not determined to be active on their First COVID-19 Test or Dx Date. Such patients could very well have been actively using cannabidiol 100 mg/mL but the medication documentation did not clearly indicate it. To again analyze the sensitivity of our results to our assumptions about who was and was not actively taking cannabidiol 100 mg/mL, we restricted the pool of feasible controls for the cannabidiol100_dur_c19 matching analysis to only include patients who *never* had a record of cannabidiol 100 mg/mL. After doing so, we found a similar odds ratio to that of the cannabidiol 100 mg/mL term in the multivariable regression model (shown in Table S5, rightmost set of columns), and it had similar statistical significance ($p=0.010$). However, the matching output reduced in quality, with 2 more characteristics having standardized mean differences (SMD) above 0.10, and a higher average SMD in general, so we report the results in the rightmost columns of Table S5 as our main findings for the patients on active cannabidiol 100 mg/mL.

Comparing COVID-19 positive test rates among patients with active cannabidiol 100 mg/mL records during their COVID-19 test and patients with prior-but-not-active cannabidiol 100 mg/mL records. To further analyze the sensitivity of our results to our assumptions about who was and was not actively taking cannabidiol 100 mg/mL, we compared the COVID-19 positive test rate between the 531 patients in our cannabidiol100_dur_c19 sample and 435 patients who only had cannabidiol 100 mg/mL records before their First COVID-19 Test or Dx Date; that is, these 435 patients only had cannabidiol 100 mg/mL records with start and end dates before their First COVID-19 Test or Dx Date. The COVID-19 positive test rate among these 435 patients was 7.6% (33 of 435) versus 4.9% (26 of 531) for the cannabidiol100_dur_c19 group, with a two-sided Fisher's exact test p-value of 0.104. In summary, the patients who were previously indicated for cannabidiol 100 mg/mL use but no longer appeared to be taking it had more frequent COVID-19 positive tests but the difference was statistically insignificant.

Comparing COVID-19 positive test rates among patients with records of cannabidiol 100 mg/mL and records of cannabidiol of unknown concentration. Within the Main Analysis Sample of 302,460 patients, 180 (0.1%) of the patients had a record of "cannabidiol" without any concentration given and no records of cannabidiol 100 mg/mL. Less than 20 of these patients had a COVID-19 Positive Status, which was a statistically comparable rate ($p>0.50$) to the positive rate of patients who ever had a cannabidiol 100 mg/mL record (75 of 1,212, or 6.2%). However, as described above, we do not assume that these 180 patients ever took a cannabidiol 100 mg/mL oral solution, because there are many other cannabidiol products available in the market that the patients could have been taking. Therefore, we do not group these 180 patients with our cannabidiol 100 mg/mL groups for any further analysis.

Comparing COVID-19 positive test rates among patients with records of cannabidiol 100 mg/mL and records of cannabidiol 100 mg/mL [Epidiolex^R]. Within the 1,212 patients who ever had a cannabidiol 100 mg/mL record, 773 (63.8%) also had the word "Epidiolex" in a cannabidiol record name, of which 46 (6.0%) had a COVID-19 Positive Status, which was a statistically comparable rate ($p=0.71$) to the 29 positive patients (6.6%) of 439 patients who ever had a cannabidiol 100 mg/mL record but did not have any record that included the word "Epidiolex". Of the 531 cannabidiol100_dur_c19 patients, 342 (64.4%) also had the word "Epidiolex" in a cannabidiol record name. Although the Ns are too small to report, the rate of COVID-19 Positive Status for these

342 of 531 patients was comparable to the 189 of 531 patients who did not have the word “Epidiolex” in a cannabidiol record name (p between 0.2 and 0.4).

Aggregations of condition concepts common among cannabidiol patients. We also used matching models and multivariable models of COVID-19 positivity with covariates that were aggregates of the condition concepts that were common among cannabidiol patients, instead of the individual common condition concept indicator covariates shown in Table S4. For example, we attempted grouping all condition concepts that appeared related to epilepsy based on their condition concept name. We observed similar findings of statistically significant negative associations between likely cannabidiol use and COVID-19 positive test rates. Since these methods required frequent and relatively-complex assumptions about which conditions should be grouped and which should be separated, we report the results of the more transparent and repeatable approach as our main results above.

Adding matching covariates of ER/hospitalization during First COVID-19 Test or Dx Date. In our main statistical analysis, we did not match on whether a patient was in the ER or in the hospital during their First COVID-19 Test or Dx Date, because we hypothesized that doing so might bias our selection towards patients with COVID-19. Instead, we only included these 2 indicator variables as covariates in our multivariable regressions with an outcome of COVID-19 test results. In this sensitivity analysis, we added these two indicators to our matching (so that all 4 ER and hospitalization indicators were now included in the matching models). After matching in this method, the p-value on the cannabidiol100_ever indicator in Table S4 changed to 0.030, and the p-value on the cannabidiol100_dur_c19 indicator changed to 0.020, both still below 0.05.

Populating values for missing BMI values. In this sensitivity analysis, we attempted to populate some of the “no qualifying BMI” values from our main analysis by allowing the inclusion of the most recent exact, calculated, or approximated BMI values from any date, not just those measured at least 15 days before the patient’s First COVID-19 Test or Dx Date. Only those patients with “no qualifying BMI” per our original procedure (described above) had missing BMIs populated in this way. It resulted in the addition of 35 populated BMI values for the cannabidiol100_ever patients in our main analysis and less than 20 of the cannabidiol100_dur_c19 patients. Notably, the p-value on the cannabidiol100_ever indicator in Table S4 changed to 0.076 and the p-value on the cannabidiol100_dur_c19 indicator changed to 0.782. We see this sensitivity analysis result as potentially misleading, however, because these patients’ BMIs could have been measured only because the patient had COVID-19, or the patients’ BMIs could have reduced in value after contracting COVID-19 as a result of the disease. Furthermore, if we include BMIs from other dates, one could argue that we should also include conditions and medications for other dates, not just 15 days or more before their First COVID-19 Test or Dx Date. Even so, as a result of this sensitivity analysis, we recommend that additional analyses be conducted among only patients with exact BMI data. Such analyses would eliminate 38% of our main analytic cohort, so we defer this work to future research.

Discussion

We find that COVID-19 positive rates were statistically significantly lower among N3C patients with a history of seizures who had likely taken cannabidiol 100 mg/mL than among other seizure-history patients who had not likely taken cannabidiol 100 mg/mL but were matched to the cannabidiol 100 mg/mL patients so that they had similar demographics, medical conditions documented before COVID-19 testing, medication records before COVID-19 testing, and other factors such as the site that contributed their data and their emergency room and hospital utilization. We especially observed this negative association among patients who were likely actively taking cannabidiol 100 mg/mL during their COVID-19 test, in that the association was slightly stronger than the association among patients ever taking cannabidiol 100 mg/mL with odds ratio confidence interval upper bounds

and point estimates that were further below 1, and the matched patients were slightly more similar. Figure S2.b summarizes our main findings that patients potentially taking cannabidiol 100 mg/mL on the date of their first known COVID-19 test or diagnosis were estimated to have between 0.29 and 0.91 odds of COVID-19 positivity compared to matched patients who we assumed were not taking cannabidiol 100 mg/mL on the date of their first known COVID-19 test or diagnosis, with a point estimate of 0.48 odds compared to the best matched group of controls.

We analyzed only patients with a prior seizure-related condition, to reduce heterogeneity in our Main Analysis Sample and to increase the likelihood that the analyzed patients with cannabidiol 100 mg/mL records were using cannabidiol 100 mg/mL and not another cannabidiol substance. However, future work should analyze cannabidiol use in other populations, under the hypothesis that COVID-19 risk or related disease risk may be associated with cannabidiol use in broader populations.

This work has limitations. Residual confounding from unmeasured covariates remains a possibility in this observational analysis that revealed a significant negative association between COVID-19 positive test rates and proximal cannabidiol medication records among a statistically matched cohort of patients. Future observational studies should seek to ensure that their cannabidiol medication data is comprehensive enough to allow for accurate accounting of the dates that a study participant takes cannabidiol and the corresponding medication trade name (if any), concentration, dose, route, quantity, and daily frequency. Also, we suggest that further analyses of large multi-site data samples utilize data that has non-missing records for all covariates (e.g., race, ethnicity, and BMI) such as other (non-N3C) electronic health record datasets. Furthermore, we suggest that researchers consider the design, safety, and efficacy of potential prospective cohort studies or randomized controlled trials that would compare rates of COVID-19 or similar disease onset between patients who are taking cannabidiol and control cohorts of non-cannabidiol patients. Such studies could provide stronger evidence for or against the potential that cannabidiol use can protect against developing COVID-19 or similar disease.

Table S4. Characteristics of Main Analysis Sample and Matched Subgroups.

	Main Analysis Sample	"Any cannabidiol 100 mg/mL" analysis**				"Active cannabidiol 100 mg/mL" analysis**			
		cannabidiol 100 ever=1	Matched controls	p*	SMD*	cannabidiol100 dur c19=1	Matched controls	p*	SMD*
Number of persons	302460 (100)	1212 (100)	1212 (100)			531 (100.0)	531 (100.0)		
COVID-19 Positive Status	42332 (14.0)	75 (6.2)	108 (8.9)	0.014	0.103	26 (4.9)	48 (9.0)	0.011	0.163
Age in years (mean, standard deviation)	35.9, 17.9	18.2, 14.7	17.4, 14.3	0.176	0.055	17.4, 13.1	17.0, 12.3	0.610	0.031
Age category				0.698	0.060			0.473	0.116
Less than 5	15840 (5.2)	181 (14.9)	185 (15.3)			72 (13.6)	64 (12.1)		
[5,18)	15318 (5.1)	544 (44.9)	559 (46.1)			256 (48.2)	260 (49.0)		
[18,30)	75325 (24.9)	265 (21.9)	269 (22.2)			119 (22.4)	124 (23.4)		
[30,40)	93740 (31.0)	106 (8.7)	103 (8.5)			48 (9.0)	58 (10.9)		
Greater than or equal to 40	92995 (30.7)	116 (9.6)	96 (7.9)			36 (6.8)	25 (4.7)		
Unknown	9242 (3.1)	0 (0.0)	0 (0.0)			0 (0.0)	0 (0.0)		
Gender				1.000	<0.001			1.000	0.004
Male	64630 (21.4)	671 (55.4)	671 (55.4)			310 (58.4)	311 (58.6)		
Female	230198 (76.1)	541 (44.6)	541 (44.6)			221 (41.6)	220 (41.4)		
Other or unknown	7632 (2.5)	0 (0.0)	0 (0.0)			0 (0.0)	0 (0.0)		
Race				0.779	0.042			0.432	0.102
White	185000 (61.2)	884 (72.9)	874 (72.1)			379 (71.4)	393 (74.0)		
Black/African-American	58063 (19.2)	139 (11.5)	144 (11.9)			63 (11.9)	68 (12.8)		
Other ⁸	12505 (4.1)	69 (5.7)	62 (5.1)			33 (6.2)	27 (5.1)		
Imprecise	46892 (15.5)	120 (9.9)	132 (10.9)			56 (10.5)	43 (8.1)		
Hispanic or Latino	44440 (14.7)	167 (13.8)	184 (15.2)	0.356	0.040	73 (13.7)	70 (13.2)	0.857	0.017
Data partners				0.123	0.129			0.773	0.098
1	7930 (2.6)	135 (11.1)	114 (9.4)			87 (16.4)	83 (15.6)		
2	8608 (2.8)	101 (8.3)	71 (5.9)			43 (8.1)	47 (8.9)		
3	1673 (0.6)	81 (6.7)	81 (6.7)			43 (8.1)	46 (8.7)		
4	6438 (2.1)	74 (6.1)	70 (5.8)						
5	4937 (1.6)	74 (6.1)	69 (5.7)			42 (7.9)	33 (6.2)		
6	13603 (4.5)	73 (6.0)	71 (5.9)			35 (6.6)	28 (5.3)		
Other data partners	259271 (85.7)	674 (55.6)	736 (60.7)			281 (52.9)	294 (55.4)		
Prior ER encounter	155161 (51.3)	729 (60.1)	690 (56.9)	0.117	0.065	326 (61.4)	307 (57.8)	0.260	0.073
Prior hospital encounter	156889 (51.9)	842 (69.5)	850 (70.1)	0.757	0.014	392 (73.8)	389 (73.3)	0.889	0.013
BMI category				0.281	0.102			0.926	0.072
Less than 16.5	10870 (3.6)	274 (22.6)	248 (20.5)			121 (22.8)	128 (24.1)		
[16.5,18.5)	9073 (3.0)	178 (14.7)	163 (13.4)			71 (13.4)	69 (13.0)		
[18.5,25)	54083 (17.9)	375 (30.9)	411 (33.9)			186 (35.0)	174 (32.8)		
[25,30)	59200 (19.6)	168 (13.9)	151 (12.5)			74 (13.9)	80 (15.1)		
Greater than or equal to 30	108244 (35.8)	122 (10.1)	125 (10.3)			48 (9.0)	53 (10.0)		
No qualifying BMI	60990 (20.2)	95 (7.8)	114 (9.4)			31 (5.8)	27 (5.1)		
Payer									
Medicaid, charity, or Hill-Burton	39841 (13.2)	352 (29.0)	302 (24.9)	0.025	0.093	133 (25.0)	140 (26.4)	0.674	0.030
Medicare	13561 (4.5)	63 (5.2)	62 (5.1)	1.000	0.004	23 (4.3)	<20	0.640	0.040
Private insurance	50703 (16.8)	173 (14.3)	169 (13.9)	0.861	0.010	67 (12.6)	62 (11.7)	0.707	0.029
Any other payer	9763 (3.2)	58 (4.8)	49 (4.0)	0.429	0.036	25 (4.7)	30 (5.6)	0.580	0.043
Self-pay	19721 (6.5)	76 (6.3)	83 (6.8)	0.623	0.023	34 (6.4)	35 (6.6)	1.000	0.008
Condition Indicators [‡]									
Conditions associated w/ high COVID-19 risk and/or severity according to the CDC									

	Main Analysis Sample	"Any cannabidiol 100 mg/mL" analysis**				"Active cannabidiol 100 mg/mL" analysis**			
		cannabidiol 100 ever=1	Matched controls	p*	SMD*	cannabidiol100 dur c19=1	Matched controls	p*	SMD*
Arthritis	51134 (16.9)	196 (16.2)	194 (16.0)	0.956	0.005	84 (15.8)	84 (15.8)	1.000	<0.001
Cancer	18026 (6.0)	51 (4.2)	40 (3.3)	0.285	0.048	<25	<25	1.000	0.010
Cerebrovascular disease or stroke	28319 (9.4)	95 (7.8)	108 (8.9)	0.379	0.039	32 (6.0)	31 (5.8)	1.000	0.008
Chronic kidney disease	17784 (5.9)	54 (4.5)	40 (3.3)	0.171	0.060	22 (4.1)	<20	0.630	0.040
Chronic lung disease	65653 (21.7)	318 (26.2)	299 (24.7)	0.401	0.036	159 (29.9)	150 (28.2)	0.589	0.037
Dementia	27680 (9.2)	134 (11.1)	132 (10.9)	0.948	0.005	58 (10.9)	60 (11.3)	0.922	0.012
Depression	78916 (26.1)	130 (10.7)	142 (11.7)	0.479	0.031	40 (7.5)	36 (6.8)	0.721	0.029
Developmental disabilities	69608 (23.0)	988 (81.5)	1013 (83.6)	0.199	0.054	446 (84.0)	448 (84.4)	0.933	0.010
Diabetes	199567 (66.0)	169 (13.9)	147 (12.1)	0.205	0.054	77 (14.5)	74 (13.9)	0.861	0.016
Smoker, former or current	46755 (15.5)	179 (14.8)	222 (18.3)	0.022	0.096	79 (14.9)	92 (17.3)	0.316	0.067
HIV or AIDS	175504 (58.0)	115 (9.5)	102 (8.4)	0.393	0.038	50 (9.4)	47 (8.9)	0.831	0.020
Hypertension	222167 (73.5)	166 (13.7)	160 (13.2)	0.766	0.015	66 (12.4)	66 (12.4)	1.000	<0.001
Liver disease	23071 (7.6)	42 (3.5)	51 (4.2)	0.398	0.039	<20	<20	0.260	0.080
Neurologic, movement disorders	19786 (6.5)	482 (39.8)	473 (39.0)	0.740	0.015	235 (44.3)	231 (43.5)	0.853	0.015
Neurologic, other disorders	195294 (64.6)	530 (43.7)	505 (41.7)	0.324	0.042	222 (41.8)	224 (42.2)	0.950	0.008
Pregnancy	165971 (54.9)	91 (7.5)	76 (6.3)	0.262	0.049	36 (6.8)	37 (7.0)	1.000	0.007
Psychoses	93852 (31.0)	199 (16.4)	205 (16.9)	0.785	0.013	75 (14.1)	71 (13.4)	0.789	0.022
Pulmonary circulatory disorders	7132 (2.4)	27 (2.2)	21 (1.7)	0.466	0.036	<20	<20	0.550	0.050
Sickle cell disease	3619 (1.2)	<20	<20	0.266	0.057	<20	0 (0.0)	0.480	0.090
Substance use disorder	186985 (61.8)	80 (6.6)	75 (6.2)	0.740	0.017	29 (5.5)	26 (4.9)	0.782	0.026
Transplant recipient status	5228 (1.7)	<20	<20	0.554	0.032	<20	<20	1.000	<0.001
Trisomy 21 / Down syndrome	280 (<0.1)	<20	<20	0.015	0.108	<20	<20	1.000	<0.001
Conditions common among cannabidiol patients									
377091: Seizure	81671 (27.0)	879 (72.5)	866 (71.5)	0.587	0.024	397 (74.8)	393 (74.0)	0.833	0.017
380378: Epilepsy	59329 (19.6)	824 (68.0)	856 (70.6)	0.172	0.057	361 (68.0)	374 (70.4)	0.425	0.053
40483317: Refractory epilepsy	7984 (2.6)	763 (63.0)	755 (62.3)	0.769	0.014	372 (70.1)	373 (70.2)	1.000	0.004
40486120: Delay in physiological development	8733 (2.9)	514 (42.4)	505 (41.7)	0.742	0.015	235 (44.3)	220 (41.4)	0.385	0.057
4046213: Lennox-Gastaut syndrome	1335 (0.4)	507 (41.8)	404 (33.3)	<0.001	0.176	275 (51.8)	261 (49.2)	0.425	0.053
4309257: Refractory localization-related epilepsy	8083 (2.7)	486 (40.1)	508 (41.9)	0.386	0.037	218 (41.1)	234 (44.1)	0.352	0.061
4101747: Localization-related(focal)(partial)idiopathic epilepsy and epileptic syndromes with seizures of localized onset	20571 (6.8)	432 (35.6)	456 (37.6)	0.332	0.041	185 (34.8)	200 (37.7)	0.372	0.059
75860: Constipation	40057 (13.2)	427 (35.2)	408 (33.7)	0.442	0.033	202 (38.0)	189 (35.6)	0.445	0.051
4102986: Disorder of psychological development	4974 (1.6)	404 (33.3)	389 (32.1)	0.545	0.026	181 (34.1)	164 (30.9)	0.295	0.068
4236312: Complex partial epileptic seizure	6376 (2.1)	398 (32.8)	414 (34.2)	0.519	0.028	180 (33.9)	187 (35.2)	0.699	0.028
4055361: Generalized epilepsy	8397 (2.8)	392 (32.3)	384 (31.7)	0.761	0.014	176 (33.1)	187 (35.2)	0.518	0.044
4201388: Gastrostomy present	6943 (2.3)	384 (31.7)	378 (31.2)	0.827	0.011	185 (34.8)	172 (32.4)	0.436	0.052
372887: Disorder of brain	21503 (7.1)	367 (30.3)	363 (30.0)	0.894	0.007	159 (29.9)	154 (29.0)	0.788	0.021
4144111: Gastroesophageal reflux disease without esophagitis	60555 (20.0)	364 (30.0)	340 (28.1)	0.303	0.044	155 (29.2)	142 (26.7)	0.412	0.055
4134120: Cerebral palsy	5465 (1.8)	339 (28.0)	337 (27.8)	0.964	0.004	160 (30.1)	157 (29.6)	0.893	0.012
437663: Fever	35267 (11.7)	328 (27.1)	305 (25.2)	0.309	0.043	153 (28.8)	143 (26.9)	0.538	0.042
438485: Postoperative state	40916 (13.5)	319 (26.3)	318 (26.2)	1.000	0.002	150 (28.2)	138 (26.0)	0.448	0.051
254761: Cough	54249 (17.9)	299 (24.7)	288 (23.8)	0.635	0.021	133 (25.0)	115 (21.7)	0.218	0.080
31317: Dysphagia	17952 (5.9)	297 (24.5)	293 (24.2)	0.887	0.008	141 (26.6)	137 (25.8)	0.834	0.017
443615: Refractory migraine	13144 (4.3)	286 (23.6)	271 (22.4)	0.499	0.029	134 (25.2)	139 (26.2)	0.779	0.022
46270367: Status epilepticus due to refractory epilepsy	1454 (0.5)	270 (22.3)	224 (18.5)	0.023	0.094	147 (27.7)	131 (24.7)	0.295	0.069
442588: Obstructive sleep apnea syndrome	22780 (7.5)	266 (21.9)	257 (21.2)	0.693	0.018	132 (24.9)	118 (22.2)	0.347	0.062
4332304: Status epilepticus	5481 (1.8)	258 (21.3)	252 (20.8)	0.803	0.012	131 (24.7)	124 (23.4)	0.666	0.031

	Main Analysis Sample	"Any cannabidiol 100 mg/mL" analysis**				"Active cannabidiol 100 mg/mL" analysis**			
		cannabidiol 100 ever=1	Matched controls	p*	SMD*	cannabidiol100 dur c19=1	Matched controls	p*	SMD*
444022: Tetraplegic cerebral palsy	3038 (1.0)	249 (20.5)	244 (20.1)	0.840	0.010	122 (23.0)	126 (23.7)	0.828	0.018
437390: Hypoxemia	11584 (3.8)	230 (19.0)	213 (17.6)	0.400	0.036	118 (22.2)	108 (20.3)	0.500	0.046
257011: Acute upper respiratory infection	41051 (13.6)	225 (18.6)	212 (17.5)	0.526	0.028	109 (20.5)	88 (16.6)	0.114	0.102
255848: Pneumonia	16039 (5.3)	219 (18.1)	204 (16.8)	0.454	0.033	106 (20.0)	97 (18.3)	0.532	0.043
40277917: Intellectual disability	4632 (1.5)	219 (18.1)	213 (17.6)	0.791	0.013	101 (19.0)	103 (19.4)	0.938	0.010
440029: Viral disease	24317 (8.0)	214 (17.7)	201 (16.6)	0.518	0.029	101 (19.0)	93 (17.5)	0.578	0.039
440530: Oropharyngeal dysphagia	6330 (2.1)	213 (17.6)	208 (17.2)	0.830	0.011	103 (19.4)	102 (19.2)	1.000	0.005
441408: Vomiting	19222 (6.4)	205 (16.9)	209 (17.2)	0.871	0.009	95 (17.9)	87 (16.4)	0.569	0.040
4274575: Idiopathic generalized epilepsy	6448 (2.1)	190 (15.7)	184 (15.2)	0.779	0.014	83 (15.6)	84 (15.8)	1.000	0.005
444070: Tachycardia	28844 (9.5)	188 (15.5)	188 (15.5)	1.000	<0.001	93 (17.5)	82 (15.4)	0.408	0.056
43530626: Simple partial seizure	2422 (0.8)	185 (15.3)	208 (17.2)	0.225	0.052	76 (14.3)	82 (15.4)	0.666	0.032
4305080: Abnormal breathing	9478 (3.1)	183 (15.1)	173 (14.3)	0.606	0.023	86 (16.2)	79 (14.9)	0.611	0.036
439780: Autistic disorder	4160 (1.4)	175 (14.4)	166 (13.7)	0.640	0.021	88 (16.6)	100 (18.8)	0.377	0.059
435796: Dehydration	20177 (6.7)	170 (14.0)	172 (14.2)	0.954	0.005	87 (16.4)	90 (16.9)	0.869	0.015
140821: Spasm	16223 (5.4)	169 (13.9)	167 (13.8)	0.953	0.005	86 (16.2)	73 (13.7)	0.302	0.069
45768910: Uncomplicated asthma	34404 (11.4)	168 (13.9)	144 (11.9)	0.163	0.059	85 (16.0)	78 (14.7)	0.610	0.037
436077: Developmental delay	2293 (0.8)	165 (13.6)	166 (13.7)	1.000	0.002	74 (13.9)	61 (11.5)	0.269	0.074
72418: Scoliosis deformity of spine	5245 (1.7)	163 (13.4)	165 (13.6)	0.953	0.005	72 (13.6)	71 (13.4)	1.000	0.006
81902: Urinary tract infectious disease	36897 (12.2)	159 (13.1)	163 (13.4)	0.858	0.010	74 (13.9)	75 (14.1)	1.000	0.005
442077: Anxiety disorder	72507 (24.0)	156 (12.9)	158 (13.0)	0.952	0.005	63 (11.9)	64 (12.1)	1.000	0.006
196523: Diarrhea	33939 (11.2)	155 (12.8)	163 (13.4)	0.674	0.020	60 (11.3)	57 (10.7)	0.845	0.018
46271795: Refractory idiopathic generalized epilepsy	1615 (0.5)	152 (12.5)	157 (13.0)	0.808	0.012	67 (12.6)	76 (14.3)	0.472	0.050
4069943: Neuromuscular scoliosis	1383 (0.5)	148 (12.2)	142 (11.7)	0.754	0.015	71 (13.4)	62 (11.7)	0.458	0.051
437986: Failure to thrive	3302 (1.1)	148 (12.2)	133 (11.0)	0.374	0.039	76 (14.3)	71 (13.4)	0.722	0.027
46271075: Acute hypoxemic respiratory failure	9939 (3.3)	147 (12.1)	142 (11.7)	0.802	0.013	78 (14.7)	73 (13.7)	0.725	0.027
135930: Musculoskeletal finding	11946 (3.9)	144 (11.9)	137 (11.3)	0.703	0.018	65 (12.2)	55 (10.4)	0.383	0.060
437677: Abnormal findings on diagnostic imaging of lung	18296 (6.0)	144 (11.9)	131 (10.8)	0.442	0.034	68 (12.8)	57 (10.7)	0.341	0.064
192450: Retention of urine	8665 (2.9)	142 (11.7)	125 (10.3)	0.299	0.045	77 (14.5)	70 (13.2)	0.594	0.038
4187218: Pneumonitis due to inhaled substance	5234 (1.7)	141 (11.6)	140 (11.6)	1.000	0.003	80 (15.1)	74 (13.9)	0.663	0.032
436222: Altered mental status	18066 (6.0)	139 (11.5)	136 (11.2)	0.898	0.008	69 (13.0)	66 (12.4)	0.854	0.017
4223659: Fatigue	36251 (12.0)	137 (11.3)	121 (10.0)	0.323	0.043	60 (11.3)	65 (12.2)	0.703	0.029
435524: Sleep disorder	8775 (2.9)	136 (11.2)	126 (10.4)	0.556	0.027	67 (12.6)	65 (12.2)	0.926	0.011
439777: Anemia	46241 (15.3)	135 (11.1)	133 (11.0)	0.948	0.005	58 (10.9)	54 (10.2)	0.764	0.025
257907: Disorder of lung	7663 (2.5)	132 (10.9)	124 (10.2)	0.644	0.022	76 (14.3)	69 (13.0)	0.592	0.038
27587: Disturbance of salivary secretion	2083 (0.7)	132 (10.9)	126 (10.4)	0.742	0.016	65 (12.2)	58 (10.9)	0.565	0.041
312437: Dyspnea	44914 (14.8)	132 (10.9)	111 (9.2)	0.176	0.058	58 (10.9)	50 (9.4)	0.477	0.050
376105: West syndrome	876 (0.3)	129 (10.6)	116 (9.6)	0.419	0.036	52 (9.8)	48 (9.0)	0.753	0.026
4029498: Seizure disorder	3028 (1.0)	129 (10.6)	125 (10.3)	0.842	0.011	64 (12.1)	63 (11.9)	1.000	0.006
200219: Abdominal pain	62828 (20.8)	126 (10.4)	118 (9.7)	0.637	0.022	55 (10.4)	58 (10.9)	0.842	0.018
381114: Microcephalus	1619 (0.5)	125 (10.3)	131 (10.8)	0.741	0.016	59 (11.1)	53 (10.0)	0.617	0.037
261880: Atelectasis	11657 (3.9)	125 (10.3)	121 (10.0)	0.840	0.011	60 (11.3)	55 (10.4)	0.693	0.030
434153: Congenital chromosomal disease	1855 (0.6)	122 (10.1)	110 (9.1)	0.448	0.034	45 (8.5)	45 (8.5)	1.000	<0.001
441840: Clinical finding	9766 (3.2)	122 (10.1)	115 (9.5)	0.682	0.019	45 (8.5)	39 (7.3)	0.570	0.042
4310999: Epilepsy, not refractory	4949 (1.6)	122 (10.1)	122 (10.1)	1.000	<0.001	55 (10.4)	60 (11.3)	0.693	0.030
46270364: Status epilepticus due to refractory complex partial seizures	1150 (0.4)	121 (10.0)	119 (9.8)	0.946	0.006	59 (11.1)	54 (10.2)	0.691	0.031
4023310: Blindness AND/OR vision impairment level	3979 (1.3)	118 (9.7)	116 (9.6)	0.945	0.006	49 (9.2)	46 (8.7)	0.830	0.020
4256228: Respiratory failure	7115 (2.4)	118 (9.7)	114 (9.4)	0.836	0.011	63 (11.9)	54 (10.2)	0.433	0.054

	Main Analysis Sample	"Any cannabidiol 100 mg/mL" analysis**				"Active cannabidiol 100 mg/mL" analysis**			
		cannabidiol 100 ever=1	Matched controls	p*	SMD*	cannabidiol100 dur c19=1	Matched controls	p*	SMD*
435232: Disorder of speech and language development	3349 (1.1)	117 (9.7)	113 (9.3)	0.835	0.011	56 (10.5)	52 (9.8)	0.761	0.025
46270366: Status epilepticus due to intractable idiopathic generalized epilepsy	1094 (0.4)	116 (9.6)	101 (8.3)	0.319	0.043	57 (10.7)	57 (10.7)	1.000	<0.001
140214: Eruption	22061 (7.3)	116 (9.6)	112 (9.2)	0.835	0.011	51 (9.6)	34 (6.4)	0.070	0.118
4191650: Acute respiratory distress	4126 (1.4)	116 (9.6)	113 (9.3)	0.890	0.009	66 (12.4)	66 (12.4)	1.000	<0.001
437758: Dependence on enabling machine or device	6081 (2.0)	114 (9.4)	107 (8.8)	0.672	0.020	54 (10.2)	47 (8.9)	0.530	0.045
4248728: Snoring	8682 (2.9)	114 (9.4)	118 (9.7)	0.836	0.011	55 (10.4)	46 (8.7)	0.403	0.058
436070: Vitamin D deficiency	30056 (9.9)	113 (9.3)	109 (9.0)	0.833	0.011	50 (9.4)	41 (7.7)	0.381	0.061
313459: Sleep apnea	10324 (3.4)	112 (9.2)	121 (10.0)	0.582	0.025	51 (9.6)	45 (8.5)	0.593	0.039
137275: Disorder of muscle	7187 (2.4)	111 (9.2)	111 (9.2)	1.000	<0.001	51 (9.6)	42 (7.9)	0.385	0.060
27674: Nausea and vomiting	37324 (12.3)	111 (9.2)	109 (9.0)	0.944	0.006	52 (9.8)	43 (8.1)	0.390	0.059
435517: Acidosis	12756 (4.2)	111 (9.2)	111 (9.2)	1.000	<0.001	51 (9.6)	48 (9.0)	0.833	0.019
79908: Muscle weakness	9240 (3.1)	109 (9.0)	108 (8.9)	1.000	0.003	65 (12.2)	50 (9.4)	0.167	0.091
42873170: Dependence on supplemental oxygen	5375 (1.8)	109 (9.0)	96 (7.9)	0.381	0.039	54 (10.2)	43 (8.1)	0.287	0.072
4168553: Electroencephalogram abnormal	5445 (1.8)	108 (8.9)	134 (11.1)	0.090	0.072	41 (7.7)	47 (8.9)	0.578	0.041
444239: Postprocedural state finding	12896 (4.3)	107 (8.8)	95 (7.8)	0.419	0.036	50 (9.4)	40 (7.5)	0.321	0.068
4141190: Malfunction of gastrostomy tube	2470 (0.8)	106 (8.7)	92 (7.6)	0.335	0.042	51 (9.6)	43 (8.1)	0.450	0.053
376229: Abnormal involuntary movement	5200 (1.7)	106 (8.7)	111 (9.2)	0.776	0.015	47 (8.9)	44 (8.3)	0.826	0.020
75909: Disorder of bone	13225 (4.4)	106 (8.7)	120 (9.9)	0.364	0.040	46 (8.7)	50 (9.4)	0.748	0.026
436962: Insomnia	21623 (7.1)	106 (8.7)	92 (7.6)	0.335	0.042	51 (9.6)	46 (8.7)	0.670	0.033
4022073: Dependence on wheelchair	3402 (1.1)	105 (8.7)	121 (10.0)	0.295	0.045	46 (8.7)	49 (9.2)	0.830	0.020
436230: Blood chemistry abnormal	24671 (8.2)	104 (8.6)	106 (8.7)	0.942	0.006	50 (9.4)	49 (9.2)	1.000	0.007
197672: Urinary incontinence	9353 (3.1)	103 (8.5)	109 (9.0)	0.719	0.018	51 (9.6)	46 (8.7)	0.670	0.033
4309063: Behavior finding	4538 (1.5)	103 (8.5)	105 (8.7)	0.942	0.006	47 (8.9)	50 (9.4)	0.831	0.020
4197485: Epileptic seizure	1453 (0.5)	102 (8.4)	105 (8.7)	0.884	0.009	49 (9.2)	47 (8.9)	0.915	0.013
320128: Essential hypertension	69319 (22.9)	102 (8.4)	111 (9.2)	0.566	0.026	37 (7.0)	38 (7.2)	1.000	0.007
132797: Sepsis	11114 (3.7)	99 (8.2)	99 (8.2)	1.000	<0.001	51 (9.6)	43 (8.1)	0.450	0.053
441277: Mixed receptive-expressive language disorder	1417 (0.5)	98 (8.1)	80 (6.6)	0.186	0.057	53 (10.0)	47 (8.9)	0.599	0.039
318800: Gastroesophageal reflux disease	9816 (3.2)	97 (8.0)	88 (7.3)	0.541	0.028	47 (8.9)	51 (9.6)	0.750	0.026
4282096: Major depression, single episode	60399 (20.0)	96 (7.9)	110 (9.1)	0.344	0.041	31 (5.8)	28 (5.3)	0.789	0.025
320136: Disorder of respiratory system	4711 (1.6)	96 (7.9)	97 (8.0)	1.000	0.003	46 (8.7)	44 (8.3)	0.912	0.014
257293: Disorder of visual pathways	721 (0.2)	95 (7.8)	78 (6.4)	0.207	0.055	42 (7.9)	35 (6.6)	0.478	0.051
4254485: Finding related to attentiveness	10024 (3.3)	95 (7.8)	99 (8.2)	0.822	0.012	40 (7.5)	41 (7.7)	1.000	0.007
378253: Headache	58027 (19.2)	94 (7.8)	88 (7.3)	0.700	0.019	43 (8.1)	43 (8.1)	1.000	<0.001
4195085: Nasal congestion	13828 (4.6)	94 (7.8)	91 (7.5)	0.878	0.009	39 (7.3)	35 (6.6)	0.718	0.030
40484102: Abnormal finding on evaluation procedure	5385 (1.8)	93 (7.7)	80 (6.6)	0.344	0.042	46 (8.7)	28 (5.3)	0.041	0.133
201956: Congenital anomaly of lower limb	1500 (0.5)	92 (7.6)	89 (7.3)	0.877	0.009	44 (8.3)	40 (7.5)	0.733	0.028
201618: Disorder of intestine	17325 (5.7)	91 (7.5)	95 (7.8)	0.819	0.012	42 (7.9)	38 (7.2)	0.727	0.029
374914: Tetraplegia	2020 (0.7)	91 (7.5)	93 (7.7)	0.939	0.006	43 (8.1)	42 (7.9)	1.000	0.007
132617: Diplegic cerebral palsy	1529 (0.5)	90 (7.4)	104 (8.6)	0.331	0.043	37 (7.0)	40 (7.5)	0.813	0.022
44784217: Cardiac arrhythmia	15809 (5.2)	90 (7.4)	83 (6.8)	0.636	0.022	52 (9.8)	39 (7.3)	0.188	0.088
4094822: Foreign body in respiratory tract	2239 (0.7)	90 (7.4)	72 (5.9)	0.167	0.060	38 (7.2)	36 (6.8)	0.904	0.015
37017436: Idiopathic scoliosis of thoracic and lumbar spine	1090 (0.4)	90 (7.4)	78 (6.4)	0.379	0.039	42 (7.9)	34 (6.4)	0.405	0.059
140673: Hypothyroidism	28234 (9.3)	89 (7.3)	87 (7.2)	0.938	0.006	34 (6.4)	34 (6.4)	1.000	<0.001
436233: Delayed milestone	1355 (0.4)	89 (7.3)	94 (7.8)	0.758	0.016	37 (7.0)	39 (7.3)	0.905	0.015
4310235: Reduced mobility	6621 (2.2)	88 (7.3)	88 (7.3)	1.000	<0.001	51 (9.6)	45 (8.5)	0.593	0.039
435515: Hypo-osmolality and or hyponatremia	12833 (4.2)	87 (7.2)	80 (6.6)	0.630	0.023	41 (7.7)	38 (7.2)	0.815	0.022
4168212: Restlessness and agitation	5474 (1.8)	86 (7.1)	91 (7.5)	0.755	0.016	41 (7.7)	41 (7.7)	1.000	<0.001

	Main Analysis Sample	"Any cannabidiol 100 mg/mL" analysis**				"Active cannabidiol 100 mg/mL" analysis**			
		cannabidiol 100 ever=1	Matched controls	p*	SMD*	cannabidiol100 dur c19=1	Matched controls	p*	SMD*
375800: Dystonia	1735 (0.6)	86 (7.1)	86 (7.1)	1.000	<0.001	41 (7.7)	44 (8.3)	0.821	0.021
4152351: Abdominal distension, gaseous	9958 (3.3)	86 (7.1)	92 (7.6)	0.697	0.019	37 (7.0)	34 (6.4)	0.806	0.023
46273463: Upper respiratory tract infection due to Influenza	13959 (4.6)	86 (7.1)	71 (5.9)	0.248	0.050	47 (8.9)	42 (7.9)	0.658	0.034
4169095: Bradycardia	10580 (3.5)	86 (7.1)	88 (7.3)	0.937	0.006	35 (6.6)	31 (5.8)	0.703	0.031
4041283: General finding of observation of patient	15638 (5.2)	85 (7.0)	88 (7.3)	0.875	0.010	34 (6.4)	35 (6.6)	1.000	0.008
437113: Asthenia	21843 (7.2)	85 (7.0)	96 (7.9)	0.440	0.035	30 (5.6)	26 (4.9)	0.680	0.034
4133224: Lobar pneumonia	5328 (1.8)	85 (7.0)	78 (6.4)	0.627	0.023	41 (7.7)	39 (7.3)	0.907	0.014
437833: Hypokalemia	21618 (7.1)	84 (6.9)	82 (6.8)	0.936	0.007	36 (6.8)	39 (7.3)	0.811	0.022
320536: Electrocardiogram abnormal	22491 (7.4)	84 (6.9)	89 (7.3)	0.752	0.016	41 (7.7)	39 (7.3)	0.907	0.014
372328: Otitis media	9062 (3.0)	84 (6.9)	80 (6.6)	0.808	0.013	34 (6.4)	29 (5.5)	0.603	0.040
133228: Dental caries	5577 (1.8)	82 (6.8)	73 (6.0)	0.507	0.030	42 (7.9)	39 (7.3)	0.817	0.021
79061: Slow transit constipation	4278 (1.4)	82 (6.8)	71 (5.9)	0.404	0.037	45 (8.5)	36 (6.8)	0.355	0.064
374915: Localization-related epilepsy	2617 (0.9)	80 (6.6)	89 (7.3)	0.524	0.029	34 (6.4)	38 (7.2)	0.714	0.030
374924: Generalized convulsive epilepsy	708 (0.2)	78 (6.4)	51 (4.2)	0.019	0.099				
4079975: Congenital malformation	2181 (0.7)	78 (6.4)	79 (6.5)	1.000	0.003	34 (6.4)	27 (5.1)	0.429	0.057
438068: Disease due to Enterovirus	1397 (0.5)	77 (6.4)	75 (6.2)	0.933	0.007	38 (7.2)	41 (7.7)	0.815	0.022
443410: Infective pneumonia	3089 (1.0)	77 (6.4)	67 (5.5)	0.439	0.035	39 (7.3)	34 (6.4)	0.628	0.037
4306496: Refractory infantile spasms	319 (0.1)	76 (6.3)	66 (5.4)	0.436	0.035	34 (6.4)	34 (6.4)	1.000	<0.001
434675: Complication of gastrostomy	1819 (0.6)	76 (6.3)	86 (7.1)	0.464	0.033	37 (7.0)	36 (6.8)	1.000	0.007
4152347: Drowsy	4885 (1.6)	75 (6.2)	70 (5.8)	0.732	0.017	33 (6.2)	40 (7.5)	0.467	0.052
4266029: Airway trauma	1333 (0.4)	74 (6.1)	62 (5.1)	0.332	0.043	30 (5.6)	28 (5.3)	0.893	0.017
432898: Severe intellectual disability	942 (0.3)	74 (6.1)	79 (6.5)	0.738	0.017	33 (6.2)	37 (7.0)	0.711	0.030
432870: Thrombocytopenic disorder	11534 (3.8)	73 (6.0)	75 (6.2)	0.932	0.007	34 (6.4)	42 (7.9)	0.405	0.059
377889: Hearing loss	6774 (2.2)	73 (6.0)	70 (5.8)	0.863	0.011	36 (6.8)	29 (5.5)	0.442	0.055
439654: Partial epilepsy with impairment of consciousness	1464 (0.5)	73 (6.0)	80 (6.6)	0.616	0.024	34 (6.4)	38 (7.2)	0.714	0.030
4166231: Genetic predisposition	2362 (0.8)	72 (5.9)	61 (5.0)	0.372	0.040	31 (5.8)	28 (5.3)	0.789	0.025
438409: Attention deficit hyperactivity disorder	8478 (2.8)	72 (5.9)	65 (5.4)	0.598	0.025	33 (6.2)	38 (7.2)	0.623	0.038
443200: Nervous system symptoms	6470 (2.1)	71 (5.9)	69 (5.7)	0.931	0.007	32 (6.0)	39 (7.3)	0.461	0.053
254061: Pleural effusion	8488 (2.8)	71 (5.9)	57 (4.7)	0.238	0.052	38 (7.2)	40 (7.5)	0.906	0.014
317002: Low blood pressure	12702 (4.2)	71 (5.9)	72 (5.9)	1.000	0.004	34 (6.4)	28 (5.3)	0.513	0.048
4329041: Pain	17357 (5.7)	70 (5.8)	79 (6.5)	0.499	0.031	34 (6.4)	32 (6.0)	0.899	0.016
4043738: Hydrocephalus	3283 (1.1)	70 (5.8)	73 (6.0)	0.863	0.011	34 (6.4)	37 (7.0)	0.806	0.023
374034: Visual disturbance	12831 (4.2)	69 (5.7)	71 (5.9)	0.931	0.007	27 (5.1)	25 (4.7)	0.887	0.018
377085: Congenital anomaly of brain	822 (0.3)	69 (5.7)	66 (5.4)	0.859	0.011	36 (6.8)	34 (6.4)	0.902	0.015
433736: Obesity	55813 (18.5)	68 (5.6)	79 (6.5)	0.395	0.038	28 (5.3)	28 (5.3)	1.000	<0.001
197675: Incontinence of feces	3181 (1.1)	68 (5.6)	67 (5.5)	1.000	0.004	36 (6.8)	32 (6.0)	0.707	0.031
253482: Cortical blindness	624 (0.2)	68 (5.6)	65 (5.4)	0.858	0.011	28 (5.3)	23 (4.3)	0.566	0.044
438733: Profound intellectual disability	755 (0.2)	67 (5.5)	56 (4.6)	0.355	0.041	33 (6.2)	34 (6.4)	1.000	0.008
436096: Chronic pain	45132 (14.9)	67 (5.5)	70 (5.8)	0.860	0.011				
197320: Acute renal failure syndrome	17050 (5.6)	66 (5.4)	67 (5.5)	1.000	0.004	30 (5.6)	32 (6.0)	0.896	0.016
4180628: Disorder of body system	11562 (3.8)	66 (5.4)	68 (5.6)	0.929	0.007	31 (5.8)	23 (4.3)	0.328	0.069
31967: Nausea	36937 (12.2)	66 (5.4)	59 (4.9)	0.582	0.026	31 (5.8)	27 (5.1)	0.685	0.033
432545: Bacterial infectious disease	22292 (7.4)	66 (5.4)	66 (5.4)	1.000	<0.001	31 (5.8)	30 (5.6)	1.000	0.008
4289309: Atrial septal defect	4658 (1.5)	64 (5.3)	69 (5.7)	0.721	0.018				
77670: Chest pain	49355 (16.3)	64 (5.3)	52 (4.3)	0.295	0.046	29 (5.5)	29 (5.5)	1.000	<0.001
319049: Acute respiratory failure	2467 (0.8)	64 (5.3)	58 (4.8)	0.642	0.023	35 (6.6)	37 (7.0)	0.903	0.015
375415: Injury of head	12545 (4.1)	64 (5.3)	56 (4.6)	0.512	0.030	26 (4.9)	33 (6.2)	0.422	0.058

	Main Analysis Sample	"Any cannabidiol 100 mg/mL" analysis**				"Active cannabidiol 100 mg/mL" analysis**			
		cannabidiol 100 ever=1	Matched controls	p*	SMD*	cannabidiol100 dur c19=1	Matched controls	p*	SMD*
257007: Allergic rhinitis	16219 (5.4)	64 (5.3)	54 (4.5)	0.396	0.038	35 (6.6)	25 (4.7)	0.232	0.082
4272240: Malaise	14392 (4.8)	64 (5.3)	58 (4.8)	0.642	0.023	26 (4.9)	28 (5.3)	0.889	0.017
76791: Acquired deformity of hip	675 (0.2)	63 (5.2)	60 (5.0)	0.853	0.011				
4177862: Chronic hypoxemic respiratory failure	2875 (1.0)	63 (5.2)	50 (4.1)	0.248	0.051	36 (6.8)	30 (5.6)	0.525	0.047
199075: Neurogenic bladder	3421 (1.1)	63 (5.2)	69 (5.7)	0.655	0.022	31 (5.8)	28 (5.3)	0.789	0.025
42537748: Acquired absence of organ	19738 (6.5)	62 (5.1)	62 (5.1)	1.000	<0.001				
4267558: Precocious puberty	527 (0.2)	62 (5.1)	53 (4.4)	0.445	0.035				
4148091: Developmental disorder of motor function	1834 (0.6)	62 (5.1)	50 (4.1)	0.287	0.047				
37016114: Acute on chronic hypoxemic respiratory failure	3270 (1.1)	62 (5.1)	59 (4.9)	0.852	0.011	33 (6.2)	26 (4.9)	0.422	0.058
438398: Leukocytosis	14493 (4.8)	62 (5.1)	55 (4.5)	0.570	0.027	29 (5.5)	30 (5.6)	1.000	0.008
441417: Incoordination	2597 (0.9)	62 (5.1)	61 (5.0)	1.000	0.004				
4150125: Persistent pain following procedure	10166 (3.4)	62 (5.1)	68 (5.6)	0.652	0.022	29 (5.5)	24 (4.5)	0.573	0.043
201620: Kidney stone	9454 (3.1)	61 (5.0)	65 (5.4)	0.784	0.015	27 (5.1)	28 (5.3)	1.000	0.009
4275359: Mental alertness - finding	7484 (2.5)	60 (5.0)	75 (6.2)	0.215	0.054	26 (4.9)	33 (6.2)	0.422	0.058
4033802: Granulomatous disorder of the skin and subcutaneous tissue	1728 (0.6)	60 (5.0)	59 (4.9)	1.000	0.004	32 (6.0)	28 (5.3)	0.690	0.033
437092: Physiological development failure	1281 (0.4)	60 (5.0)	54 (4.5)	0.631	0.023				
4042889: Finding relating to drug misuse behavior	11282 (3.7)	60 (5.0)	40 (3.3)	0.052	0.083				
436659: Iron deficiency anemia	22179 (7.3)	60 (5.0)	61 (5.0)	1.000	0.004	31 (5.8)	35 (6.6)	0.703	0.031
432795: Traumatic AND/OR non-traumatic injury	11133 (3.7)	60 (5.0)	39 (3.2)	0.040	0.088	27 (5.1)	24 (4.5)	0.774	0.026
4009585: Abnormal urine	9017 (3.0)	60 (5.0)	46 (3.8)	0.197	0.057	27 (5.1)	28 (5.3)	1.000	0.009
4075828: Global developmental delay	465 (0.2)					40 (7.5)	29 (5.5)	0.213	0.084
374375: Impacted cerumen	8696 (2.9)					34 (6.4)	34 (6.4)	1.000	<0.001
439935: Abnormal posture	1534 (0.5)					32 (6.0)	<20	0.090	0.110
4306572: Hypoxia	1024 (0.3)					31 (5.8)	28 (5.3)	0.789	0.025
194081: Acute cystitis	14727 (4.9)					30 (5.6)	28 (5.3)	0.893	0.017
46273390: Dependence on respirator	2898 (1.0)					30 (5.6)	27 (5.1)	0.785	0.025
4201387: Tracheostomy present	2390 (0.8)					30 (5.6)	27 (5.1)	0.785	0.025
4041136: Intellectual functioning disability	395 (0.1)					29 (5.5)	28 (5.3)	1.000	0.008
4047124: Expressive language disorder	1591 (0.5)					28 (5.3)	26 (4.9)	0.889	0.017
4306923: Chronic idiopathic constipation	2455 (0.8)					28 (5.3)	27 (5.1)	1.000	0.009
312940: Acute-on-chronic respiratory failure	1674 (0.6)					28 (5.3)	24 (4.5)	0.670	0.035
24609: Hypoglycemia	6129 (2.0)					28 (5.3)	27 (5.1)	1.000	0.009
260125: Acute bronchiolitis	2532 (0.8)					28 (5.3)	31 (5.8)	0.789	0.025
4146581: Mild intermittent asthma	13728 (4.5)					27 (5.1)	27 (5.1)	1.000	<0.001
45766714: Inflammatory dermatosis	13246 (4.4)					27 (5.1)	30 (5.6)	0.785	0.025
43530807: Allergic disposition	11352 (3.8)					27 (5.1)	21 (4.0)	0.460	0.054
317376: Tachypnea	2206 (0.7)					26 (4.9)	28 (5.3)	0.889	0.017
Non-cannabidiol medications									
acetaminophen	174910 (57.8)	757 (62.5)	735 (60.6)	0.381	0.037	342 (64.4)	327 (61.6)	0.374	0.059
albuterol	66457 (22.0)	437 (36.1)	436 (36.0)	1.000	0.002	214 (40.3)	212 (39.9)	0.950	0.008
amoxicillin	59621 (19.7)	364 (30.0)	338 (27.9)	0.263	0.047	162 (30.5)	147 (27.7)	0.344	0.062
bacitracin	15039 (5.0)					103 (19.4)	101 (19.0)	0.938	0.010
bacitracin zinc	5149 (1.7)	92 (7.6)	96 (7.9)	0.820	0.012	46 (8.7)	51 (9.6)	0.670	0.033
baclofen	10145 (3.4)	154 (12.7)	146 (12.0)	0.666	0.020	74 (13.9)	62 (11.7)	0.312	0.068
barium sulfate	5801 (1.9)	77 (6.4)	62 (5.1)	0.221	0.053	40 (7.5)	36 (6.8)	0.721	0.029
bisacodyl	33270 (11.0)	168 (13.9)	176 (14.5)	0.684	0.019	85 (16.0)	86 (16.2)	1.000	0.005
botulinum toxin type A	3381 (1.1)					30 (5.6)	22 (4.1)	0.320	0.070

	Main Analysis Sample	"Any cannabidiol 100 mg/mL" analysis**				"Active cannabidiol 100 mg/mL" analysis**			
		cannabidiol 100 ever=1	Matched controls	p*	SMD*	cannabidiol100 dur c19=1	Matched controls	p*	SMD*
budesonide	12005 (4.0)	123 (10.1)	117 (9.7)	0.734	0.017	68 (12.8)	63 (11.9)	0.709	0.029
calcium chloride	105873 (35.0)	532 (43.9)	514 (42.4)	0.486	0.030	241 (45.4)	221 (41.6)	0.240	0.076
cefazolin	48738 (16.1)	283 (23.3)	271 (22.4)	0.595	0.024	143 (26.9)	120 (22.6)	0.118	0.101
ceftriaxone	23970 (7.9)	207 (17.1)	200 (16.5)	0.744	0.016	107 (20.2)	102 (19.2)	0.758	0.024
cephalexin	39320 (13.0)	173 (14.3)	172 (14.2)	1.000	0.002	74 (13.9)	75 (14.1)	1.000	0.005
cetirizine hydrochloride	15771 (5.2)	176 (14.5)	166 (13.7)	0.600	0.024	94 (17.7)	91 (17.1)	0.872	0.015
chlorhexidine gluconate	10280 (3.4)	108 (8.9)	105 (8.7)	0.886	0.009	58 (10.9)	54 (10.2)	0.764	0.025
cholecalciferol	34339 (11.4)	331 (27.3)	320 (26.4)	0.647	0.021	162 (30.5)	151 (28.4)	0.501	0.045
clobazam	3987 (1.3)	627 (51.7)	636 (52.5)	0.745	0.015	293 (55.2)	296 (55.7)	0.902	0.011
clonazepam	15318 (5.1)	512 (42.2)	495 (40.8)	0.510	0.029	253 (47.6)	254 (47.8)	1.000	0.004
clonidine hydrochloride	8487 (2.8)	143 (11.8)	130 (10.7)	0.441	0.034	66 (12.4)	77 (14.5)	0.369	0.061
dexamethasone	54039 (17.9)	419 (34.6)	401 (33.1)	0.466	0.031	197 (37.1)	174 (32.8)	0.157	0.091
dexamethasone phosphate	33651 (11.1)	271 (22.4)	262 (21.6)	0.695	0.018	122 (23.0)	100 (18.8)	0.113	0.102
dexamethasone sodium phosphate	3906 (1.3)					26 (4.9)	21 (4.0)	0.551	0.046
dexmedetomidine	11194 (3.7)	168 (13.9)	157 (13.0)	0.551	0.027	87 (16.4)	71 (13.4)	0.196	0.085
diazepam	22857 (7.6)	688 (56.8)	654 (54.0)	0.178	0.057	333 (62.7)	320 (60.3)	0.449	0.050
diphenhydramine hydrochloride	62925 (20.8)	263 (21.7)	253 (20.9)	0.655	0.020	128 (24.1)	120 (22.6)	0.612	0.036
divalproex sodium	9242 (3.1)	265 (21.9)	251 (20.7)	0.519	0.028	109 (20.5)	119 (22.4)	0.501	0.046
docusate sodium	82383 (27.2)	205 (16.9)	199 (16.4)	0.785	0.013				
enoxaparin sodium	31406 (10.4)	131 (10.8)	137 (11.3)	0.746	0.016	59 (11.1)	58 (10.9)	1.000	0.006
erythromycin	15708 (5.2)					51 (9.6)	48 (9.0)	0.833	0.019
famotidine	52892 (17.5)	251 (20.7)	259 (21.4)	0.727	0.016	123 (23.2)	120 (22.6)	0.884	0.013
fentanyl	99096 (32.8)	446 (36.8)	446 (36.8)	1.000	<0.001	196 (36.9)	177 (33.3)	0.247	0.075
ferrous sulfate	44955 (14.9)					73 (13.7)	68 (12.8)	0.718	0.028
fluticasone propionate	34469 (11.4)	280 (23.1)	257 (21.2)	0.282	0.046	133 (25.0)	118 (22.2)	0.312	0.067
gabapentin	36394 (12.0)	184 (15.2)	167 (13.8)	0.356	0.040	90 (16.9)	70 (13.2)	0.103	0.105
gadobenate	7837 (2.6)	91 (7.5)	78 (6.4)	0.339	0.042	38 (7.2)	34 (6.4)	0.714	0.030
gadobutrol	10255 (3.4)	67 (5.5)	58 (4.8)	0.463	0.034	29 (5.5)	27 (5.1)	0.891	0.017
glucose	60777 (20.1)	514 (42.4)	506 (41.7)	0.773	0.013	239 (45.0)	239 (45.0)	1.000	<0.001
glycerin	16763 (5.5)	205 (16.9)	193 (15.9)	0.546	0.027	97 (18.3)	92 (17.3)	0.748	0.025
glycopyrrolate	15416 (5.1)	299 (24.7)	285 (23.5)	0.537	0.027	153 (28.8)	140 (26.4)	0.410	0.055
heparin sodium	30612 (10.1)	177 (14.6)	201 (16.6)	0.198	0.055	85 (16.0)	86 (16.2)	1.000	0.005
ibuprofen	117803 (38.9)	498 (41.1)	489 (40.3)	0.741	0.015	230 (43.3)	228 (42.9)	0.951	0.008
influenza	89163 (29.5)	359 (29.6)	357 (29.5)	0.965	0.004	159 (29.9)	135 (25.4)	0.115	0.101
iohexol	19290 (6.4)	128 (10.6)	103 (8.5)	0.097	0.070	55 (10.4)	52 (9.8)	0.838	0.019
iopamidol	21966 (7.3)	91 (7.5)	81 (6.7)	0.477	0.032	35 (6.6)	36 (6.8)	1.000	0.008
ipratropium bromide	22202 (7.3)	177 (14.6)	173 (14.3)	0.862	0.009	84 (15.8)	85 (16.0)	1.000	0.005
ketamine	10463 (3.5)					59 (11.1)	51 (9.6)	0.481	0.050
ketorolac	67545 (22.3)	226 (18.6)	219 (18.1)	0.753	0.015	94 (17.7)	79 (14.9)	0.245	0.077
lacosamide	7420 (2.5)	339 (28.0)	353 (29.1)	0.559	0.026	157 (29.6)	147 (27.7)	0.541	0.042
lactaseibacillus rhamnosus GG	1636 (0.5)					56 (10.5)	47 (8.9)	0.407	0.057
lactulose	8139 (2.7)	100 (8.3)	101 (8.3)	1.000	0.003	50 (9.4)	47 (8.9)	0.831	0.020
lamotrigine	15319 (5.1)	269 (22.2)	306 (25.2)	0.086	0.072	120 (22.6)	127 (23.9)	0.663	0.031
lansoprazole	4564 (1.5)					73 (13.7)	66 (12.4)	0.585	0.039
levetiracetam	38760 (12.8)	585 (48.3)	594 (49.0)	0.745	0.015	258 (48.6)	254 (47.8)	0.854	0.015
levocarnitine	1620 (0.5)	195 (16.1)	169 (13.9)	0.155	0.060	96 (18.1)	79 (14.9)	0.186	0.086
lidocaine	125530 (41.5)	570 (47.0)	554 (45.7)	0.541	0.027	246 (46.3)	249 (46.9)	0.902	0.011

	Main Analysis Sample	"Any cannabidiol 100 mg/mL" analysis**				"Active cannabidiol 100 mg/mL" analysis**			
		cannabidiol 100 ever=1	Matched controls	p*	SMD*	cannabidiol100 dur c19=1	Matched controls	p*	SMD*
lidocaine hydrochloride	98491 (32.6)	486 (40.1)	465 (38.4)	0.405	0.036	215 (40.5)	212 (39.9)	0.900	0.012
lorazepam	46247 (15.3)	601 (49.6)	592 (48.8)	0.745	0.015	290 (54.6)	291 (54.8)	1.000	0.004
melatonin	21472 (7.1)	270 (22.3)	248 (20.5)	0.298	0.044	139 (26.2)	132 (24.9)	0.673	0.030
metoclopramide	45112 (14.9)	103 (8.5)	89 (7.3)	0.328	0.043	47 (8.9)	41 (7.7)	0.578	0.041
midazolam	58026 (19.2)	597 (49.3)	582 (48.0)	0.569	0.025	277 (52.2)	256 (48.2)	0.220	0.079
morphine sulfate	37618 (12.4)	188 (15.5)	202 (16.7)	0.472	0.031	79 (14.9)	63 (11.9)	0.176	0.089
mupirocin	16351 (5.4)	151 (12.5)	146 (12.0)	0.804	0.013	87 (16.4)	78 (14.7)	0.498	0.047
neostigmine methylsulfate	8321 (2.8)					51 (9.6)	43 (8.1)	0.450	0.053
nystatin	19185 (6.3)	174 (14.4)	177 (14.6)	0.908	0.007	79 (14.9)	75 (14.1)	0.794	0.021
omeprazole	36069 (11.9)	208 (17.2)	190 (15.7)	0.351	0.040	105 (19.8)	105 (19.8)	1.000	<0.001
onabotulinumtoxinA	4161 (1.4)	83 (6.8)	83 (6.8)	1.000	<0.001	44 (8.3)	41 (7.7)	0.821	0.021
ondansetron	144932 (47.9)	643 (53.1)	615 (50.7)	0.272	0.046	294 (55.4)	271 (51.0)	0.176	0.087
oseltamivir	18617 (6.2)					65 (12.2)	61 (11.5)	0.776	0.023
oxcarbazepine	6634 (2.2)					71 (13.4)	77 (14.5)	0.658	0.033
oxycodone hydrochloride	61633 (20.4)	283 (23.3)	279 (23.0)	0.885	0.008	113 (21.3)	121 (22.8)	0.604	0.036
oxymetazoline hydrochloride	4490 (1.5)	89 (7.3)	78 (6.4)	0.423	0.036	47 (8.9)	41 (7.7)	0.578	0.041
pantoprazole	38630 (12.8)	168 (13.9)	149 (12.3)	0.278	0.047	68 (12.8)	74 (13.9)	0.652	0.033
perampanel	713 (0.2)	141 (11.6)	115 (9.5)	0.099	0.070	64 (12.1)	67 (12.6)	0.852	0.017
phenobarbital	5124 (1.7)	163 (13.4)	152 (12.5)	0.546	0.027	76 (14.3)	66 (12.4)	0.417	0.055
polyethylene glycol	63516 (21.0)	542 (44.7)	531 (43.8)	0.683	0.018	261 (49.2)	246 (46.3)	0.390	0.057
potassium chloride	126529 (41.8)	655 (54.0)	622 (51.3)	0.193	0.055	293 (55.2)	278 (52.4)	0.389	0.057
prednisolone	35828 (11.8)	210 (17.3)	204 (16.8)	0.787	0.013	101 (19.0)	86 (16.2)	0.259	0.074
propofol	51629 (17.1)	478 (39.4)	467 (38.5)	0.677	0.019	228 (42.9)	207 (39.0)	0.212	0.081
ranitidine	21790 (7.2)	192 (15.8)	197 (16.3)	0.825	0.011	97 (18.3)	88 (16.6)	0.518	0.045
rocuronium bromide	21187 (7.0)	282 (23.3)	276 (22.8)	0.809	0.012	139 (26.2)	122 (23.0)	0.254	0.074
rufinamide	599 (0.2)					81 (15.3)	87 (16.4)	0.674	0.031
sennosides	49624 (16.4)	298 (24.6)	277 (22.9)	0.340	0.041	147 (27.7)	130 (24.5)	0.264	0.073
simethicone	61298 (20.3)	138 (11.4)	133 (11.0)	0.797	0.013	64 (12.1)	63 (11.9)	1.000	0.006
sodium chloride	160406 (53.0)	836 (69.0)	818 (67.5)	0.458	0.032	380 (71.6)	367 (69.1)	0.420	0.054
sodium phosphate	17781 (5.9)					93 (17.5)	85 (16.0)	0.565	0.040
sugammadex	12652 (4.2)	111 (9.2)	115 (9.5)	0.834	0.011				
sulfamethoxazole	25831 (8.5)	149 (12.3)	156 (12.9)	0.713	0.017	77 (14.5)	82 (15.4)	0.731	0.026
topiramate	16989 (5.6)	241 (19.9)	216 (17.8)	0.213	0.053	119 (22.4)	111 (20.9)	0.602	0.037
valproic acid	3470 (1.1)	210 (17.3)	191 (15.8)	0.325	0.042	87 (16.4)	80 (15.1)	0.613	0.036
vancomycin	19898 (6.6)	165 (13.6)	181 (14.9)	0.384	0.038	83 (15.6)	73 (13.7)	0.435	0.053
vitamin B	25480 (8.4)	164 (13.5)	172 (14.2)	0.681	0.019	83 (15.6)	70 (13.2)	0.294	0.070
water	8819 (2.9)	108 (8.9)	102 (8.4)	0.718	0.018	56 (10.5)	46 (8.7)	0.349	0.064
zonisamide	5045 (1.7)	201 (16.6)	210 (17.3)	0.665	0.020	91 (17.1)	110 (20.7)	0.159	0.091

*The bold font in the p-value and SMD columns highlight all p-values less than 0.05 and all SMDs greater than 0.10.

**The gray shaded cells indicate that the variable was not included in that column's matching model due to cell size being <=5% of the cannabidiol 100 mg/mL sample. For those variables that were also not included in the multivariable outcomes models, we omit their frequencies because the variables were not used in any of the matching or outcomes models.

§Of the 69 cannabidiol100_ever patients with "Other Race", 26 had a race_concept_name of "Asian" and the remaining 43 had a race_concept_name of "Other". Of the 120 cannabidiol100_ever patients with "Race Imprecise", 98 had a race_concept_name of "No matching concept" and the remainder were counts of less than 20 spread for the remaining concept IDs in this category (described in the Methods section).

REFERENCES AND NOTES

1. S. R. Weiss, J. L. Leibowitz, Coronavirus pathogenesis. *Adv. Virus Res.* 81, 85–164 (2011).
2. S. E. Galloway, P. Paul, D. R. MacCannell, M. A. Johansson, J. T. Brooks, A. M. Neil, R. B. Slayton, S. Tong, B. J. Silk, G. L. Armstrong, M. Biggerstaff, V. G. Dugan, Emergence of SARS-CoV-2 B.1.1.7 lineage - United States, December 29, 2020-January 12, 2021. *MMWR Morb. Mortal. Wkly Rep.* 70, 95–99 (2021).
3. M. Hoffmann, H. Kleine-Weber, S. Schroeder, N. Krüger, T. Herrler, S. Erichsen, T. S. Schiergens, G. Herrler, N.-H. Wu, A. Nitsche, M. A. Müller, C. Drosten, S. Pöhlmann, SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* 181, 271–280.e8 (2020).
4. J. Jaimes, J. Millet, G. Whittaker, Proteolytic cleavage of the SARS-CoV-2 spike protein and the role of the novel S1/S2 Site. *iScience* 23, 101212 (2020).
5. S. Matsuyama, N. Nao, K. Shirato, M. Kawase, S. Saito, I. Takayama, N. Nagata, T. Sekizuka, H. Katoh, F. Kato, M. Sakata, M. Tahara, S. Kutsuna, N. Ohmagari, M. Kuroda, T. Suzuki, T. Kageyama, M. Takeda, Enhanced isolation of SARS-CoV-2 by TMPRSS2-expressing cells. *Proc. Natl. Acad. Sci. U.S.A.* 117, 7001–7003 (2020).
6. R. Wang, C. R. Simoneau, J. Kulsuptrakul, M. Bouhaddou, K. A. Travisano, J. M. Hayashi, J. Carlson-Stevermer, J. R. Zengel, C. M. Richards, P. Fozouni, J. Oki, L. Rodriguez, B. Joehnk, K. Walcott, K. Holden, A. Sil, J. E. Carette, N. J. Krogan, M. Ott, Genetic screens identify host factors for SARS-CoV-2 and common cold coronaviruses. *Cell* 184, 106–119.e14 (2021).
7. W. M. Schneider, J. M. Luna, H.-H. Hoffmann, F. J. Sánchez-Rivera, A. A. Leal, A. W. Ashbrook, J. L. Pen, I. Ricardo-Lax, E. Michailidis, A. Peace, A. F. Stenzel, S. W. Lowe, M. R. MacDonald, C.

- M. Rice, J. T. Poirier, Genome-scale identification of SARS-CoV-2 and pan-coronavirus host factor networks. *Cell* 184, 120–132.e14 (2021).
8. Z. Daniloski, T. X. Jordan, H.-H. Wessels, D. A. Hoagland, S. Kasela, M. Legut, S. Maniatis, E. P. Mimitou, L. Lu, E. Geller, O. Danziger, B. R. Rosenberg, H. Phatnani, P. Smibert, T. Lappalainen, B. R. tenOever, N. E. Sanjana, Identification of required host factors for SARS-CoV-2 infection in human cells. *Cell* 184, 92–105.e16 (2021).
9. D. W. Kneller, G. Phillips, H. M. O'Neill, R. Jedrzejczak, L. Stols, P. Langan, A. Joachimiak, L. Coates, A. Kovalevsky, Structural plasticity of SARS-CoV-2 3CL Mpro active site cavity revealed by room temperature x-ray crystallography. *Nat. Commun.* 11, 3202 (2020).
10. J. Osipiuk, S. A. Azizi, S. Dvorkin, M. Endres, R. Jedrzejczak, K. A. Jones, S. Kang, R. S. Kathayat, Y. Kim, V. G. Lisnyak, S. L. Maki, V. Nicolaescu, C. A. Taylor, C. Tesar, Y.A. Zhang, Z. Zhou, G. Randall, K. Michalska, S. A. Snyder, B. C. Dickinson, A. Joachimiak, Structure of papain-like protease from SARS-CoV-2 and its complexes with non-covalent inhibitors. *Nat. Commun.* 12, 743 (2021).
11. F. Shahbazi, V. Grandi, A. Banerjee, J. F. Trant, Cannabinoids and cannabinoid receptors: The story so far. *iScience* 23, 101301 (2020).
12. K. M. Nelson, J. Bisson, G. Singh, J. G. Graham, S.N. Chen, J. B. Friesen, J. L. Dahlin, M. Niemitz, M. A. Walters, G. F. Pauli, The essential medicinal chemistry of cannabidiol (CBD). *J. Med. Chem.* 63, 12137–12155 (2020).
13. K. Sekar, A. Pack, Epidiolex as adjunct therapy for treatment of refractory epilepsy: A comprehensive review with a focus on adverse effects. *F1000Res* 8, F1000 (2019).
14. H. I. Lowe, N. J. Toyang, W. McLaughlin, Potential of cannabidiol for the treatment of viral hepatitis. *Pharm. Res.* 9, 116–118 (2017).

15. L. Taylor, B. Gidal, G. Blakey, B. Tayo, G. Morrison, A phase I, randomized, double-blind, placebo-controlled, single ascending dose, multiple dose, and food effect trial of the safety, tolerability and pharmacokinetics of highly purified cannabidiol in healthy subjects. *CNS Drugs* 32, 1053–1067 (2018).
16. S. M. Anil, N. Shalev, A. C. Vinayaka, S. Nadarajan, D. Namdar, E. Belausov, I. Shoval, K. A. Mani, G. Mechrez, H. Koltai, *Cannabis* compounds exhibit anti-inflammatory activity in vitro in COVID-19-related inflammation in lung epithelial cells and pro-inflammatory activity in macrophages. *Sci. Rep.* 11, 1462 (2021).
17. K. H. D. Crawford, R. Eguia, A. S. Dingens, A. N. Loes, K. D. Malone, C. R. Wolf, H. Y. Chu, M. A. Tortorici, D. Veessler, M. Murphy, D. Pettie, N. P. King, A. B. Balazs, J. D. Bloom, Protocol and reagents for pseudotyping lentiviral particles with SARS-CoV-2 spike protein for neutralization assays. *Viruses* 12, 513 (2020).
18. C. Hetz, E. Chevet, S. A. Oakes, Proteostasis control by the unfolded protein response. *Nat. Cell Biol.* 17, 829–838 (2015).
19. R. Dash, M. C. Ali, I. Jahan, Y. A. Munni, S. Mitra, M. A. Hannan, B. Timalisina, D. F. Oktaviani, H. J. Choi, I. S. Moon, Emerging potential of cannabidiol in reversing proteinopathies. *Ageing Res. Rev.* 65, 101209 (2021).
20. J. Bechill, Z. Chen, J. W. Brewer, S. C. Baker, Mouse hepatitis virus infection activates the Ire1/XBP1 pathway of the unfolded protein response. *Adv. Exp. Med. Biol.* 581, 139–144 (2006).
21. T. S. Fung, D. X. Liu, The ER stress sensor IRE1 and MAP kinase ERK modulate autophagy induction in cells infected with coronavirus infectious bronchitis virus. *Virology* 533, 34–44 (2019).
22. C. Hetz, K. Zhang, R. J. Kaufman, Mechanisms, regulation and functions of the unfolded protein response. *Nat. Rev. Mol. Cell Biol.* 21, 421–438 (2020).

23. O. Haller, G. Kochs, F. Weber, The interferon response circuit: Induction and suppression by pathogenic viruses. *Virology* 344, 119–130 (2006).
24. D. Blanco-Melo, B. E. Nilsson-Payant, W.-C. Liu, S. Uhl, D. Hoagland, R. Møller, T. X. Jordan, K. Oishi, M. Panis, D. Sachs, T. T. Wang, R. E. Schwartz, J. K. Lim, R. A. Albrecht, B. R. tenOever, Imbalanced host response to SARS-CoV-2 drives development of COVID-19. *Cell* 181, 1036–1045.e9 (2020).
25. H. Di, H. Elbahesh, M. A. Brinton, Characteristics of human OAS1 isoform proteins. *Viruses* 12, (2020).
26. T. A. Tummino, V. V. Rezelj, B. Fischer, A. Fischer, M. J. O’Meara, B. Monel, T. Vallet, Z. Zhang, A. Alon, Henry R O’Donnell, J. Lyu, H. Schadt, K. M. White, N. J. Krogan, L. Urban, K. M. Shokat, A. C. Kruse, A. García-Sastre, O. Schwartz, F. Moretti, M. Vignuzzi, F. Pognan, B. K. Shoichet, Phospholipidosis is a shared mechanism underlying the in vitro antiviral activity of many repurposed drugs against SARS-CoV-2. *bioRxiv* 10.1101/2021.03.23.436648 (2021).
27. P. B. McCray Jr, L. Pewe, C. Wohlford-Lenane, M. Hickey, L. Manzel, L. Shi, J. Netland, H. P. Jia, C. Halabi, C. D. Sigmund, D. K. Meyerholz, P. Kirby, D. C. Look, S. Perlman, Lethal infection of K18-hACE2 mice infected with severe acute respiratory syndrome coronavirus. *J. Virol.* 81, 813–821 (2007).
28. M. A. Haendel, C. G. Chute, T. D. Bennett, D. A. Eichmann, J. Guinney, W. A. Kibbe, P. R. O. Payne, E. R. Pfaff, P. N. Robinson, J. H. Saltz, H. Spratt, C. Suver, J. Wilbanks, A. B. Wilcox, A. E. Williams, C. Wu, C. Blacketer, R. L. Bradford, J. J. Cimino, M. Clark, E. W. Colmenares, P. A. Francis, D. Gabriel, A. Graves, R. Hemadri, S. S. Hong, G. Hripscak, D. Jiao, J. G. Klann, K. Kostka, A. M. Lee, H. P. Lehmann, L. Lingrey, R. T. Miller, M. Morris, S. N. Murphy, K. Natarajan, M. B. Palchuk, U. Sheikh, H. Solbrig, S. Visweswaran, A. Walden, K. M. Walters, G. M.

- Weber, X. T. Zhang, R. L. Zhu, B. Amor, A. T. Girvin, A. Manna, N. Qureshi, M. G. Kurilla, S. G. Michael, L. M. Portilla, J. L. Rutter, C. P. Austin, K. R. Gersing; The N3C Consortium, The National COVID Cohort Collaborative (N3C): Rationale, design, infrastructure, and deployment. *J. Am. Med. Inform. Assoc.* 28, 427–443 (2021).
29. Centers for Disease Control and Prevention, *COVID-19, People at Increased Risk, People with Certain Medical Conditions* (CDC, 2021).
30. Q. Zhou, V. Chen, C. P. Shannon, X.S. Wei, X. Xiang, X. Wang, Z.H. Wang, S. J. Tebbutt, T. R. Kollmann, E. N. Fish, Interferon- α 2b treatment for COVID-19. *Front. Immunol.* 11, 1061 (2020).
31. T. Muthumalage, I. Rahman, Cannabidiol differentially regulates basal and LPS-induced inflammatory responses in macrophages, lung epithelial cells, and fibroblasts. *Toxicol. Appl. Pharmacol.* 382, 114713 (2019).
32. Food and Drug Administration (FDA), Application number 210365Orig1s000, *Clinical Pharmacology and Biopharmaceutics Reviews* (2017).
33. B. C. Sorkin, A. J. Kuszak, G. Bloss, N. K. Fukagawa, F. A. Hoffman, M. Jafari, B. Barrett, P. N. Brown, F. D. Bushman, S. J. Casper, F. H. Chilton, C. S. Coffey, M. G. Ferruzzi, D. C. Hopp, M. Kiely, D. Lakens, J. B. MacMillan, D. O. Meltzer, M. Pahor, J. Paul, K. Pritchett-Corning, S. K. Quinney, B. Rehermann, K. D. R. Setchell, N. S. Sipes, J. M. Stephens, D. L. Taylor, H. Tiriach, M. A. Walters, D. Xi, G. Zappalá, G. F. Pauli, Improving natural product research translation: From source to clinical trial. *FASEB J.* 34, 41–65 (2020).
34. E. Afgan, D. Baker, B. Batut, M. van den Beek, D. Bouvier, M. Čech, J. Chilton, D. Clements, N. Coraor, B. A. Grüning, A. Guerler, J. Hillman-Jackson, S. Hiltemann, V. Jalili, H. Rasche, N. Soranzo, J. Goecks, J. Taylor, A. Nekrutenko, D. Blankenberg, The Galaxy platform for accessible,

reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* 46, W537–W544 (2018).

35. F. Krueger, Trim Galore, A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files. (2015); https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/.
36. A. Dobin, C.A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T.R. Gingeras, STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21 (2013).
37. Y. Liao, G. K. Smyth, W. Shi, featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930 (2014).
38. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550 (2014).
39. J. T. Robinson, H. Thorvaldsdóttir, W. Winckler, M. Guttman, E. S. Lander, G. Getz, J. P. Mesirov, Integrative genomics viewer. *Nat. Biotechnol.* 29, 24–26 (2011).
40. S. B. Yoon, Y.H. Park, S.A. Choi, H.J. Yang, P.S. Jeong, J.J. Cha, S. Lee, S. H. Lee, J.H. Lee, B.W. Sim, B.S. Koo, S.J. Park, Y. Lee, Y.H. Kim, J. J. Hong, J.S. Kim, Y. B. Jin, J.W. Huh, S.R. Lee, B.S. Song, S.U. Kim, Real-time PCR quantification of spliced X-box binding protein 1 (XBP1) using a universal primer method. *PLOS ONE* 14, e0219978 (2019).
41. Y. Zhou, B. Zhou, L. Pache, M. Chang, A. H. Khodabakhshi, O. Tanaseichuk, C. Benner, S. K. Chanda, Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* 10, 1523 (2019).
42. A. Subramanian, P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub, E. S. Lander, J. P. Mesirov, Gene set enrichment analysis: A

- knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* 102, 15545–15550 (2005).
43. C. Hetz, F. Martinon, D. Rodriguez, L. H. Glimcher, The unfolded protein response: Integrating stress signals through the stress sensor IRE1 α . *Physiol. Rev.* 91, 1219–1243 (2011).
44. L. J. Reed, H. Muench, A simple method of estimating 50% end points. *Am. J. Epidemiol.* 27, 493–497 (1938).
45. Observational Health Data Sciences and Informatics, ATLAS (2021); www.ohdsi.org/software-tools/atlas/.
46. Observational Medical Outcomes Partnership (OMOP), OMOP Clinical Data Model v5.3.1, Standardized Derived Elements, Drug Era table (2021); https://ohdsi.github.io/CommonDataModel/cdm53.html#Standardized_Derived_Elements
47. Centers for Disease Control and Prevention, *COVID-19: When to Quarantine* (CDC, 2021).
48. Text Processing Services Library, *Regular Expression Operations* (Python Software Foundation, 2021); <https://docs.python.org/3/library/re.html>.
49. J. J. Beltran-Montoya, T. Herrerias-Canedo, A. Arzola-Paniagua, F Vadillo-Ortega, O. F. Dueñas-Garcia, H. Rico-Olvera, A randomized, clinical trial of ketorolac tromethamine vs ketorolac trometamine plus complex B vitamins for cesarean delivery analgesia. *Saudi J Anaesth* 6, 207–212 (2012).
50. Centers for Disease Control and Prevention, *COVID-19-Hospitalization and Death by Age* (CDC, 2021).
51. Observational Medical Outcomes Partnership (OMOP), OMOP Clinical Data Model v5.3.1, Clinical Data Tables, Person table (2021); <https://ohdsi.github.io/CommonDataModel/cdm53.html>

52. L. Kompaniyets, A.B. Goodman, B. Belay, D.S. Freedman, M.S. Sucusky, S.J. Lange, A.V. Gundlapalli, T.K. Boehmer, H.M. Blanck, Body mass index and risk for COVID-19–related hospitalization, intensive care unit admission, invasive mechanical ventilation, and Death—United States, March–December 2020. *MMWR Morb. Mortal. Wkly Rep.* 70, 355–361 (2021).
53. C. B. Weir, A. Jan, BMI classification percentile and cut off points. In: *StatPearls [Internet]* (StatPearls Publishing, January 2021, updated 9 May 2021); <https://www.ncbi.nlm.nih.gov/books/NBK541070/>.
54. US Food and Drug Administration, FDA approves new indication for drug containing an active ingredient derived from cannabis to treat seizures in rare genetic Disease (FDA news release, 31 July 2020); <https://www.fda.gov/news-events/press-announcements/fda-approves-new-indication-drug-containing-active-ingredient-derived-cannabis-treatseizures-rare>.
55. E. C. Leas, E. M. Hendrickson, A. L. Nobles, R. Todd, D. M. Smith, M. Dredze, J. W. Ayers, Self-reported cannabidiol (CBD) use for conditions with proven therapies. *JAMA Netw. Open* 3, e2020977 (2020).
56. M. Hepburn, N. Mullaguri, P. George, S. Hantus, V. Punia, A. Bhimraj, C. R. Newey, Acute symptomatic seizures in critically ill patients with COVID-19: Is there an association? *Neurocrit. Care* 34, 139–143 (2021).
57. StataCorp, Tabulate two-way reference manual (2021); <https://www.stata.com/manuals/rtabulatetwoway.pdf>.
58. D. Ho, K. Imai, G. King, E. A. Stuart, MatchIt: nonparametric preprocessing for parametric causal inference. *J. Stat. Softw.* 42, 1–28 (2021).

59. P. C. Austin, Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat. Med.* 28, 3083–3107 (2009).
60. D. Pregibon, Data Analytic Methods for Generalized Linear Models. Thesis, University of Toronto (1979).
61. L. Blizzard, D. W. Hosmer, Parameter estimation and goodness-of-fit in log binomial regression. *Biom. J.* **48**, 5–22 (2006).
62. John Hopkins Coronavirus Resource Center, Covid-19 Dashboard (2021); <https://coronavirus.jhu.edu/map.html>.
63. National COVID Cohort Collaborative, *COVID-19 Phenotype 3.0 Additional Information* (N3C, 2021). https://github.com/National-COVID-Cohort-Collaborative/Phenotype_Data_Acquisition/wiki/Phenotype-3.0---Additional-Information.