Plan Overview

A Data Management Plan created using DMPTool-Stage

DMP ID: https://doi.org/10.48321/D12G72

Title: Open Covid Trials

Creator: John Borghi - **ORCID:** 0000-0001-9570-4163

Affiliation: Stanford University (stanford.edu)

Principal Investigator: John Borghi

Data Manager: John Borghi

Contributor: Lily Ren, Christopher Stave, Amanda Woodward, Connie Wong, Cheyenne Payne

Funder: Stanford University (stanford.edu)

Funding opportunity number: n/a

Grant: n/a

Template: Stanford University Template

Project abstract:

The COVID-19 pandemic has brought substantial attention to the systems used to communicate biomedical research. In particular, the need to rapidly and credibly communicate research findings has led many stakeholders to encourage researchers to adopt open science practices such as posting preprints and sharing data. To examine the degree to which this has led to the adoption of such practices, we examined the "openness" of a sample of 539 published papers describing the results of randomized controlled trials testing interventions to prevent or treat COVID-19.

For this study, we defined "openness" quite broadly to include the accessibility of the publications themselves, if they were preceded by preprints, and if they provided information about how to obtain related datasets and other materials.

Start date: 08-31-2021

End date: 08-10-2022

Last modified: 08-07-2023

Copyright information:

Open Covid Trials

Data Types

Over the course of this study, we will collect and synthesize information related to publications describing the results of randomized clinical trials examining pharmaceutical interventions intended to prevent or treat COVID-19.

This information will be drawn from several sources, including the Unpaywall database, PubMed Central, Europe PMC, and manual inspection of the publications themselves. Information from these sources will be combined into a spreadsheet- first a Google sheet that will facilitate our collaborative data collection process and then as a CSV file that will be shared (along with all relevant documentation) publicly via the Dryad data repository.

We expect that the final dataset for this project will be less than a megabyte in total size.

Related Tools, Software, and Code

We will be using R for all of our data analysis and visualization-related activities. Our final dataset will be shared in CSV format, an open (non-proprietary) file format. CSV files can be opened and used by computational tools that are in widespread use and freely available (e.g. Excel, Python, R), so no specialized tools will be needed to access or manipulate our dataset after it is shared.

Standards

Formal metadata standards have not yet been widely adopted for the type of data we will be working with over the course of this study. However, our data and other materials will be structured and described according to best practices.

Data and related documentation will be stored in common and open formats (e.g. CSV, TXT). Information needed to make use of this data (e.g. the meaning of variable names, codes, information about missing data, other metadata etc) will be recorded in a data dictionary that will be accessible to every member of the research team and will subsequently be shared alongside the final dataset. Information about our research process, including the details of our analysis pipeline will be maintained contemporaneously using Protocols.io, an open-access platform for recording, sharing, and reusing study protocols before, during, and after the publication of research results

The Dryad data repository uses the Datacite metadata scheme and mints digital object identifiers (DOIs) to enhance discoverability for shared datasets.

Data Preservation, Access, and Associated Timelines

We will share our final dataset alongside all relevant documentation and code upon publication of a preprint describing our results. These materials will be deposited into the Dryad data repository.

Dryad is an open-source tool for data publication and digital preservation. Datasets deposited into Dryad are permanently archived in a CoreTrustSeal-certified repository. Data files are regularly audited to ensure fixity and authenticity and are replicated with multiple copies in multiple geographic locations. Professional curators examine all Dryad deposits to ensure the validity of the data, apply robust metadata, and make certain that highly sensitive information has not been inadvertently included. Datasets deposited in Dryad are automatically assigned a Digital Object Identifier (DOI) and are indexed by Google Dataset Search and other tools to enhance discoverability.

Access, Distribution, or Reuse Considerations

We do not believe there are any concerns related to privacy, security, informed consent, or proprietary issues that apply to our data. On Dryad, all datasets are assigned a CC0 license, which allows for unrestricted reuse. We will publish our dataset before or alongside the publication of a manuscript describing our results.

Oversight of Data Management and Sharing

Data management-related practices and strategies will be documented in materials available to the entire research team, will be actively monitored in real-time, and will be discussed at regular team meetings. We will update the contents of this DMSP as needed. Compliance with these practices will be managed by the following individuals…

Name	Position	Role
John Borghi, PhD	Project lead	Oversee data management and sharing-related practices and procedures. Update DMSP as needed. Day-to-day monitoring of compliance with data management and sharing practices and procedures.

Planned Research Outputs

Dataset - "Dataset for Open Covid Trials "

We will deposit a dataset containing all of the values that underlie our study's claims in a publicly accessible repository.

Planned research output details

Title	Type	Anticipated release date	Initial access level	Intended repository(ies)	Anticipated file size	License	Metadata standard(s)	May contain sensitive data?	May contain PII?
Dataset for Open Covid Trials	Dataset	2022-08-31	Open	DRYAD	1 MB	Creative Commons Zero v1.0 Universal	None specified	No	No