What is the goal of the RDB?

https://covid19researchdatabase.org
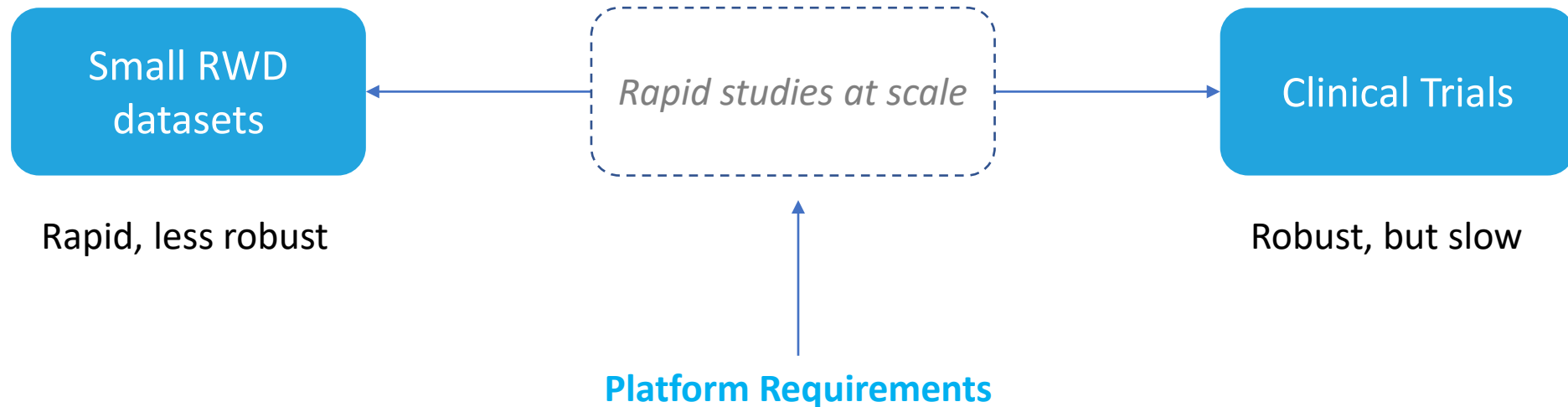
The COVID-19 Research Database seeks to accelerate the amount of real-world data that researchers can freely access to better understand and mitigate the effects of the COVID-19 pandemic.
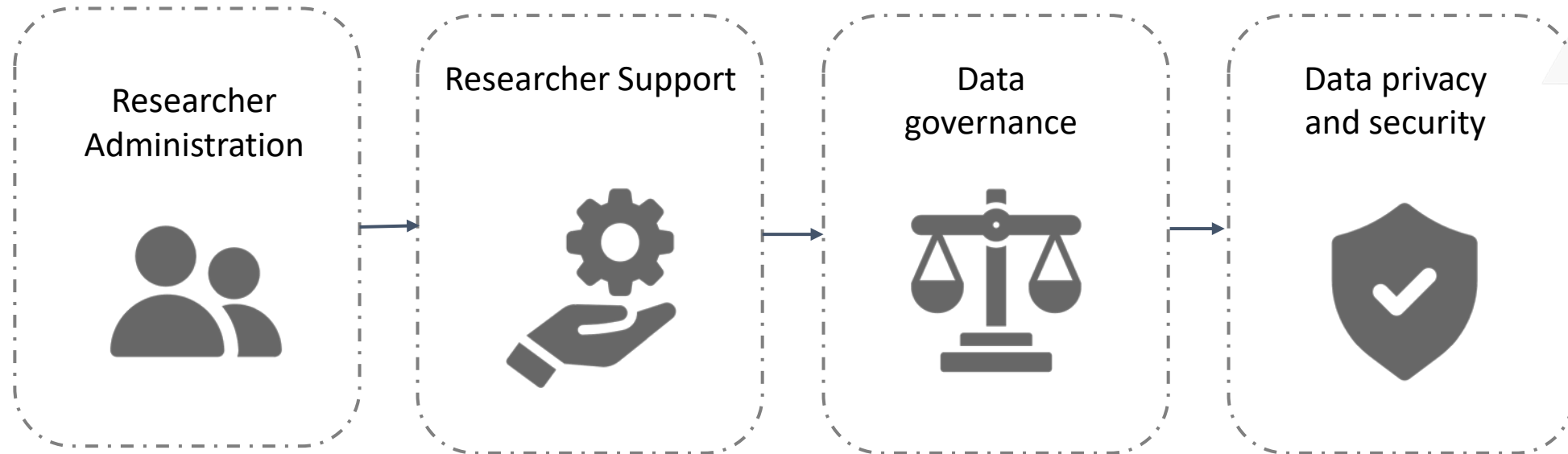
Better insights will enable faster cures, more tailored mitigation plans, and support reopening America.

COVID-19
RESEARCH DATABASE

Need: An Open Platform for Research

Small RWD datasets → *Rapid studies at scale* → Clinical Trials

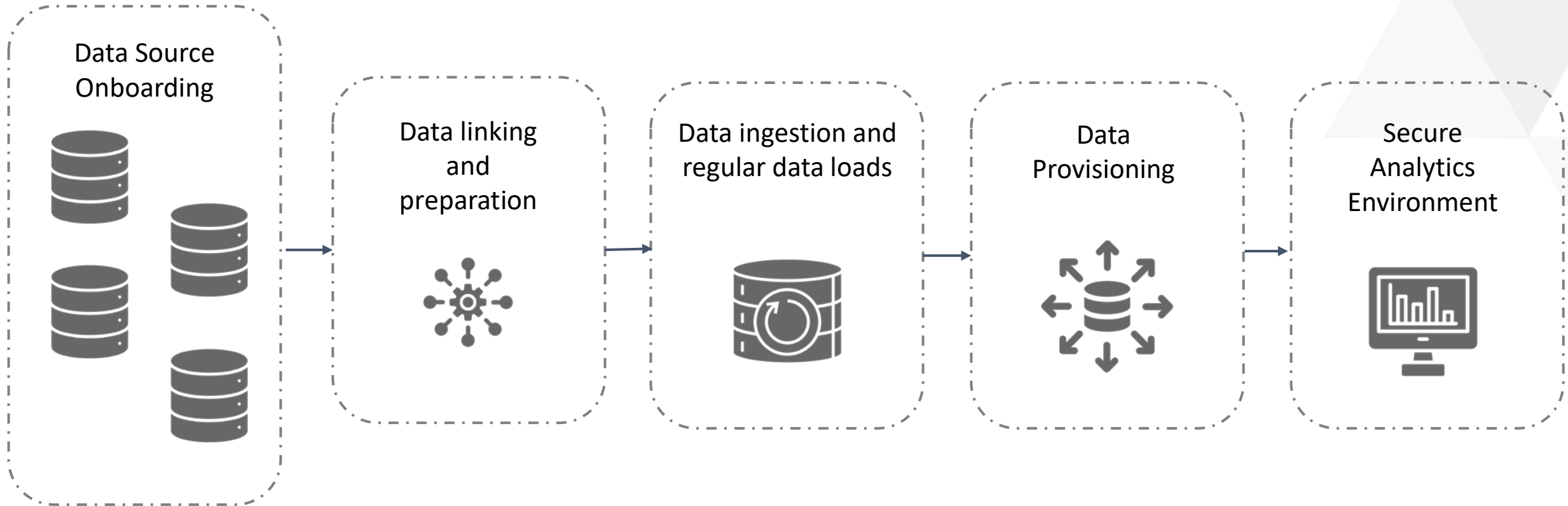Rapid, less robust

Robust, but slow

Platform Requirements

1. Core real-world data sets that can be joined in a privacy-preserving manner

2. Open, scalable technology to host data, and link to any additional real-world data set

3. Robust governance structure to control access to data sets

Governance and Research Management

Researcher Administration → Researcher Support → Data governance → Data privacy and security

Technical Infrastructure

Data Source Onboarding → **Data linking and preparation** → **Data ingestion and regular data loads** → **Data Provisioning** → **Secure Analytics Environment**
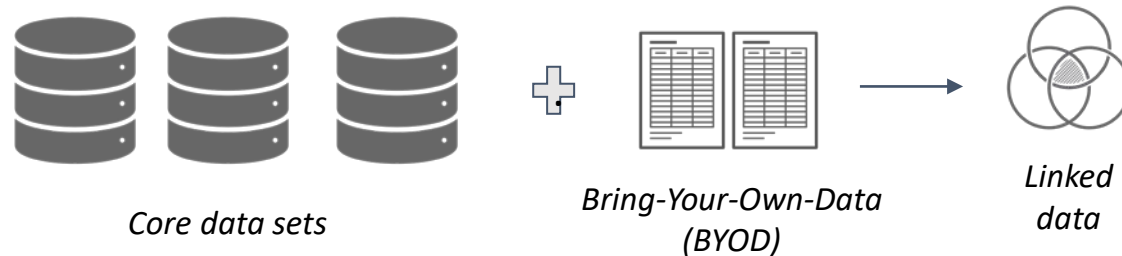
Solution: The Open Research Database Stack

HCCI manages data governance and researcher engagement

MEDIDATA manages data provisioning and the analytics environment

SNOWFLAKE provides Cloud Hosting on AWS

Core data sets
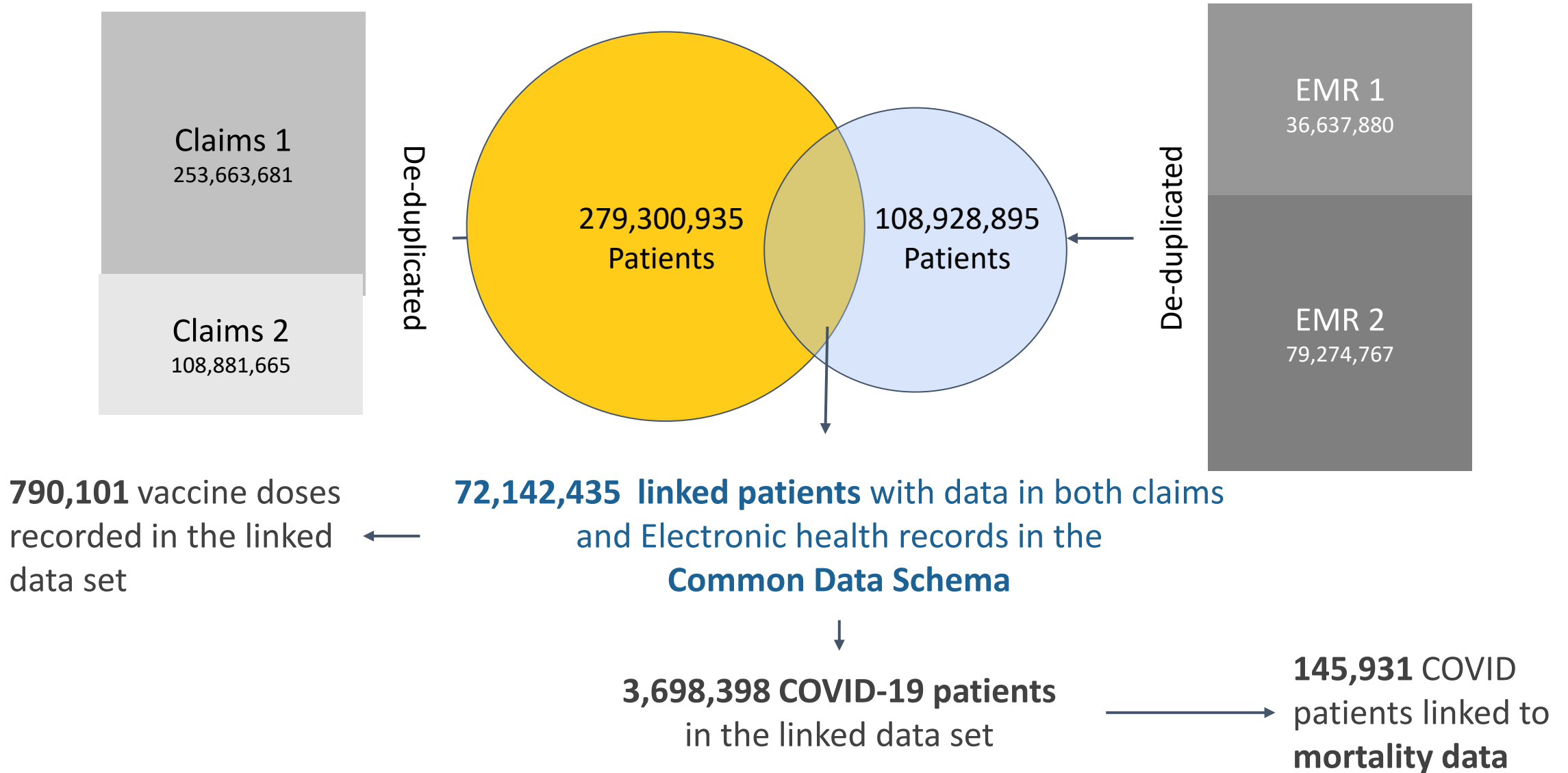
Bring-Your-Own-Data (BYOD)

Linked data

MEDIDATA manages data refreshes and loads

MIRADOR ANALYTICS certifies individual and linked data sets as de-identified

DATAVANT onboards data sources and links datasets

The COVID-19 Research Database currently covers **316,087,395** individuals with claims or electronic records (EHR) data

Claims 1
253,663,681

Claims 2
108,881,665

De-duplicated

279,300,935
Patients

108,928,895
Patients

De-duplicated

EMR 1
36,637,880

EMR 2
79,274,767

790,101 vaccine doses recorded in the linked data set

72,142,435 linked patients with data in both claims and Electronic health records in the **Common Data Schema**

3,698,398 COVID-19 patients in the linked data set

145,931 COVID patients linked to **mortality data**

1 year post launch

>2200

Researchers have registered

>150

Research groups actively using the Research Database for a major project

~400

Researchers have applied for access to datasets as part of research applications

Registered Organizations

Government Organizations, Non-Profits, and Labs: CDC, NIH, FDA, VA, CBO, HHS, American Cancer Society, IHME, Autoimmune Registry, Urban Institute, NASA, EPA, Alzheimer's Association, Government Accountability Office, Pacific Northwest National Lab, Berkeley Lab, Los Alamos National Laboratory, Federal Reserve Bank

State Governments: California, Massachusetts, Michigan, Pennsylvania, Utah, Wisconsin, Virginia

Medical Centers: Mayo Clinic, Cleveland Clinic, and all of the top 30 medical schools according to the 2021 US News Ranking (Harvard, Johns Hopkins, Perelman (Penn), NYU, Stanford, Columbia, Mayo, UCLA, UCSF, WashU, Weill (Cornell), Duke, University of Washington, University of Pittsburgh, University of Michigan, Yale, University of Chicago, Northwestern, Vanderbilt, Icahn School, UCSD, Baylor, University of North Carolina (Chapel Hill), Case Western, Emory, University of Texas Southwestern, University of Wisconsin Madison, Oregon Health and Science University, BU, and UVA)

Countries: 45 different countries (Afghanistan, Algeria, Australia, Austria, Bahrain, Brazil, Bulgaria, Canada, Chile, China, Colombia, Ecuador, Egypt, England, France, Germany, Guatemala, India, Indonesia, Iran, Iraq, Italy, Japan, Kenya, Lebanon, Macedonia, Mexico, Netherlands, Nigeria, North Macedonia, Pakistan, Peru, Philippines, Portugal, Romania, Saudi Arabia, Senegal, Singapore, South Africa, Spain, Switzerland, Taiwan, United Arab Emirates, United Kingdom, United States)

Example Studies Submitted by Researchers

Prediction and Treatment Recommendations for Patients with Severe COVID-19 Infection

THE UNIVERSITY OF **CHICAGO**

Risk of COVID-19 Amongst Lupus and Dermatomyositis Patients Taking Hydroxychloroquine

HARVARD MEDICAL SCHOOL

A Geospatial and Temporal Analysis Comparing COVID-19 Related Outcomes in Patients Routinely Treated with IL-6 inhibitors, JAK inhibitors, or Monoclonal Antibody to Patients Routinely Receiving Hydroxychloroquine.

McMaster University **Cytel**

Surgery During the COVID19 Pandemic: A Unique Opportunity to Define US Baseline Essential Surgical Needs

Stanford University   JOHNS HOPKINS UNIVERSITY

The Prevalence of Hypercoagulopathy in COVID-19 and Potential Mitigation Strategies: A Retrospective Cohort Analysis

The University of Vermont

The Effects of COVID-19 on Health Care Utilization and Mortality

UCSF   Berkeley UNIVERSITY OF CALIFORNIA

The Effects of Health Insurance on COVID-19 Mortality: A Regression Discontinuity Design

Yale University