

## Title

The United States COVID-19 Forecast Hub dataset

## Authors

Estee Y Cramer<sup>1\*</sup>, Yuxin Huang<sup>1\*</sup>, Yijin Wang<sup>1\*</sup>, Evan L Ray<sup>1</sup>, Matthew Cornell<sup>1</sup>, Johannes Bracher<sup>2,3</sup>, Andrea Brennen<sup>4</sup>, Alvaro J Castero Rivadeneira<sup>1</sup>, Aaron Gerding<sup>1</sup>, Katie House<sup>1</sup>, Dasuni Jayawardena<sup>1</sup>, Abdul H Kanji<sup>1</sup>, Ayush Khandelwal<sup>1</sup>, Khoa Le<sup>1</sup>, Jarad Niemi<sup>5</sup>, Ariane Stark<sup>1</sup>, Apurv Shah<sup>1</sup>, Nutch Wattanchit<sup>1</sup>, Martha W Zorn<sup>1</sup>, Nicholas G Reich<sup>1</sup>, on behalf of the US COVID-19 Forecast Hub Consortium\*\*

\*These three authors contributed equally

\*\*See group authorship list as appendix

## Affiliations

1. University of Massachusetts Amherst
2. Chair of Econometrics and Statistics, Karlsruhe Institute of Technology
3. Computational Statistics Group, Heidelberg Institute for Theoretical Studies
4. IQT Labs
5. Iowa State University

corresponding author:

Nicholas G Reich ([nick@umass.edu](mailto:nick@umass.edu))

Disclaimer: The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

## Abstract

Academic researchers, government agencies, industry groups, and individuals have produced forecasts at an unprecedented scale during the COVID-19 pandemic. To leverage these forecasts, the United States Centers for Disease Control and Prevention (CDC) partnered with an academic research lab at the University of Massachusetts Amherst to create the US COVID-19 Forecast Hub. Launched in April 2020, the Forecast Hub is a dataset with point and probabilistic forecasts of incident hospitalizations, incident cases, incident deaths, and cumulative deaths due to COVID-19 at national, state, and county levels in the United States. Included forecasts represent a variety of modeling approaches, data sources, and assumptions regarding the spread of COVID-19. The goal of this dataset is to establish a standardized and comparable set of short-term forecasts from modeling teams. These data can be used to develop ensemble models, communicate forecasts to the public, create visualizations, compare models, and inform policies regarding COVID-19 mitigation.

These open-source data are available via download from GitHub, through an online API, and through R packages.

## Background & Summary

To understand how the COVID-19 pandemic would progress in the United States, dozens of academic research groups, government agencies, industry groups, and individuals produced probabilistic forecasts for COVID-19 outcomes starting in March 2020.<sup>1</sup> We have collected forecasts from over 82 modeling teams in a data repository, thus making forecasts easily accessible for COVID-19 response efforts and forecast evaluation. The data repository is called the United States (US) COVID-19 Forecast Hub (hereafter, Forecast Hub) and was created through a partnership between the US Centers for Disease Control and Prevention (CDC) and an academic research lab at the University of Massachusetts Amherst.

The Forecast Hub was launched in early April 2020 and contains real-time forecasts of reported COVID-19 cases, hospitalizations, and deaths. As of September 8, 2021, the Forecast Hub had collected nearly 65 million individual point or quantile predictions contained within over 4,600 submitted forecast files from over 100 unique models. The forecasts submitted each week reflected a variety of forecasting approaches, data sources, and underlying assumptions. There were no restrictions in place regarding the underlying information or code used to generate real-time forecasts. Each week, the latest forecasts were combined into an ensemble forecast (Figure 1) and all recent forecast data were updated on an official COVID-19 Forecasting page hosted by the US CDC.<sup>2</sup> The ensemble models were also used in the weekly reports that are posted on the Forecast Hub website (<https://covid19forecasthub.org/doc/reports/>).

Forecasts are quantitative predictions about future observations. Forecasts differ from scenario-based projections, which examine feasible outcomes conditional on a variety of future assumptions. Because forecasts are unconditional estimates of future observations, they can be evaluated. An important feature of the Forecast Hub is that submitted forecasts are time-stamped so that the exact time at which a forecast was made public can be verified. In this way, the Forecast Hub serves as a public, independent registration system for these forecast model outputs. Data from the Forecast Hub have served as the basis for research articles for forecast evaluation<sup>3</sup> and forecast combination.<sup>4-6</sup> These studies can be used to determine how well models have performed at various points during the pandemic, which can, in turn, guide best practices for utilizing forecasts in practice and inform future forecasting efforts.<sup>3</sup>

Any modeling team was eligible to submit forecast data to the Forecast Hub, provided they submitted data in the correct format. Teams submitted predictions in a structured format to facilitate data validation, storage, and analysis. Teams also submitted a

metadata file and license for their model's data. Forecast data, ground truth data from the Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE),<sup>7</sup> New York Times (NYTimes),<sup>8</sup> and USA Facts,<sup>9</sup> as well as model metadata were stored in the public Forecast Hub GitHub repository.<sup>10</sup>

The forecasts were automatically synchronized with an online database called Zoltr via calls to a REpresentational State Transfer (REST) application programming interface (API)<sup>11</sup> every six hours (Figure 2). Forecast data may be downloaded directly from GitHub, via the *covidHubUtils* R package,<sup>12</sup> the *zoltr* R package<sup>13</sup> or *zoltpy* python library.<sup>14</sup>

This dataset of real-time forecasts created during the COVID-19 pandemic can provide insights into the shortcomings and successes of predictions and improve forecasting efforts in years to come. Though these data are restricted to forecasts for COVID-19 in the United States, the structure of this dataset has been used to create datasets of COVID-19 forecasts in the EU and the UK, and longer-term scenario projections in the US.<sup>15–18</sup> The general structure of this data collection could be applied to additional diseases or forecasting outcome in the future.<sup>11</sup>

This large collaborative effort has provided data on short-term forecasts for over a year of forecasting efforts. These data were collected in real-time and therefore are not subject to retrospective biases. The data are also openly available to the public, thus fostering a transparent, open science approach to support public health efforts.

## Methods

### *Data Acquisition*

Beginning in April of 2020, the Reich Lab at the University of Massachusetts, Amherst, in partnership with the CDC, began collecting probabilistic forecasts of key COVID-19 outcomes in the United States (Table 1). The effort began by collecting forecasts of deaths and hospitalizations at the weekly and daily scale for the 50 US states, Washington DC, and 4 territories (Puerto Rico, US Virgin Islands, Guam, and the Northern Mariana Islands) as well as the aggregated US national level. In July 2020, the effort expanded to include forecasts of weekly incident cases at the county, state, and national levels. Forecasts may include a point prediction and/or quantiles of a predictive distribution.

Any team was eligible to submit data to the Forecast Hub. Upon initial submission of forecast data, teams were required to upload a metadata file that briefly described the

methods used to create the forecasts and specified a license under which their forecast data were released. No model code was stored by the Forecast Hub.

During the first month of operation, members of the Forecast Hub team downloaded forecasts made available by teams publicly online, transformed these into the correct format, and pushed them into the Forecast Hub repository. Starting in May 2020, all teams were required to format and submit their own forecasts.

### *Repository structure*

The dataset is stored in two locations, and all data can be accessed through either source. The first is the COVID-19 Forecast Hub GitHub repository and the second is an online database, Zoltar, which can be accessed via a REST API.<sup>11</sup> Details about data format and access are documented in the subsequent sections.

### *Zoltar: data backend*

The data can be accessed through the Zoltar forecast repository REST API. Through the API, subsets of submitted forecasts can be queried directly from a PostgreSQL database. This eliminates the need to access individual CSV files and facilitates access to versions of forecasts in cases when they are updated.

### *Outcomes and locations*

The Forecast Hub dataset stores forecasts for four different outcomes: incident hospitalizations, incident cases, incident deaths, and cumulative deaths (Table 1). Incident hospitalizations can be submitted for a horizons of 1 - 130 days in the future, incident cases can be submitted for 1 - 8 weeks in the future, and incident and cumulative deaths can be submitted for 1 - 20 weeks into the future. For all outcomes, forecasts can be submitted on a national and state level. Incident case forecasts were first introduced as a forecast outcome several months after the Hub started and have several key differences with other predicted outcomes. They are the only outcome for which the Hub accepts county-level forecasts in addition to the state and national level. Because there are over 3,000 counties in the US, this required some compromises on the scale of data collected for these forecasts in other ways. Specifically, case forecasts are required to have fewer quantiles (seven quantiles) compared to other outcomes which can have up to twenty-three quantiles. This gives a coarser representation of the forecast (see the section on Forecast format below).

Weekly targets follow the standard of epidemiological weeks (EW) used by the CDC, which defines a week to start on a Sunday and end on the following Saturday.<sup>19</sup>

Forecasts of cumulative deaths target the number of cumulative deaths reported by the Saturday ending a given week. Forecasts of weekly incident cases or deaths target the difference between reported cumulative cases or deaths on consecutive Saturdays. As an example of a forecast and the corresponding observation, forecasts submitted between Tuesday, October 6, 2020 (day 3 of EW41) and Monday, October 12, 2020 (day 2 of EW42) contained a “1 week ahead” forecast of incident deaths that corresponded to the change in cumulative reported deaths observed in EW42 (i.e., the difference between the cumulative reported deaths on Saturday, October 17, 2020, and Saturday, October 10, 2020), a “2 week ahead” forecast that corresponded to the change in cumulative reported deaths in week EW43. In this paper, we refer to the “forecast week” of a submitted forecast as the week corresponding to a “0-week ahead” horizon. In the example above, the forecast week would be EW41. Daily incident hospitalization horizons are for the number of reported hospitalizations a specified number of days after the forecast was generated.

### *Forecast assumptions*

Forecasters used a variety of assumptions to build models and generate predictions. Forecasting approaches include statistical or machine learning models, mechanistic models incorporating disease transmission dynamics, and combinations of multiple approaches.<sup>3</sup> Teams have also included varying assumptions regarding future changes in policies and physical distancing measures, the transmissibility of COVID-19, vaccination rates, and the spread of new virus variants throughout the United States.

### *Weekly submissions*

A forecast submission consists of a single comma-separated value (CSV) file submitted via pull request to the GitHub repository. Forecast submissions are validated for technical accuracy and formatting (see exclusion criteria below) before being merged. To be included in the weekly ensemble model, teams were required to submit their forecast on Sunday or prior to a deadline on Monday. The majority of teams contributing to the dataset submitted forecasts to the Hub repository on Sunday or Monday, although some teams submitted at other times depending on their model production schedule.

### *Model designation*

Each model stored in the repository must have a classification of “primary”, “secondary” or “other”. Each team must only have one “primary” model. Teams submitting multiple models with similar forecasting approaches can use the designations “secondary” or “other” for their models. Models with the designation “primary” are included in

evaluations, the weekly ensemble, and the visualization. The “secondary” label is designed for models that have a substantive methodological difference than a team’s “primary” model. Models with the designation “secondary” are included only in the weekly ensemble and the visualization. The “other” label is designed for models that are small variations on a team’s “primary” model. Models with the designation “other” are not included in evaluations, the ensemble build, or the visualization.

### *Ensemble and baseline forecasts*

Several models have a special status, either as a baseline or as an ensemble that combines multiple models from the Hub to create a single forecast.

The COVIDhub-baseline model was created by the Hub in May 2020 as a benchmarking model. Its point forecast is the most recent observed value as of the forecast creation date with a probability distribution around that based on weekly differences in previous observations.<sup>3</sup> The baseline model initially produced forecasts for case and death outcomes. Hospitalization baseline forecasts were added in September 2021.

The COVIDhub-ensemble model creates a combination of submitted forecasts to the Hub. Other work details the methods used for determining the appropriate combination approach.<sup>4,5</sup> Starting in February 2021, GitHub tags were created to document the exact version of the repository used each week to create the COVIDhub-ensemble forecast. This creates an auditable trail in the repository so the correct version of the used forecasts could be recovered even in cases when some forecasts were subsequently updated.

Several other models also are combinations of some or all models submitted to the Forecast Hub. As of August 1, 2021, these models are COVIDhub-trained\_ensemble, FDANIHASU-Sweight, JHUAPL-SLPHospEns, and KITmetricslab-select\_ensemble. These models are flagged in the metadata using the Boolean metadata field, “ensemble\_of\_hub\_models”.

### *Exclusion criteria*

No forecasts were excluded from the dataset due to the forecast values or the background experience of the forecasters. Forecast files were only rejected if they did not meet the automatic formatting criteria implemented through automatic GitHub checks.<sup>20</sup> These included checks to ensure that, among other criteria:

- A forecast file is submitted no more than 2 days after it has been created (to ensure forecasts submitted were truly prospective). The creation date is based on the date in the filename created by the submitting team.
- The forecast dates in the content of the file are in the format YYYY-MM-DD and match the creation date.
- Quantile forecasts do not contain any quantiles at probability levels other than those required (see Forecast Format section below).

### *Updates to files*

To ensure that forecasting is done in real-time, all forecasts are required to be submitted to the Hub within 2 days of the forecast date, which is listed in a column within each forecast file. Though occasional late submissions were accepted up through January of 2021, the policy was updated to not accept late forecasts due to missed deadlines, updated modeling methods, or other reasons.

Exceptions to this policy were made if there were programming or data errors that affected the forecasts in the original submission or if a new team joined. If there was an error, teams were required to submit a comment with their updated submission affirming that there was a bug and that the forecast was only produced using data that were available at the time of the original submission. In the case of updates to forecast data, both the old and updated versions of the forecasts can be accessed either through the GitHub commit history or through time-stamped queries of the forecasts in the Zoltar database. Note that an updated forecast can include “retracting” a particular set of predictions in the case when an initial forecast was not able to be updated. When new teams join the Hub, they can submit late forecasts if they can provide publicly available evidence that the forecasts were made in real-time (e.g. GitHub commit history).

### *Ground truth data*

Data from the JHU CSSE dataset<sup>21</sup> are used as the ground truth data for reported cases and deaths. Data from the HealthData.gov system for state-level hospitalizations are used for the hospitalization outcome. JHU CSSE obtained counts of cases and deaths by collecting and aggregating reports from state and local health departments. HealthData.gov contains reports of hospitalizations assembled by the U.S. Department of Health and Human Services. Teams were encouraged to use these sources to build models. Although hospitalization forecasts were collected starting in March 2020, the hospitalizations data from HealthData.gov were only available later, and we started encouraging teams to target these data in November 2020. Some teams used alternate data sources including the NYTimes, USAFacts, US Census data, and other signals.<sup>3</sup> Versions of truth data from JHU CSSE, USAFacts, and the NYTimes are stored in the GitHub repository.

Previous reports of ground truth data for past time points were occasionally updated as new records became available, definitions of reportable cases, deaths, or hospitalizations changed, or errors in data collection were identified and corrected. These revisions to the data are sometimes quite substantial, and for some purposes such as retrospective ensemble construction, it is necessary to use the data that would have been available in real-time. The historically versioned data can be accessed either through GitHub commit records, data versions released on HealthData.gov, or third-party tools such as the covidcast API provided by the Delphi group at Carnegie Mellon University or the *covidData* R package.<sup>22</sup>

## Data Records

### *Summary of forecast data collected*

In the initial weeks of submission, there were fewer than 10 models providing forecasts. As the pandemic spread, the number of teams submitting forecasts increased to 82; as of July 2021, 82 primary, 4 secondary models and 15 models with the designation “other” had been submitted to the Forecast Hub. In the first six months of 2021, a median of 35.5 teams (range: 30 to 38) contributed incident case forecasts (Fig 3a), a median of 12 teams (range: 9 to 14) contributed incident hospitalizations (Fig 3b), a median of 43 teams (range 37 to 49) contributed incident death forecasts (Fig 3c), and a median of 44 teams (range 34 to 46) contributed cumulative death forecasts (Fig 3d). As of September 8 2021, the dataset contained 4,602 forecast files with 64,902,239 point or quantile predictions for unique combinations of targets and locations.

### *GitHub repository data structure*

Forecasts in the GitHub repository are available in subfolders organized by model. Folders are named with a team name and model name, and each folder includes a metadata file and forecast files. Forecast CSV files are named using the format “<YYYY-MM-DD>-<team abbreviation>-<model abbreviation>.csv”. In these files, each row contains data for a single outcome, location, horizon, and point or quantile prediction as described above.

The metadata file for each team, named using the format “metadata-<team abbreviation>-<model abbreviation>.txt”, contains relevant information about the team and the model that the team is using to generate forecasts.



### *Forecast format*

Forecasts were required to be submitted in the format of point predictions and/or quantile predictions. Point predictions represented single “best” predictions with no uncertainty, typically representing a mean or median prediction from the model. Quantile predictions are an efficient format for storing predictive distributions of a wide range of outcomes.

Quantile representations of predictive distributions lend themselves to natural computations of, for example, pinball loss or a weighted interval score, both strictly proper scoring rules that can be used to evaluate forecasts<sup>23</sup> However, they do not capture the structure of the tails of the predictive distribution beyond the reported quantiles. Also, the quantile format does not preserve any information on correlation structures between different outcomes.

### *Variable descriptions*

The forecast data in this dataset are stored in seven columns:

1. **forecast\_date** - the date the forecast was made in the format YYYY-MM-DD.
2. **target** - a character string giving the number of days/weeks ahead that are being forecasted (horizon) and the outcome. Horizons must be one of the following:
  - a. “N wk ahead cum death” where N is a number between 1 and 20
  - b. “N wk ahead inc death” where N is a number between 1 and 20
  - c. “N wk ahead inc case” where N is a number between 1 and 8
  - d. “N day ahead inc hosp” where N is a number between 0 and 130
3. **target\_end\_date** - a character string representing the date for the forecast target in the format YYYY-MM-DD. For “k day-ahead” targets, target\_end\_date will be k days after forecast\_date. For “k week ahead” targets, target\_end\_date will be the Saturday at the end of the specified epidemic week, as described above.
4. **location** - character string of Federal Information Processing Standard Publication (FIPS) codes identifying U.S. states, counties, territories, and districts as well as “US” for national forecasts. The values for the FIPS codes are available in a CSV file in the repository and as a data object in the *covidHubUtils* R package for convenience.
5. **type** - character value of “point” or “quantile” indicating whether the row corresponds to a point forecast or a quantile forecast.
6. **quantile** - the probability level for a quantile forecast. For death and hospitalization forecasts, forecasters can submit quantiles at 23 probability levels: 0.01, 0.025, 0.05, 0.10, 0.15, ..., 0.95, 0.975, 0.99. For cases, teams can submit up to 7 quantiles at levels .025, 0.100, 0.250, 0.5, 0.750, 0.900 and 0.975.

If the forecast “type” is equal to “point”, the value in the quantile column is equal to “NA”.

7. **value** - non-negative numbers indicating the “point” or “quantile” prediction for the row. For a “point” prediction, value is simply the value of that point prediction for the target and location associated with that row. For a “quantile” prediction, the model predicts that the eventual observation will be less than this value with the probability given by the quantile probability level.

### *Metadata format*

Each team documents their model information in a metadata file which is required along with the first forecast submission. Each team is asked to record their model’s design and assumptions, the model contributors, the team’s website, information regarding the team’s data sources, and a brief model description. Teams may update their metadata file periodically to keep track of minor changes to a model.

### *Variable descriptions*

A standard metadata file should be a YAML file with the following required fields in a specific order:

1. **team\_name** - the name of the team (less than 50 characters).
2. **model\_name** - the name of the model (less than 50 characters).
3. **model\_abbr** - an abbreviated and uniquely identified name for the model that is less than 30 alphanumeric characters. The model abbreviation must be in the format of `[team\_abbr]-[model\_abbr]` where each of the `[team\_abbr]` and `[model\_abbr]` are text strings that are each less than 15 alphanumeric characters that do not include a hyphen or whitespace.
4. **model\_contributors** - a list of all individuals involved in the forecasting effort, affiliations, and email addresses. At least one contributor needs to have a valid email address. The syntax of this field should be name1 (affiliation1) <user@address>, name2 (affiliation2) <user2@address2>
5. **website\_url\*** - a URL to a website that has additional data about the model. We encourage teams to submit the most user-friendly version of the model, e.g. a dashboard, or similar, that displays the model forecasts. If there is an additional data repository where forecasts and other model code are stored, this can be included in the methods section. If only a more technical site, e.g. GitHub repo, exists that link should be included here.
6. **license** - one of the acceptable license types in the Forecast Hub. We encourage teams to submit as a "cc-by-4.0" to allow the broadest possible use, including private vaccine production (which would be excluded by the "cc-by-nc-4.0")

license). If the value is "LICENSE.txt", then a LICENSE.txt file must exist within the model folder and provide a license.

7. **team\_model\_designation** - upon initial submission this field should be one of "primary", "secondary" or "other".
8. **methods** - a brief description of the forecasting methodology that is less than 200 Characters.
9. **ensemble\_of\_hub\_models** - a Boolean value (`true` or `false`) that indicates whether a model combines multiple hub models into an ensemble.

\*previously named **model\_output**

Teams are also encouraged to add model information with optional fields described in Supplement 1.

## Technical Validation

Two similar but distinct validation processes were used to validate data on the GitHub repository and on Zoltar.

### *GitHub repository*

Validations were set up using GitHub Actions to manage continuous integration and automated data checking.<sup>20</sup> Teams submitted their metadata files and forecasts through pull requests on GitHub. Each time a new pull request was submitted, a validation script ran on all new or updated files in the pull request to test for their validity. Separate checks ran on metadata file changes and forecast data file changes.

The metadata file for each team was required to be in valid YAML format, and a set of specific checks were required before a new metadata file could be merged into the repository. Checks included ensuring that the proposed team and model names do not conflict with existing names, that a valid license for data reuse is specified, and that a valid model designation was present. A list of specific validations for metadata may be found in Supplement 2.

New or changed forecast data files for each team were required to pass a series of checks for data formatting and validity. These checks also ensured that the forecast data files did not meet any of the exclusion criteria (see the Methods section for specific rules). Furthermore, a list of specific validations for forecast data files is provided in Supplement 2.

## *Zoltar*

When a new forecast file is uploaded to Zoltar, unit tests are run on the file to make sure that forecast elements contain valid structure. (For a detailed specification of the structure of forecast elements, see <https://docs.zoltardata.com/validation/>.) If a forecast file does not pass all unit tests, the upload will fail and the forecast file will not be added to the database; only when all tests pass will the new forecast be added to Zoltar. The validations in place on GitHub ensure that only valid forecasts will be uploaded to Zoltar.

## *Observed data*

Raw observed data from multiple sources including JHU, NYTimes, USAFacts, and Healthdata.gov is downloaded and reformatted using the scripts in the R packages *covidHubUtils* (<https://github.com/reichlab/covidHubUtils>) and *covidData* (<https://github.com/reichlab/covidData>). This data generating process is automated by GitHub Actions every week and the results (called “truth data”) are directly uploaded to the Forecast Hub repository and Zoltar. In specific, case and death raw observed data are aggregated to a weekly level and all three outcomes (cases, deaths, and hospitalization) are reformatted for use within the Hub.

## **Usage Notes**

We have developed the *covidHubUtils* R package (<https://github.com/reichlab/covidHubUtils>) to facilitate bulk retrieval of forecasts for analysis and evaluation. Examples of how to use the *covidHubUtils* package and its functions can be found at <https://reichlab.io/covidHubUtils/>. The package supports loading forecasts from a local clone of the GitHub repository or by querying data from Zoltar. The package supports common actions for working with the data, such as loading in specific subsets of forecasts, plotting forecasts, scoring forecasts, retrieving ground truth data, and many other utility functions to simplify working with the data.

## *Communicating results from the COVID-19 Forecast Hub*

Communication of probabilistic forecasts to the public is challenging,<sup>24,25</sup> and the best practices regarding the communication of outbreaks are still developing.<sup>26</sup> Starting in April 2020, the CDC published weekly summaries of these forecasts on their public website<sup>27</sup>, and these forecasts were occasionally used in public briefings by the CDC director.<sup>28</sup> Additional examples of the communication of Forecast Hub data can be viewed through weekly reports generated by the Hub team for dissemination to the general public, including state and local departments of health. (<https://covid19forecasthub.org/doc/reports/>)

## Acknowledgements

This work has been supported in part by the US Centers for Disease Control and Prevention (1U01IP001122) and the National Institutes of General Medical Sciences (R35GM119582). The content is solely the responsibility of the authors and does not necessarily represent the official views of CDC, FDA, NIGMS or the National Institutes of Health. Johannes Bracher was supported by the Helmholtz Foundation via the SIMCARD Information & Data Science Pilot Project. Tilmann Gneiting gratefully acknowledges support by the Klaus Tschira Foundation.

## Funding

For teams that reported receiving funding for their work, we report the sources and disclosures below.

Alpert-pwllnod: Natural Sciences and Engineering Research Council of Canada

Caltech-CS156: Gary Clinard Innovation Fund

CEID-Walk: University of Georgia

CMU-TimeSeries: CDC Center of Excellence, gifts from Google and Facebook

COVIDhub: This work has been supported by the US Centers for Disease Control and Prevention (1U01IP001122) and the National Institutes of General Medical Sciences (R35GM119582). The content is solely the responsibility of the authors and does not necessarily represent the official views of CDC, NIGMS or the National Institutes of Health. Johannes Bracher was supported by the Helmholtz Foundation via the SIMCARD Information & Data Science Pilot Project. Tilmann Gneiting gratefully acknowledges support by the Klaus Tschira Foundation.

CU-select: NSF DMS-2027369 and a gift from the Morris-Singer Foundation.

DDS-NBDS: NSF III-1812699

epiforecasts-ensemble1: Wellcome Trust (210758/Z/18/Z)

FDANIHASU: supported by the Intramural Research Program of the NIH/NIDDK

GT\_CHHS-COVID19: William W. George Endowment, Virginia C. and Joseph C. Mello Endowment, NSF DGE-1650044, NSF MRI 1828187, research cyberinfrastructure resources and services provided by the Partnership for an Advanced Computing Environment (PACE) at Georgia Tech, and the following benefactors at Georgia Tech: Andrea Laliberte, Joseph C. Mello, Richard "Rick" E. & Charlene Zalesky, and Claudia & Paul Raines, CDC MInD-Healthcare U01CK000531-Supplement.

IHME: This work was supported by the Bill & Melinda Gates Foundation, as well as funding from the state of Washington and the National Science Foundation (award no. FAIN: 2031096)

Imperial-ensemble1: SB acknowledges funding from the Wellcome Trust (219415).

Institute of Business Forecasting: IBF

IowaStateLW-STEM: NSF DMS-1916204, Iowa State University Plant Sciences Institute Scholars Program, NSF DMS-1934884, Laurence H. Baker Center for Bioinformatics and Biological Statistics.

IUPUI CIS: NSF

JHU CSSE-DECOM: JHU CSSE: National Science Foundation (NSF) RAPID “Real-time Forecasting of COVID-19 risk in the USA”. 2021-2022. Award ID: 2108526. National Science Foundation (NSF) RAPID “Development of an interactive web-based dashboard to track COVID-19 in real-time”. 2020. Award ID: 2028604

JHU IDD-CovidSP: State of California, US Dept of Health and Human Services, US Dept of Homeland Security, Johns Hopkins Health System, Office of the Dean at Johns Hopkins Bloomberg School of Public Health, Johns Hopkins University Modeling and Policy Hub, Centers for Disease Control and Prevention (5U01CK000538-03), University of Utah Immunology, Inflammation, & Infectious Disease Initiative (26798 Seed Grant).

JHU UNC GAS-StatMechPool: NIH NIGMS: R01GM140564

JHUAPL-Bucky: US Dept of Health and Human Services

KITmetricslab-select\_ensemble: Daniel Wolffram gratefully acknowledges support by the Klaus Tschira Foundation.

LANL-GrowthRate: LANL LDRD 20200700ER

MIT-Cassandra: MIT Quest for Intelligence

MOBS-GLEAM COVID: COVID Supplement CDC-HHS-6U01IP001137-01; CA NU38OT000297 from the Council of State and Territorial Epidemiologists (CSTE)

NotreDame-FRED: NSF RAPID DEB 2027718

NotreDame-mobility: NSF RAPID DEB 2027718

PSI-DRAFT: NSF RAPID Grant # 2031536

QJHong-Encounter: NSF DMR-2001411 and DMR-1835939

SDSC ISG-TrendModel: The development of the dashboard was partly funded by the Fondation Privée des Hôpitaux Universitaires de Genève

UA-EpiCovDA: NSF RAPID Grant # 2028401

UChicagoCHATTOPADHYAY-UNIT: Defense Advanced Research Projects Agency (DARPA) #HR00111890043/P00004 (I. Chattopadhyay, University of Chicago).

UCSB-ACTS: NSF RAPID IIS 2029626

UCSD NEU-DeepGLEAM: Google Faculty Award, W31P4Q-21-C-0014

UMass-MechBayes: NIGMS #R35GM119582, NSF #1749854, NIGMS #R35GM119582

UMich-RidgeTfReg: This project is funded by the University of Michigan Physics Department and the University of Michigan Office of Research.

UVA-Ensemble: National Institutes of Health (NIH) Grant 1R01GM109718, NSF BIG DATA Grant IIS-1633028, NSF Grant No.: OAC-1916805, NSF Expeditions in Computing Grant CCF-1918656, CCF-1917819, NSF RAPID CNS-2028004, NSF RAPID OAC-2027541, US Centers for Disease Control and Prevention 75D30119C05935, a grant from Google, University of Virginia Strategic Investment Fund award number SIF160, Defense Threat Reduction Agency (DTRA) under Contract No. HDTRA1-19-D-0007, and Virginia Dept of Health Grant VDH-21-501-0141

Wadnwan AI-BayesOpt: This study is made possible by the generous support of the American People through the United States Agency for International Development (USAID). The work described in this article was implemented under the TRACETB Project, managed by WIAI under the terms of Cooperative Agreement Number 72038620CA00006. The contents of this manuscript are the sole responsibility of the authors and do not necessarily reflect the views of USAID or the United States Government.

WalmartLabsML-LogForecasting: Team acknowledges Walmart to support this study

## Author Consortium

Estee Y Cramer<sup>1</sup>, Yuxin Huang<sup>1</sup>, Yijin Wang<sup>1</sup>, Evan L Ray<sup>1</sup>, Matthew Cornell<sup>1</sup>, Johannes Bracher<sup>2,3</sup>, Andrea Brennen<sup>4</sup>, Alvaro J Castro Rivadeneira<sup>1</sup>, Aaron Gerding<sup>1</sup>, Katie House<sup>1</sup>, Dasuni Jayawardena<sup>1</sup>, Abdul H Kanji<sup>1</sup>, Ayush Khandelwal<sup>1</sup>, Khoa Le<sup>1</sup>, Jarad Niemi<sup>5</sup>, Ariane Stark<sup>1</sup>, Apurv Shah<sup>1</sup>, Nutch Wattanachit<sup>1</sup>, Martha W Zorn<sup>1</sup>, Tilmann Gneiting<sup>2</sup>, Anja Mühlemann<sup>6</sup>, Youyang Gu<sup>7</sup>, Yixian Chen<sup>8</sup>, Krishna Chintanippu<sup>8</sup>, Viresh Jivane<sup>8</sup>, Ankita Khurana<sup>8</sup>, Ajay Kumar<sup>8</sup>, Anshul Lakhani<sup>8</sup>, Prakhar Mehrotra<sup>8</sup>, Sujitha Pasumarty<sup>8</sup>, Monika Shrivastav<sup>8</sup>, Jialu You<sup>8</sup>, Nayana Bannur<sup>9</sup>, Ayush Deva<sup>9</sup>, Sansiddh Jain<sup>9</sup>, Mihir Kulkarni<sup>9</sup>, Srujana Merugu<sup>9</sup>, Alpan Raval<sup>9</sup>, Siddhant Shingi<sup>9</sup>, Avtansh Tiwari<sup>9</sup>, Jerome White<sup>9</sup>, Aniruddha Adiga<sup>10</sup>, Benjamin Hurt<sup>10</sup>, Bryan Lewis<sup>10</sup>, Madhav Marathe<sup>10</sup>, Akhil Sai Peddireddy<sup>10</sup>, Przemyslaw Porebski<sup>10</sup>, Srinivasan Venkatramanan<sup>10</sup>, Lijing Wang<sup>10</sup>, Maytal Dahan<sup>11</sup>, Spencer Fox<sup>12</sup>, Kelly Gaither<sup>11</sup>, Michael Lachmann<sup>13</sup>, Lauren Ancel Meyers<sup>12</sup>, James G Scott<sup>12</sup>, Mauricio Tec<sup>12</sup>, Spencer Woody<sup>12</sup>, Ajitesh Srivastava<sup>14</sup>, Tianjian Xu<sup>14</sup>, Jeffrey C Cegan<sup>15</sup>, Ian D Dettwiller<sup>15</sup>, William P England<sup>15</sup>, Matthew W Farthing<sup>15</sup>, Glover E George<sup>15</sup>, Robert H Hunter<sup>15</sup>, Brandon Lafferty<sup>15</sup>, Igor Linkov<sup>15</sup>, Michael L Mayo<sup>15</sup>, Matthew D Parno<sup>15</sup>, Michael A Rowland<sup>15</sup>, Benjamin D Trump<sup>15</sup>, Samuel Chen<sup>16</sup>, Stephen V Faraone<sup>16</sup>, Jonathan Hess<sup>16</sup>, Christopher P Morley<sup>16</sup>, Asif Salekin<sup>17</sup>, Dongliang Wang<sup>16</sup>, Yanli Zhang-James<sup>16</sup>, Thomas M Baer<sup>18</sup>, Sabrina M Corsetti<sup>19</sup>, Marisa C Eisenberg<sup>19</sup>, Karl Falb<sup>19</sup>, Yitao Huang<sup>19</sup>, Emily T Martin<sup>19</sup>, Ella McCauley<sup>19</sup>, Robert L Myers<sup>19</sup>, Tom Schwarz<sup>19</sup>, Graham Casey Gibson<sup>1</sup>, Daniel Sheldon<sup>1</sup>, Liyao Gao<sup>20</sup>, Yian Ma<sup>21</sup>, Dongxia Wu<sup>21</sup>, Rose Yu<sup>22,21</sup>, Xiaoyong Jin<sup>23</sup>, Yu-Xiang Wang<sup>23</sup>, Xifeng Yan<sup>23</sup>, YangQuan Chen<sup>24</sup>, Lihong Guo<sup>25</sup>, Yanting Zhao<sup>26</sup>, Jinghui Chen<sup>27</sup>, Quanquan Gu<sup>27</sup>, Lingxiao Wang<sup>27</sup>, Pan Xu<sup>27</sup>, Weitong Zhang<sup>27</sup>, Difan Zou<sup>27</sup>, Ishanu Chattopadhyay<sup>28</sup>, Yi Huang<sup>28</sup>, Guoqing Lu<sup>29</sup>, Ruth Pfeiffer<sup>30</sup>, Timothy Sumner<sup>31</sup>, Liqiang Wang<sup>31</sup>, Dongdong Wang<sup>31</sup>, Shunpu Zhang<sup>31</sup>, Zihang Zou<sup>31</sup>, Hannah Biegel<sup>32</sup>, Joceline Lega<sup>32</sup>, Fazle Hussain<sup>33</sup>, Zeina Khan<sup>33</sup>, Frank Van Bussel<sup>33</sup>, Steve McConnell<sup>34</sup>, Stephanie L Guertin<sup>35</sup>, Christopher Hulme-Lowe<sup>35</sup>, VP Nagraj<sup>35</sup>, Stephen D Turner<sup>35</sup>, Benjamín Bejar<sup>36</sup>, Christine Choirat<sup>36</sup>, Antoine Flahault<sup>37</sup>, Ekaterina Krymova<sup>36</sup>, Gavin Lee<sup>36</sup>, Elisa Manetti<sup>37</sup>, Kristen Namigai<sup>37</sup>, Guillaume Obozinski<sup>36</sup>, Tao Sun<sup>36</sup>, Dorina Thanou<sup>38</sup>, Xuegang Ban<sup>20</sup>, Yunfeng Shi<sup>39</sup>, Robert Walraven<sup>7</sup>, Qi-Jun Hong<sup>40,41</sup>, Axel van de Walle<sup>41</sup>, Michal Ben-Nun<sup>42</sup>, Steven Riley<sup>43</sup>, Pete Riley<sup>42</sup>, James A Turtle<sup>42</sup>, Duy Cao<sup>44</sup>, Joseph Galasso<sup>44</sup>, Jae H Cho<sup>7</sup>, Areum Jo<sup>7</sup>, David DesRoches<sup>45</sup>, Pedro Forli<sup>45</sup>, Bruce Hamory<sup>45</sup>, Ugur Koyluoglu<sup>45</sup>, Christina Kyriakides<sup>45</sup>, Helen Leis<sup>45</sup>, John Milliken<sup>45</sup>, Michael Moloney<sup>45</sup>, James Morgan<sup>45</sup>, Ninad Nirgudkar<sup>45</sup>, Gokce Ozcan<sup>45</sup>, Noah Pivonka<sup>45</sup>, Matt Ravi<sup>45</sup>, Chris Schrader<sup>45</sup>, Elizabeth Shakhnovich<sup>45</sup>, Daniel Siegel<sup>45</sup>, Ryan Spatz<sup>45</sup>, Chris Stiefeling<sup>45</sup>, Barrie Wilkinson<sup>45</sup>, Alexander Wong<sup>45</sup>, Sean Cavany<sup>46</sup>, Guido España<sup>46</sup>, Sean Moore<sup>46</sup>, Rachel Oidtman<sup>28,46</sup>, Alex Perkins<sup>46</sup>, Andrea Kraus<sup>47</sup>, David Kraus<sup>47</sup>, Jiang Bian<sup>48</sup>, Wei Cao<sup>48</sup>, Zhifeng Gao<sup>48</sup>, Juan Lavista Ferres<sup>48</sup>, Chaozhuo Li<sup>48</sup>, Tie-Yan Liu<sup>48</sup>, Xing Xie<sup>48</sup>, Shun Zhang<sup>48</sup>, Shun Zheng<sup>48</sup>, Matteo Chinazzi<sup>49</sup>,

Alessandro Vespignani<sup>50,49</sup>, Xinyue Xiong<sup>49</sup>, Jessica T Davis<sup>49</sup>, Kunpeng Mu<sup>49</sup>, Ana Pastore y Piontti<sup>49</sup>, Jackie Baek<sup>51</sup>, Vivek Farias<sup>52</sup>, Andreea Georgescu<sup>51</sup>, Retsef Levi<sup>52</sup>, Deeksha Sinha<sup>51</sup>, Joshua Wilde<sup>51</sup>, Andrew Zheng<sup>51</sup>, Amine Bennouna<sup>51</sup>, David Nze Ndong<sup>52</sup>, Georgia Perakis<sup>53</sup>, Divya Singhvi<sup>54</sup>, Ioannis Spantidakis<sup>51</sup>, Leann Thayaparan<sup>51</sup>, Asterios Tsiourvas<sup>51</sup>, Shane Weisberg<sup>51</sup>, Ali Jadbabaie<sup>55</sup>, Arnab Sarker<sup>55</sup>, Devavrat Shah<sup>55</sup>, Leo A Celi<sup>56</sup>, Nicolas D Penna<sup>56</sup>, Saketh Sundar<sup>57</sup>, Russ Wolfinger<sup>58</sup>, Lauren Castro<sup>59</sup>, Geoffrey Fairchild<sup>59</sup>, Isaac Michaud<sup>59</sup>, Dave Osthus<sup>59</sup>, Daniel Wolfram<sup>2,3</sup>, Dean Karlen<sup>60,61</sup>, Mark J Panaggio<sup>62</sup>, Matt Kinsey<sup>62</sup>, Luke C. Mullany<sup>62</sup>, Kaitlin Rainwater-Lovett<sup>62</sup>, Lauren Shin<sup>62</sup>, Katharine Tallaksen<sup>62</sup>, Shelby Wilson<sup>62</sup>, Michael Brenner<sup>63,64</sup>, Marc Coram<sup>64</sup>, Jessie K Edwards<sup>65</sup>, Keya Joshi<sup>66</sup>, Ellen Klein<sup>64</sup>, Juan Dent Hulse<sup>67</sup>, Kyra H Grantz<sup>67</sup>, Alison L Hill<sup>68</sup>, Joshua Kaminsky<sup>67</sup>, Kathryn Kaminsky<sup>7</sup>, Lindsay T Keegan<sup>69</sup>, Stephen A Lauer<sup>67</sup>, Elizabeth C Lee<sup>67</sup>, Joseph C Lemaitre<sup>70</sup>, Justin Lessler<sup>71</sup>, Hannah R Meredith<sup>67</sup>, Javier Perez-Saez<sup>67</sup>, Sam Shah<sup>7</sup>, Claire P Smith<sup>67</sup>, Shaun A Truelove<sup>67</sup>, Josh Wills<sup>7</sup>, Lauren Gardner<sup>68</sup>, Maximilian Marshall<sup>68</sup>, Kristen Nixon<sup>68</sup>, John C. Burant<sup>7</sup>, Wen-Hao Chiang<sup>72</sup>, George Mohler<sup>72</sup>, Junyi Gao<sup>73</sup>, Lucas Glass<sup>74</sup>, Cheng Qian<sup>74</sup>, Justin Romberg<sup>75</sup>, Rakshith Sharma<sup>74</sup>, Jeffrey Spaeder<sup>76</sup>, Jimeng Sun<sup>73</sup>, Cao Xiao<sup>77</sup>, Lei Gao<sup>6</sup>, Zhiling Gu<sup>6</sup>, Myungjin Kim<sup>6</sup>, Xinyi Li<sup>78</sup>, Guannan Wang<sup>79</sup>, Lily Wang<sup>6</sup>, Yueying Wang<sup>6</sup>, Shan Yu<sup>10</sup>, Chaman Jain<sup>80</sup>, Sangeeta Bhatia<sup>81</sup>, Pierre Nouvellet<sup>81,82</sup>, Ryan Barber<sup>20</sup>, Emmanuela Gaikedu<sup>20</sup>, Simon Hay<sup>20</sup>, Steve Lim<sup>20</sup>, Chris Murray<sup>20</sup>, David Pigott<sup>20</sup>, Robert C Reiner<sup>20</sup>, Prasith Baccam<sup>83</sup>, Heidi L Gurung<sup>83</sup>, Steven A Stage<sup>83</sup>, Bradley T Suchoski<sup>83</sup>, Chung-Yan Fong<sup>84</sup>, Dit-Yan Yeung<sup>84</sup>, Bijaya Adhikari<sup>85</sup>, Jiaming Cui<sup>75</sup>, B. Aditya Prakash<sup>75</sup>, Alexander Rodriguez<sup>75</sup>, Anika Tabassum<sup>75,86</sup>, Jiajia Xie<sup>75</sup>, John Asplund<sup>87</sup>, Arden Baxter<sup>88</sup>, Pinar Keskinocak<sup>88</sup>, Buse Eylul Oruc<sup>88</sup>, Nicoleta Serban<sup>88</sup>, Sercan O Arik<sup>89</sup>, Mike Dusenberry<sup>89</sup>, Arkady Epshteyn<sup>89</sup>, Elli Kanal<sup>89</sup>, Long T Le<sup>89</sup>, Chun-Liang Li<sup>89</sup>, Tomas Pfister<sup>89</sup>, Rajarishi Sinha<sup>89</sup>, Thomas Tsai<sup>90</sup>, Nate Yoder<sup>89</sup>, Jinsung Yoon<sup>89</sup>, Leyou Zhang<sup>89</sup>, Daniel Wilson<sup>91</sup>, Artur A Belov<sup>92</sup>, Carson C Chow<sup>93</sup>, Richard C Gerkin<sup>40</sup>, Osman N Yogurtcu<sup>92</sup>, Mark Ibrahim<sup>94</sup>, Timothee Lacroix<sup>94</sup>, Matthew Le<sup>94</sup>, Jason Liao<sup>95</sup>, Maximilian Nickel<sup>94</sup>, Levent Sagun<sup>94</sup>, Sam Abbott<sup>96</sup>, Nikos I Bosse<sup>96</sup>, Sebastian Funk<sup>96</sup>, Joel Hellewell<sup>96</sup>, Sophie R Meakin<sup>96</sup>, Katharine Sherratt<sup>96</sup>, Rahi Kalantari<sup>97</sup>, Mingyuan Zhou<sup>97</sup>, Sen Pei<sup>98</sup>, Jeffrey Shaman<sup>98</sup>, Teresa K Yamana<sup>98</sup>, Omar Skali Lami<sup>51</sup>, Dimitris Bertsimas<sup>52</sup>, Michael L Li<sup>51</sup>, Saksham Soni<sup>51</sup>, Hamza Tazi Bouardi<sup>51</sup>, Madeline Adee<sup>99</sup>, Turgay Ayer<sup>100,88</sup>, Jagpreet Chhatwal<sup>101</sup>, Ozden O Dalgic<sup>102</sup>, Mary A Ladd<sup>99</sup>, Benjamin P Linas<sup>103</sup>, Peter Mueller<sup>99</sup>, Jade Xiao<sup>88</sup>, Qinxia Wang<sup>98</sup>, Yuanjia Wang<sup>98</sup>, Shanghong Xie<sup>98</sup>, Donglin Zeng<sup>104</sup>, Jacob Bien<sup>14</sup>, Logan Brooks<sup>105</sup>, Alden Green<sup>105</sup>, Addison J Hu<sup>105</sup>, Maria Jahja<sup>105</sup>, Daniel McDonald<sup>106</sup>, Balasubramanian Narasimhan<sup>107</sup>, Collin Politsch<sup>105</sup>, Samyak Rajanala<sup>107</sup>, Aaron Rumack<sup>105</sup>, Noah Simon<sup>20</sup>, Ryan J Tibshirani<sup>105</sup>, Rob Tibshirani<sup>107</sup>, Valerie Ventura<sup>105</sup>, Larry Wasserman<sup>105</sup>, John M Drake<sup>108</sup>, Eamon B O'Dea<sup>108</sup>, Yaser Abu-Mostafa<sup>109</sup>, Rahil Bathwal<sup>109</sup>, Nicholas A Chang<sup>109</sup>, Pavan Chitta<sup>109</sup>, Anne Erickson<sup>109</sup>, Sumit Goel<sup>109</sup>, Jethin Gowda<sup>109</sup>, Qixuan Jin<sup>109</sup>, HyeongChan Jo<sup>109</sup>, Juhyun Kim<sup>109</sup>, Pranav Kulkarni<sup>109</sup>, Samuel M Lushtak<sup>109</sup>, Ethan Mann<sup>109</sup>, Max Popken<sup>109</sup>, Connor Soohoo<sup>109</sup>, Kushal Tirumala<sup>109</sup>, Albert Tseng<sup>109</sup>, Vignesh Varadarajan<sup>109</sup>, Jagath Vytheeswaran<sup>109</sup>, Christopher Wang<sup>109</sup>, Akshay Yeluri<sup>109</sup>, Dominic Yurk<sup>109</sup>, Michael Zhang<sup>109</sup>, Alexander Zlokapa<sup>109</sup>, Robert Pagano<sup>110</sup>, Chandini Jain<sup>111</sup>, Vishal Tomar<sup>111</sup>, Lam Ho<sup>112</sup>, Huong Huynh<sup>113,114</sup>, Quoc Tran<sup>113,115</sup>, Velma K Lopez<sup>116</sup>, Jo W Walker<sup>116</sup>, Rachel B Slayton<sup>116</sup>, Michael A Johansson<sup>116</sup>, Matthew Biggerstaff<sup>116</sup>, Nicholas G Reich<sup>1</sup>

<sup>1</sup>University of Massachusetts Amherst

<sup>2</sup>Chair of Econometrics and Statistics, Karlsruhe Institute of Technology



- <sup>3</sup>Computational Statistics Group, Heidelberg Institute for Theoretical Studies
- <sup>4</sup>IQT Labs
- <sup>5</sup>Iowa State University
- <sup>6</sup>Institute of Mathematical Statistics and Actuarial Science, University of Bern
- <sup>7</sup>Unaffiliated
- <sup>8</sup>Walmart
- <sup>9</sup>Wadhvani Institute of Artificial Intelligence
- <sup>10</sup>University of Virginia
- <sup>11</sup>Texas Advanced Computing Center
- <sup>12</sup>University of Texas at Austin
- <sup>13</sup>Santa Fe Institute
- <sup>14</sup>University of Southern California
- <sup>15</sup>US Army Engineer Research and Development Center
- <sup>16</sup>State University of New York Upstate Medical University
- <sup>17</sup>Syracuse University
- <sup>18</sup>Trinity University, San Antonio
- <sup>19</sup>University of Michigan - Ann Arbor
- <sup>20</sup>University of Washington
- <sup>21</sup>University of California, San Diego
- <sup>22</sup>Northeastern University
- <sup>23</sup>University of California at Santa Barbara
- <sup>24</sup>University of California, Merced
- <sup>25</sup>Jilin University
- <sup>26</sup>University of Science and Technology of China
- <sup>27</sup>University of California, Los Angeles
- <sup>28</sup>University of Chicago
- <sup>29</sup>University of Nebraska Omaha
- <sup>30</sup>National Cancer Institute (NCI), NIH
- <sup>31</sup>University of Central Florida
- <sup>32</sup>University of Arizona
- <sup>33</sup>Texas Tech University
- <sup>34</sup>Construx
- <sup>35</sup>Signature Science, LLC
- <sup>36</sup>Swiss Data Science Center, EPFL & ETHZ
- <sup>37</sup>Institute of Global Health, Faculty of Medicine, University of Geneva
- <sup>38</sup>Center for Intelligent Systems, EPFL
- <sup>39</sup>Rensselaer Polytechnic Institute
- <sup>40</sup>Arizona State University
- <sup>41</sup>Brown University
- <sup>42</sup>Predictive Science, Inc
- <sup>43</sup>Imperial College, London
- <sup>44</sup>University of Dallas
- <sup>45</sup>Oliver Wyman
- <sup>46</sup>University of Notre Dame
- <sup>47</sup>Masaryk University
- <sup>48</sup>Microsoft
- <sup>49</sup>Laboratory for the Modeling of Biological and Socio-technical Systems, Northeastern University
- <sup>50</sup>ISI Foundation
- <sup>51</sup>Operations Research Center, Massachusetts Institute of Technology
- <sup>52</sup>Sloan School of Management, Massachusetts Institute of Technology

- 53 Sloan School of Management and Operations Research Center, Massachusetts Institute of Technology
- 54 Leonard N Stern School of Business, New York University
- 55 Institute for Data, Systems, and Society, Massachusetts Institute of Technology
- 56 Laboratory for Computational Physiology, Massachusetts Institute of Technology
- 57 River Hill High School
- 58 SAS Institute Inc
- 59 Los Alamos National Laboratory
- 60 TRIUMF
- 61 University of Victoria
- 62 Johns Hopkins University Applied Physics Lab
- 63 School of Engineering and Applied Sciences, Harvard University
- 64 Google Research
- 65 Department of Epidemiology, UNC Gillings School of Public Health, University of North Carolina at Chapel Hill
- 66 Harvard TH Chan School of Public Health
- 67 Johns Hopkins Bloomberg School of Public Health
- 68 Johns Hopkins University
- 69 University of Utah
- 70 École Polytechnique Fédérale de Lausanne
- 71 Department of Epidemiology, University of North Carolina Gillings School of Global Public Health
- 72 Indiana University–Purdue University Indianapolis
- 73 University of Illinois at Urbana-Champaign
- 74 Analytics Center of Excellence, IQVIA
- 75 Georgia Institute of Technology
- 76 IQVIA
- 77 Amplitude
- 78 Clemson University
- 79 College of William & Mary
- 80 Institute of Business Forecasting
- 81 Imperial College London
- 82 School of Life Sciences, University of Sussex
- 83 IEM, Inc.
- 84 The Hong Kong University of Science and Technology
- 85 University of Iowa
- 86 Virginia Tech
- 87 Metron, Inc.
- 88 Georgia Institute of Technology
- 89 Google Cloud
- 90 Harvard University
- 91 Federal Reserve Bank of San Francisco
- 92 Food and Drug Administration, Center for Biologics Evaluation and Research
- 93 NIH
- 94 Facebook AI Research
- 95 Facebook
- 96 London School of Hygiene & Tropical Medicine
- 97 The University of Texas at Austin
- 98 Columbia University
- 99 Massachusetts General Hospital
- 100 Emory University Medical School

<sup>101</sup>Massachusetts General Hospital, Harvard Medical School

<sup>102</sup>Value Analytics Labs

<sup>103</sup>Boston University School of Medicine

<sup>104</sup>UNC Chapel Hill

<sup>105</sup>Carnegie Mellon University

<sup>106</sup>University of British Columbia

<sup>107</sup>Stanford University

<sup>108</sup>University of Georgia

<sup>109</sup>California Institute of Technology

<sup>110</sup>No affiliation

<sup>111</sup>Auquan

<sup>112</sup>Dalhousie University

<sup>113</sup>Alpert

<sup>114</sup>Virtual Power System

<sup>115</sup>Walmart Inc.

<sup>116</sup>Centers for Disease Control and Prevention

## References

1. Haghani, M. & Bliemer, M. C. J. Covid-19 pandemic and the unprecedented mobilisation of scholarly efforts prompted by a health crisis: Scientometric comparisons across SARS, MERS and 2019-nCoV literature. *Scientometrics* **125**, 2695–2726 (2020).
2. CDC. COVID-19 Mathematical Modeling. <https://www.cdc.gov/coronavirus/2019-ncov/covid-data/mathematical-modeling.html> (2021).
3. Cramer, E. Y. *et al.* Evaluation of individual and ensemble probabilistic forecasts of COVID-19 mortality in the US. *medRxiv* 2021.02.03.21250974 (2021) doi:10.1101/2021.02.03.21250974.
4. Brooks, L. C. *et al.* Comparing ensemble approaches for short-term probabilistic COVID-19 forecasts in the U.S. *International Institute of Forecasters* (2020).
5. Ray, E. L. *et al.* Challenges in training ensembles to forecast COVID-19 cases and deaths in the United States. *International Institute of Forecasters* (2021).
6. Taylor, J. W. & Taylor, K. S. Combining Probabilistic Forecasts of COVID-19 Mortality in the United States. *Eur. J. Oper. Res.* (2021) doi:10.1016/j.ejor.2021.06.044.

7. CSSEGISandData/COVID-19. *GitHub* <https://github.com/CSSEGISandData/COVID-19>.
8. *covid-19-data*, <https://github.com/CSSEGISandData/COVID-19>. (Github).
9. US COVID-19 cases and deaths by state. <https://usafacts.org/visualizations/coronavirus-covid-19-spread-map/> (2021).
10. Cramer, E. *et al.* *reichlab/covid19-forecast-hub: release for Zenodo, 20210816*. (2021). doi:10.5281/zenodo.5208210.
11. Reich, N. G., Cornell, M., Ray, E. L., House, K. & Le, K. The Zoltar forecast archive, a tool to standardize and store interdisciplinary prediction research. *Sci Data* **8**, 59 (2021).
12. Wang, S. Y. *et al.* *reichlab/covidHubUtils: repository release for Zenodo*. (2021). doi:10.5281/zenodo.5207940.
13. Cornell, M., Gruson, H., Wang, S. Y. & Ray, E. *reichlab/zoltr: Release for Zenodo, 20210816*. (2021). doi:10.5281/zenodo.5207856.
14. Cornell, M. *et al.* *reichlab/zoltpy: Release for Zenodo, 20210816*. (2021). doi:10.5281/zenodo.5207932.
15. *covid19-scenario-modeling-hub*, <https://github.com/midas-network/covid19-scenario-modeling-hub>. (Github).
16. *covid19-forecast-hub-europe: European Covid-19 Forecast Hub*, <https://github.com/epiforecasts/covid19-forecast-hub-europe>. (Github).
17. *covid19-forecast-hub-de: German and Polish COVID-19 Forecast Hub*, <https://github.com/KITmetricslab/covid19-forecast-hub-de>. (Github).
18. Borchering, R. K. *et al.* Modeling of Future COVID-19 Cases, Hospitalizations, and Deaths, by Vaccination Rates and Nonpharmaceutical Intervention Scenarios - United States, April-September 2021. *MMWR Morb. Mortal. Wkly. Rep.* **70**, 719–724 (2021).
19. MMWR Weeks. CDC [https://wwwn.cdc.gov/nndss/document/MMWR\\_Week\\_overview.pdf](https://wwwn.cdc.gov/nndss/document/MMWR_Week_overview.pdf).
20. Hannan, A., Huang, Y. D. & Wang, S. Y. *reichlab/covid19-forecast-hub-validations: Release for Zenodo, 20210816*. (2021). doi:10.5281/zenodo.5207934.

21. Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.* **20**, 533–534 (2020).
22. Ray, E. *et al. reichlab/covidData: repository release for Zenodo.* (2021).  
doi:10.5281/zenodo.5208224.
23. Bracher, J., Ray, E. L., Gneiting, T. & Reich, N. G. Evaluating epidemic forecasts in an interval format. *PLoS Comput. Biol.* **17**, e1008618 (2021).
24. Gigerenzer, G., Hertwig, R., van den Broek, E., Fasolo, B. & Katsikopoulos, K. V. ‘A 30% chance of rain tomorrow’: how does the public understand probabilistic weather forecasts? *Risk Anal.* **25**, 623–629 (2005).
25. Raftery, A. E. Use and Communication of Probabilistic Forecasts. *Stat. Anal. Data Min.* **9**, 397–410 (2016).
26. Risk-Communication-and-Behavior-Best-Practices-and-Research-Findings-July-2016.pdf.
27. CDC. COVID-19 Forecasts: Deaths. <https://www.cdc.gov/coronavirus/2019-ncov/covid-data/forecasting-us.html> (2021).
28. Waldrop, T., Andone, D. & Holcombe, M. CDC warns new Covid-19 variants could accelerate spread in US. *CNN* (2021).

## Tables and Figures

Figure 1: Time series of weekly incident deaths at the national level and forecasts from the COVID-19 Forecast Hub ensemble model for selected weeks in 2020 and 2021. Ensemble forecasts (blue) with 50%, 80% and 95% prediction intervals shown in shaded regions, and the ground-truth data (black) for incident cases (A), incident hospitalizations (B), incident deaths (C) and cumulative deaths (D). The truth data come from JHU CSSE (panels A, C, D) and HealthData.gov (panel B).

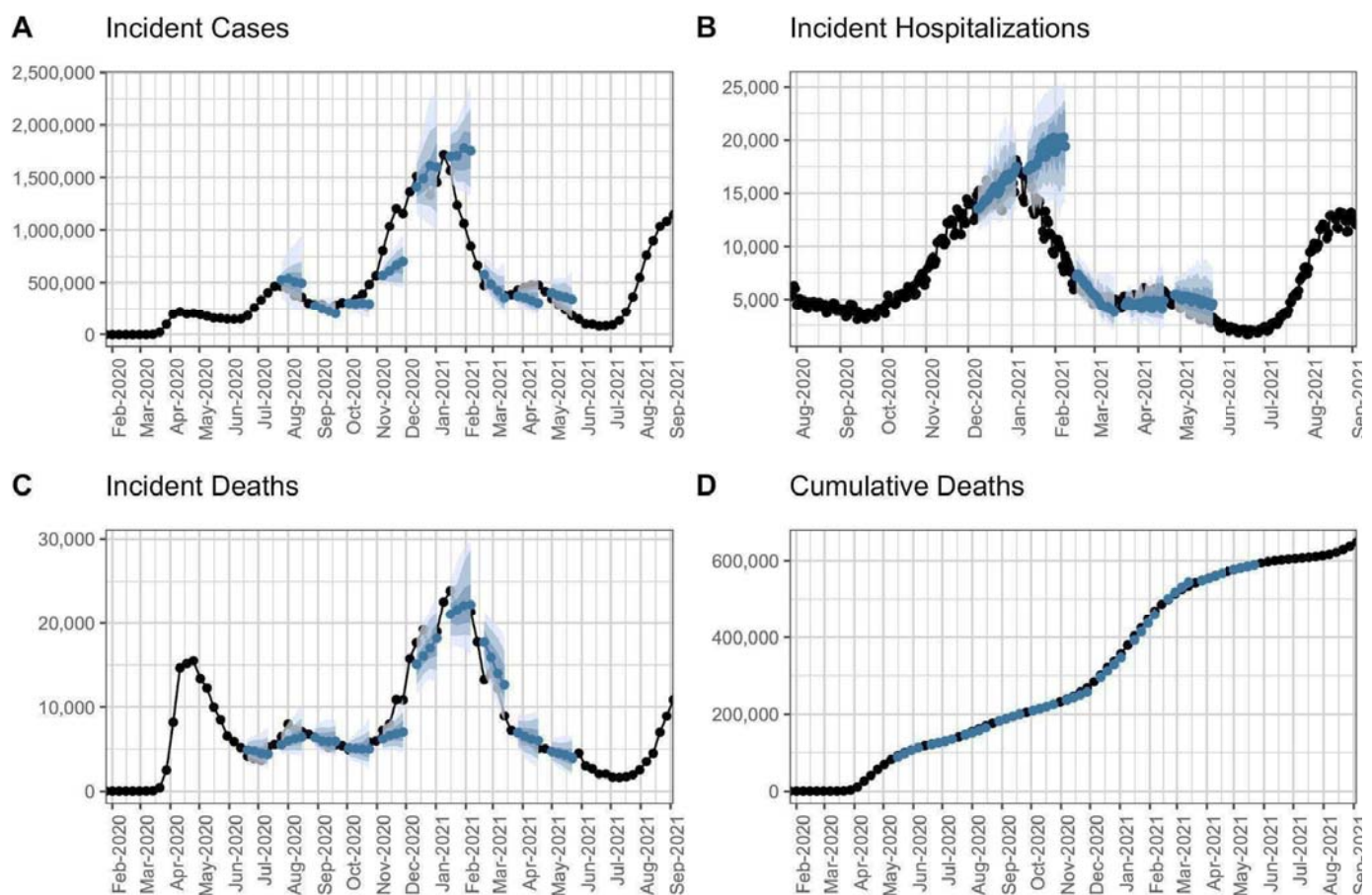


Figure 2: Schematic of the data storage and related infrastructure surrounding the COVID-19 Forecast Hub. (A) Forecasts are submitted to the COVID-19 Forecast Hub GitHub repository and undergo data format validation before being accepted into the system. (B) A continuous integration service ensures that the GitHub repository and PostgreSQL database stay in sync with essentially mirrored versions of the data. (C) Truth data for visualization, evaluation, and ensemble building are retrieved once per week using both the *covidHubUtils* and the *covidData* R packages. Truth data are stored in both repositories. (D) Once per week, an ensemble forecast submission is made using the *covidEnsembles* R package. It is submitted to the GitHub repository and undergoes the same validation as other submissions. (E) Using the *covidHubUtils* R package, forecast and truth data may be extracted from either the GitHub or PostgreSQL database in a standard format for tasks such as scoring or plotting.

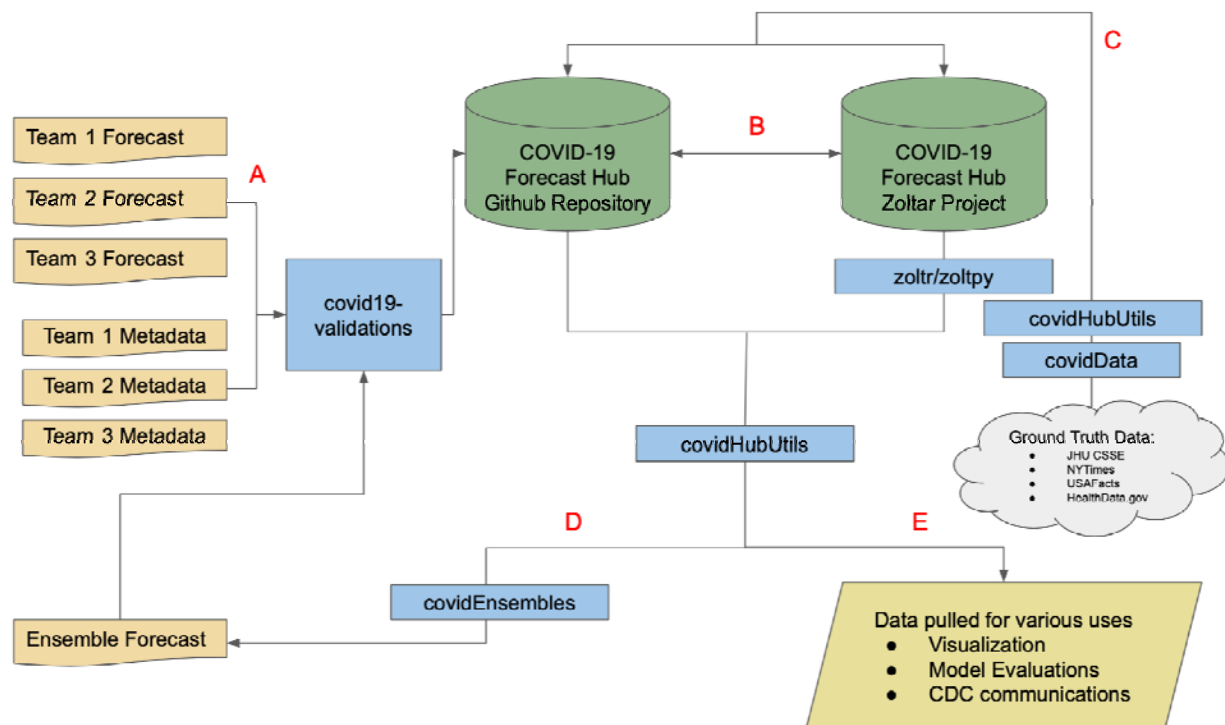


Figure 3: Number of primary forecasts submitted for each outcome per week from April 27th 2020 through July 31st 2021. In the initial weeks of submission, fewer than 10 models provided forecasts. Over time, the number of teams submitting forecasts for each forecasted outcome increased into early 2021 and then saw a small decline through the summer of 2021.

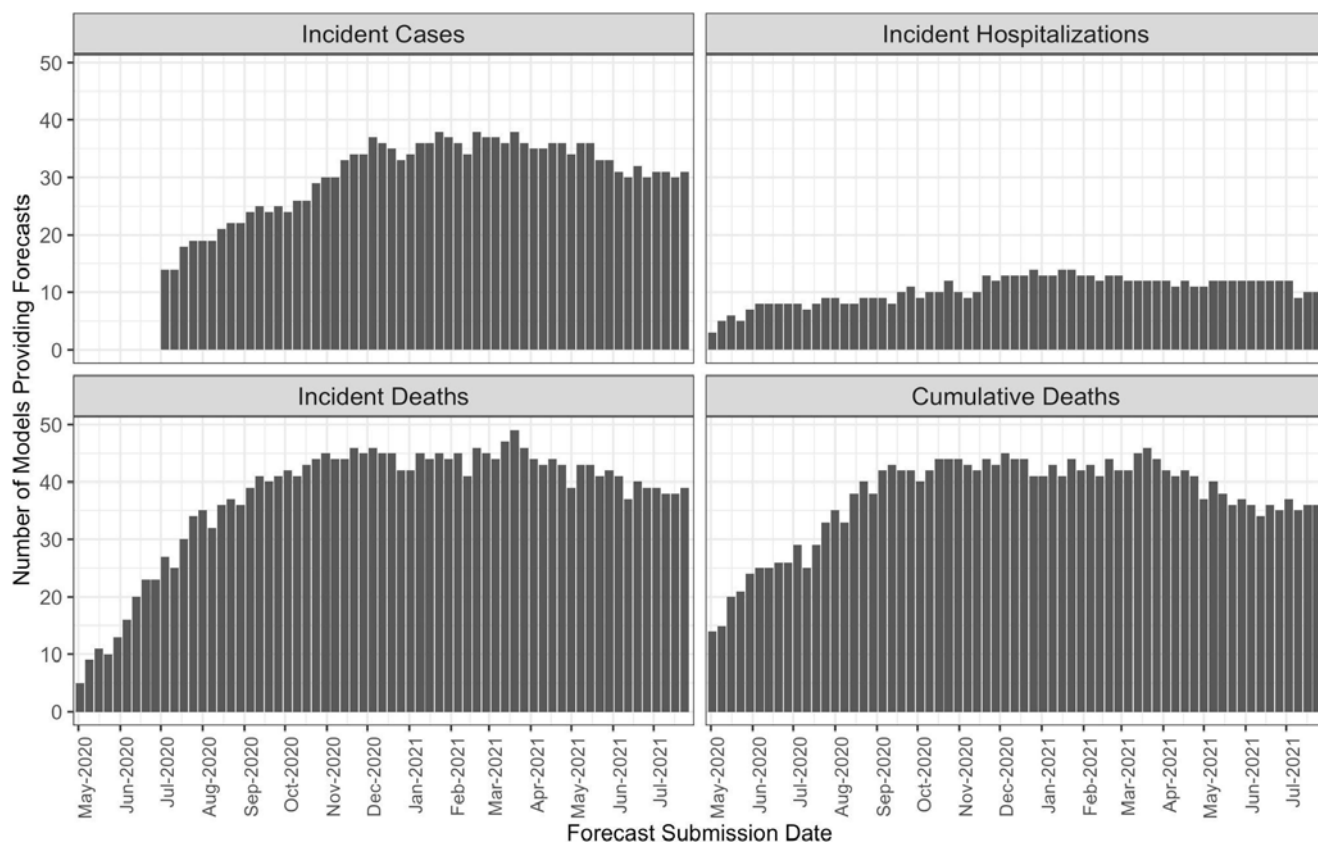




Table 1: Forecast characteristics for all four outcomes. The table shows the temporal scale, spatial scale of locations, horizons stored, number of quantiles, and the dates of the earliest forecast, earliest standardized truth data, and the earliest ensemble build.

| Outcome                   | Scale  | Locations |       |          | Horizons Stored | Number of quantiles for probabilistic forecasts | Earliest Forecast Date | First date of standardized truth data | Date of first ensemble forecast |
|---------------------------|--------|-----------|-------|----------|-----------------|---|------------------------|---------------------------------------|---------------------------------|
|                           |        | County    | State | National |                 |   |                        |                                       |                                 |
| Incident Cases            | Weekly | X         | X     | X        | 1 - 8 weeks     | 7   | 2020-07-05             | 2020-03-15                            | 2020-07-18                      |
| Incident Hospitalizations | Daily  |           | X     | X        | 1 - 130 days    | 23  | 2020-03-27             | 2020-11-16                            | 2020-12-05                      |
| Incident Deaths           | Daily  |           | X     | X        | 1 - 130 days    | 23  | 2020-03-15             | 2020-03-15                            | NA                              |
| Incident Deaths           | Weekly |           | X     | X        | 1-20 weeks      | 23  | 2020-03-15             | 2020-03-15                            | 2020-06-20                      |
| Cumulative Deaths         | Daily  |           | X     | X        | 1 - 130 days    | 23  | 2020-03-15             | 2020-03-15                            | NA                              |
| Cumulative Deaths         | Weekly |           | X     | X        | 1-20 weeks      | 23  | 2020-03-15             | 2020-03-15                            | 2020-04-13                      |

## Supplemental Information:

### Supplement 1: Optional fields in each metadata file:

1. **institution\_affil** - University or company names, if relevant.
2. **team\_funding** - Like an acknowledgement in a manuscript, you can acknowledge funding here.
3. **repo\_url** - A github repository url or something similar.
4. **twitter\_handles** - one or more twitter handles (without the @) separated by commas.
5. **data\_inputs** - A description of the data sources used to inform the model and the truth data targeted by model forecasts. Common data sources are NYTimes, JHU CSSE, COVIDTracking, Google mobility, HHS hospitalization etc. An example description could be “cases forecasts use NYTimes data and target JHU CSSE truth data, hospitalization forecasts use and target HHS hospitalization data”
6. **citation** - a url (doi link preferred) to an extended description of your model, e.g. blog post, website, preprint, or peer-reviewed manuscript.
7. **methods\_long** - An extended description of the methods used in the model. If the model is modified, this field can be used to provide the date of the modification and a description of the change.

### Supplement 2: Validations for pull requests and metadata files submitted to the forecast hub

Each time an update (pull request) from a forecast team is submitted to the Hub, a set of validation rules that enforce the metadata requirements as outlined in the Data Records section are applied to the contents of the update. Any file in an update that fails to conform to the rules will cause the entire update to fail.

### Forecasts

Each forecast file is subject to the validation rules documented at: <https://github.com/reichlab/covid19-forecast-hub/wiki/Forecast-Checks>.

### Miscellaneous

Additionally, each team must have their files under a folder named consistently with their *model\_abbrev*, and they must only have one *primary* model.