

# Meta and consensus forecast of COVID-19 targets

Allison Codi,<sup>1,\*</sup> Damon Luk,<sup>1,\*</sup> Juan Cambeiro,<sup>2</sup> David Braun,<sup>1</sup> Tamay Besiroglu,<sup>2</sup>  
Eva Chen,<sup>3</sup> Luis Enrique Urtubey de C  sar  s,<sup>3</sup> and Thomas McAndrew<sup>1,†</sup>

<sup>1</sup> *College of Health, Lehigh University, Bethlehem, Pennsylvania, United States of America*

<sup>2</sup> *Metaculus*

<sup>3</sup> *Good Judgment Inc.*

(Dated: April 7, 2021)

## SUMMARY FINDINGS

As of Apr 4, 2021 there are a reported 32,192,971 cumulative confirmed COVID-19 cases, 567,484 cumulative deaths due to COVID-19 in the United States, according to the [COVID-19 Data Repository by the Center for Systems Science and Engineering \(CSSE\) at Johns Hopkins University](#), and on Apr 3, 2021 there were 5,541 US previous day adult and pediatric admissions to the hospital because of COVID-19 as reported by the [Department of Health and Human Services](#).

Our aim is to provide two types of probabilistic forecasts of the evolving US COVID-19 outbreak to support public health officials, infectious disease modeling groups, and the general public. The first forecast is an aggregate of probabilistic predictions from experts in the modeling of infectious disease, trained forecasters, and generalist forecasters called a *consensus forecast* (Fig. 1). The second forecast is a combination of our consensus forecast and the COVIDhub-ensemble, an ensemble of computational models hosted by the [COVID-19 Forecast Hub](#). This combination of predictions from computational models, from experts, and from trained and generalist forecasters is called a *metaforecast* (Fig. 4).

A consensus of experts in the modeling of infectious disease, trained forecasters from [Metaculus](#), and generalist forecasters from [Good Judgment Open](#)—a forecasting platform open to any interested member of the public, predicted for the week beginning March 21 and ending March 27 that the median number of pediatric and adult hospital admissions would be 25,808 (80% CI = [16,646, 34,572]), the median number of new confirmed cases of COVID-19 would be 340,988 (80% CI = [200,874, 532,109]), and the median number of new deaths due to COVID-19 would be 7,690 (80% CI: [4,816, 10,473]). These numbers are lower than predicted a month prior with March median predictions of deaths and hospitalizations approximately one half of predictions made for the end of February.

To help public health officials better predict the effect of vaccination on the trajectory of the US COVID-19 pandemic, we asked forecasters to predict the cumulative number of people who would have received one or more doses and who would have received two doses (fully vaccinated) of a COVID-19 vaccine by Mar 31. The consensus median prediction is that by Mar 31, 101,993,827 people would have received at least one dose of a vaccine (80% CI = [90,418,102, 117,260,754]), and 60,375,895 would have received two doses or be fully vaccinated (80% CI = [52,808,421, 65,886,093]).

In the past two months, three SARS-COV-2 variants of concern (B.1.1.7, B.1.351, P.1) have been publicly monitored by the CDC. A consensus forecast assigned the highest probability to the CDC monitoring 4 variants of concern by Apr 4, one additional variant than currently reported at [US COVID-19 Cases Caused By Variants](#). To increase awareness of community transmission of these variants, subject matter experts and trained forecasters recommended a minimum of 11.5% of positive COVID-19 cases be sequenced (80% CI = [4.0, 30.5]). The rate of B.1.1.7—the variant identified in the U.K.—in the United States has the potential to impact decisions about non-pharmaceutical interventions and changes in the pace of vaccination. A consensus of experts and trained forecasters from Metaculus predicted the rolling 7-day average of % B.1.1.7 among all sequences samples of COVID-19 in the United States on March 27 would be 44% (80% CI = [24, 62]).

Metaforecasts were generated for the number of incident confirmed cases and incident deaths for the week of Mar 21, 2021 (Fig. 4).

---

\* equal contribution

† [mcandrew@lehigh.edu](mailto:mcandrew@lehigh.edu)

## CONSENSUS FORECASTS

### Consensus predictions for posed questions

- The median consensus prediction of the number of new US COVID-19 hospital admissions beginning Mar 21, 2021 and ending Mar 27, 2021 is 26,775 (80% CI:[17,850, 36,125]).
- The median consensus prediction of the number of new US deaths due to COVID-19 beginning Mar 21, 2021 and ending Mar 27, 2021 is 8,125 (80% CI:[4,750, 11,875]).
- The median consensus prediction of the number of new US confirmed COVID-19 cases beginning Mar 21, 2021 and ending Mar 27, 2021 is 350,000 (80% CI:[200,000, 570,000]).
- The median consensus prediction of the seven day rolling average of the % B.1.1.7 in the United States on Mar 27, 2021 is 44 (80% CI:[24, 62]).
- The median consensus prediction of the minimum percentage of positive COVID-19 cases that should be sequenced is 11.5% (80% CI:[4.0%, 30.5%]).
- A consensus of experts and trained forecasters assigned the highest probability on the CDC monitoring 4 variants of concern by Apr 4, 2021. 4.0 (80% CI:[3.0, 6.5])
- The median consensus prediction of the cumulative number of people who will have received one or more doses of a COVID-19 vaccine in the U.S by Mar 31, 2021 is 100,415,000 (80% CI:[86,060,000, 114,335,000]).
- The median consensus prediction of the cumulative number of people who will be fully vaccinated in the United States by Mar 31, 2021 is 58,175,000 (80% CI:[50,865,000, 64,840,000]).

## CONSENSUS PREDICTIVE DENSITIES

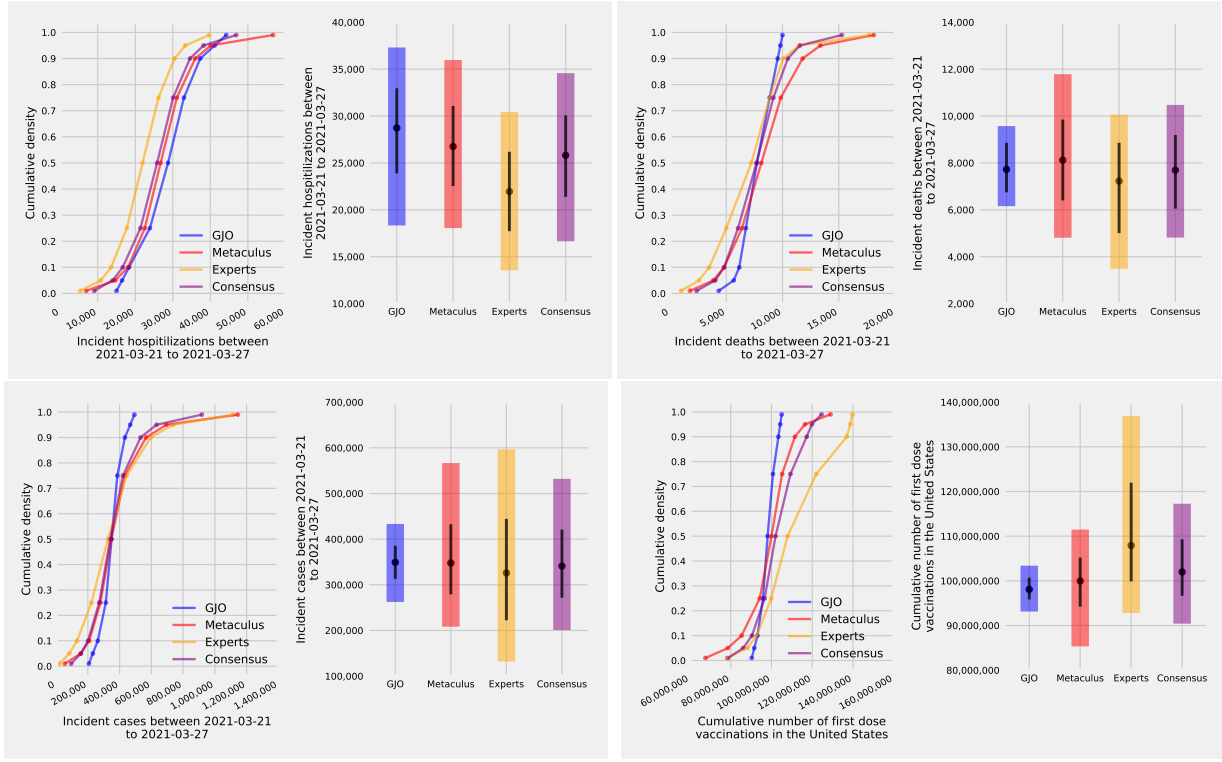


FIG. 1: (Left) Cumulative density functions of a consensus of trained forecasters from Metaculus, generalist forecasters from Good Judgment Open, and subject matter experts, and an equally weighted combination of all three consensus predictions (a *consensus*). (Right) An 80% prediction interval (the 10th and 90th percentile), 50% prediction interval (25th and 75th percentile), and the median (50th) percentile for all four consensus models.

# CONSENSUS PREDICTIVE DENSITIES (METACULUS PLATFORM)

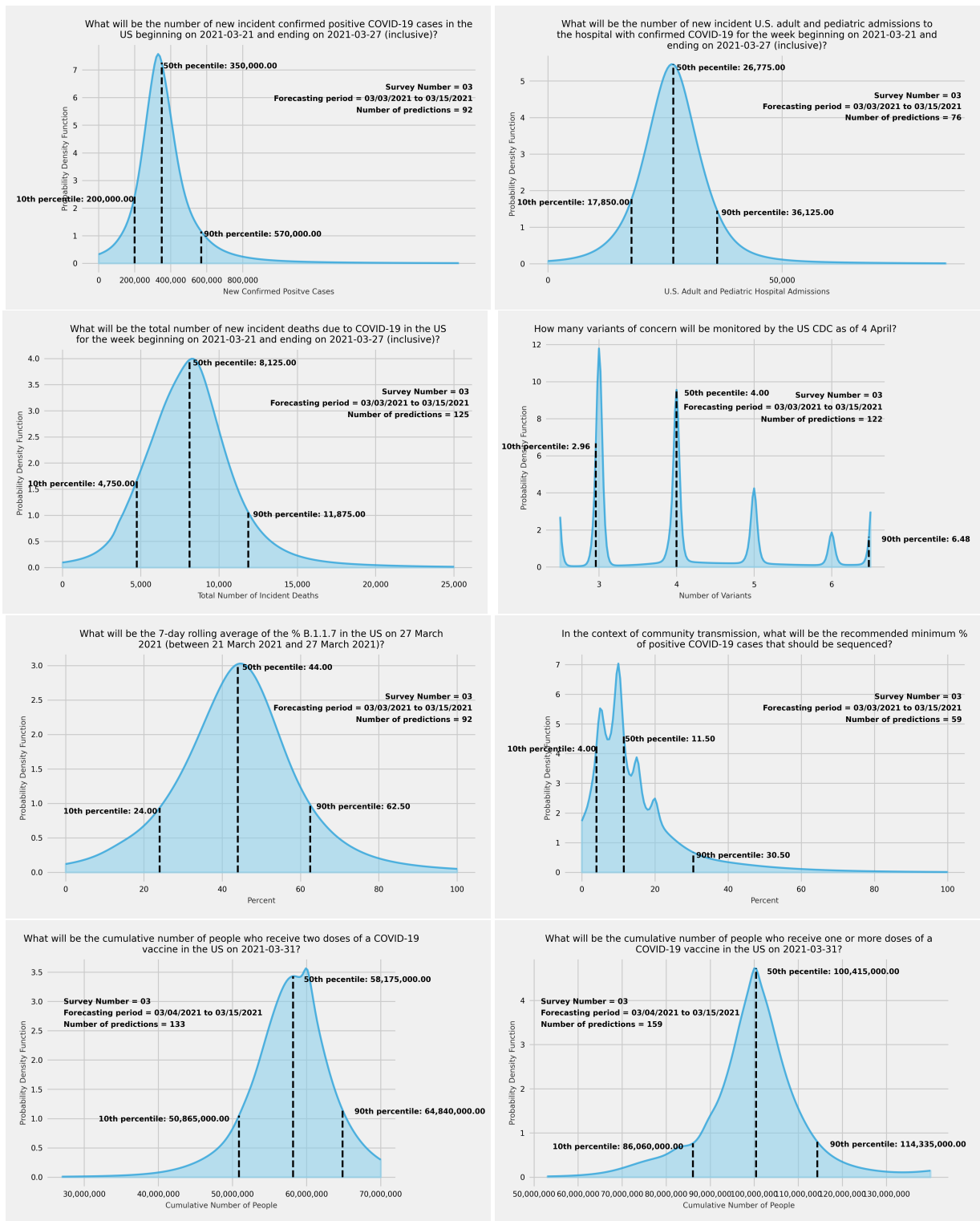


FIG. 2: Consensus predictive densities of eight targets of the US COVID-19 outbreak for those who used the Metaculus platform. Black dotted lines represent (from left to right) the tenth, fiftieth, and ninetieth percentiles of the distribution. Individual plots are available for download at [https://github.com/computationalUncertaintyLab/aggStatModelsAndHumanJudgment\\_PUBL/tree/main/densities](https://github.com/computationalUncertaintyLab/aggStatModelsAndHumanJudgment_PUBL/tree/main/densities).

## CONSENSUS PREDICTIVE DENSITIES (GOOD JUDGMENT OPEN PLATFORM)

One hundred and forty nine generalist forecasters who work on the public forecasting website Good Judgment Open (GJO) made, over a two-week period, 734 distinct forecasts on four separate public health targets.

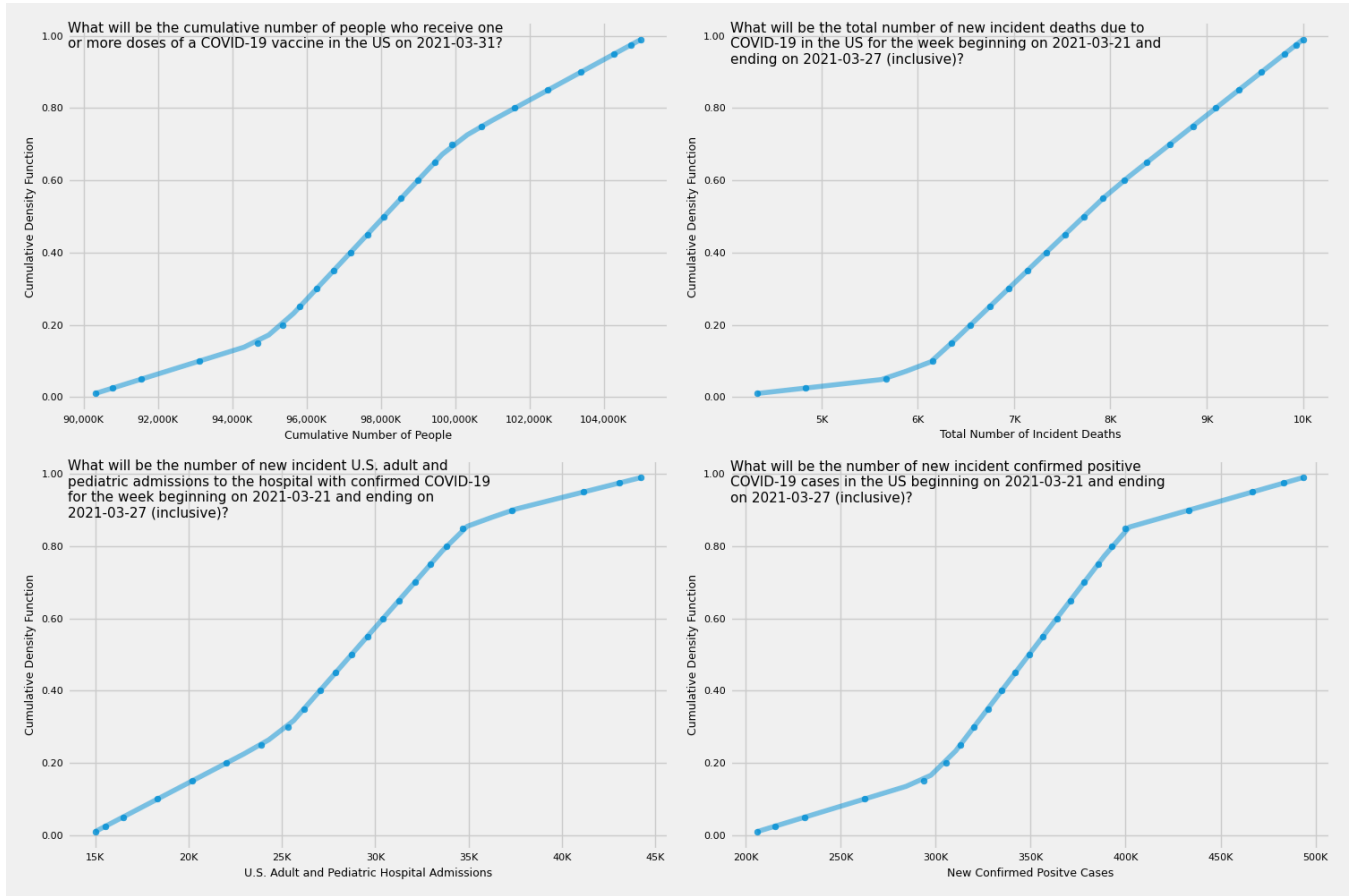


FIG. 3: Cumulative density functions of four targets of the US COVID-19 outbreak for those who used the Good Judgment Open platform. A cumulative density function is the probability that the variable takes a value less than or equal to  $x$ . Plots are available for download at [https://github.com/computationalUncertaintyLab/aggStatModelsAndHumanJudgment\\_PUBL/tree/main/data/GJOforecasts](https://github.com/computationalUncertaintyLab/aggStatModelsAndHumanJudgment_PUBL/tree/main/data/GJOforecasts).

## METAFORECASTS

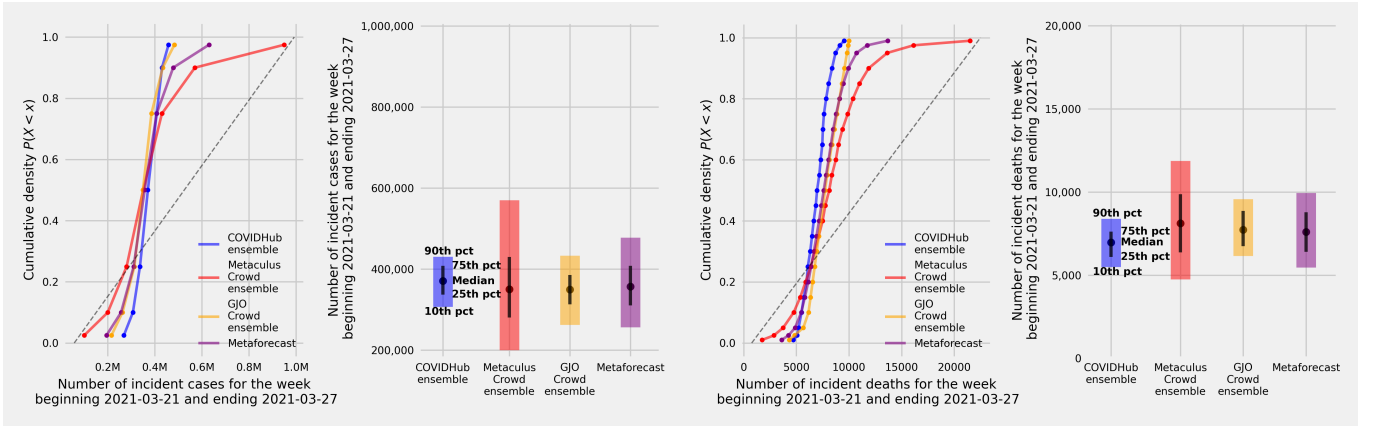


FIG. 4: Metaforecasts of the number of incident cases and incident deaths for the week beginning on Mar. 21, 2021 and ending on Mar 27, 2021. A metaforecast is a combination of an ensemble of computational models from the [COVID-19 Forecasthub](https://github.com/computationalUncertaintyLab/aggStatModelsAndHumanJudgment_PUBL/blob/main/metaforecasts/cases_survey3.png) and a consensus of predictions from trained forecasters and experts. (A) Cumulative density functions of an ensemble of computational models, consensus of predictions from trained forecasters and experts, and a metaforecast. (B) An 80% prediction interval (the 10th and 90th percentile), 50% prediction interval (25th and 75th percentile), and the median (50th) percentile for all three ensemble models. This plot is available for download at [https://github.com/computationalUncertaintyLab/aggStatModelsAndHumanJudgment\\_PUBL/blob/main/metaforecasts/cases\\_survey3.png](https://github.com/computationalUncertaintyLab/aggStatModelsAndHumanJudgment_PUBL/blob/main/metaforecasts/cases_survey3.png)

## MECHANICS OF A CONSENSUS FORECAST

To build a forecast we (i) pose questions to a crowd of subject matter experts and trained generalist forecasters, (ii) collect predictive densities over potential future values from each member of the crowd, and (iii) we aggregate this set of predictive densities into a single *consensus forecast*. We expect predictions from subject matter experts and trained forecasters to be accurate and calibrated because they have access to structured data, the same data computational models use, and because they have access to subjective, unstructured data often unavailable to computational models.

Suppose we collect a set  $F = \{f_1, f_2, f_3, \dots, f_C\}$  of predictive densities from  $C$  members of the crowd. Our consensus predictive density ( $f_{\text{consensus}}$ ) is generated as

$$f_{\text{consensus}}(v) = \sum_{i=1}^C \pi_i f_i(v)$$

such that  $\sum_{i=1}^C \pi_i = 1$

where  $f_i$  is the  $i$ th individual density from a crowd member,  $C$  is the number of individual densities, and  $\pi_i$  is the corresponding weight assigned to that member of the crowd. As of now, this summary report assigns equal weights ( $\pi_i = \frac{1}{C}$ ) to each member who proposed a predictive density  $f_i$

## MECHANICS OF A METAFORECAST

A metaforecast is built by combining an ensemble of computational models and a consensus of predictions from humans, in our case, predictions from trained forecasters and experts. The [COVID-19 forecast hub ensemble](https://github.com/computationalUncertaintyLab/aggStatModelsAndHumanJudgment_PUBL/blob/main/metaforecasts/cases_survey3.png) cumulative

predictive density is divided into 23 quantiles for incident deaths and 7 quantiles for incident cases. We can extract the same number of quantiles from our consensus forecast and compute quantiles for our metaforecast as

$$Q_{\text{metaforecast}} = \frac{1}{2}Q_{\text{COVID-19 ensemble}} + \frac{1}{2}Q_{\text{Consensus forecast}}$$

where  $Q_x$  is a quantile from distribution  $x$ .

## ONGOING METAFORECASTING WORK

There is ongoing work to generate a full density for the meta forecast by (i) estimating a probability density for the COVID-19 ensemble from quantile information and (ii) combining predictive densities generated by the COVID-19 ensemble and consensus forecast. We plan to generate a cumulative density function (cdf)  $F_{\text{COVID-19 ensemble}}$  for the COVID-19 ensemble using (i) monotonic cubic interpolation and (ii) fitting a mixture model that minimizes the sum square error between the fitted distribution and COVID-19 ensemble quantiles. We will combine the COVID-19 ensemble cumulative density and a linearly interpolated cumulative density from the consensus forecast as

$$F_{\text{metaforecast}}(v) = \pi_{\text{COVID-19 ensemble}} F_{\text{COVID-19 ensemble}} + \pi_{\text{Consensus forecast}} F_{\text{Consensus forecast}}$$

where  $\pi$  is the weight assigned to each ensemble. We can use the linearity of the derivative and our ability to differentiate the above interpolated cdfs to generate a probability density function for the metaforecast as

$$f_{\text{metaforecast}}(v) = \frac{dF_{\text{metaforecast}}}{dv} = \pi_{\text{COVID-19 ensemble}} \frac{dF_{\text{COVID-19 ensemble}}}{dv} + \pi_{\text{Consensus forecast}} \frac{dF_{\text{Consensus forecast}}}{dv}$$

## QUESTIONS AND RESOLUTION CRITERIA

Each month, we store all question and resolution text, the launch and close data to make forecasts, and urls corresponding to question in a dataset available for download at [https://github.com/computationalUncertaintyLab/aggStatModelsAndHumanJudgment\\_PUBL/blob/main/questions/questionData.csv](https://github.com/computationalUncertaintyLab/aggStatModelsAndHumanJudgment_PUBL/blob/main/questions/questionData.csv).

### • Question 1

- Question: What will be the number of new incident US adult and pediatric admissions to the hospital with confirmed COVID-19 for the week beginning on 2021-03-21 and ending on 2021-03-27 (inclusive)?
- Resolution Criteria: This question will resolve as the total number of adult plus pediatric previous day admissions with confirmed COVID-19 as recorded in the Department of Health and Human Services report of COVID-19 reported patient impact and hospital capacity for the dates from 2021-03-22 to 2021-03-28, corresponding to the number of hospitalizations from 2021-03-21 to 2021-03-27. Daily updates are provided by the Department of Health and Human Services. The total previous day admissions is computed using two variables in this report: `previous_day_admission_adult_covid_confirmed` and `previous_day_admission_pediatric_covid_confirmed` and stored in Lehigh University’s Computational Uncertainty Lab Github data repository. This report, and the resolution criteria, includes data on all 50 US states, Washington DC, Puerto Rico, and the US Virgin Islands (53 states and territories). The report will be accessed no sooner than (2021-04-04).
- Range:[0-80K]
- Metaculus question URL: <https://pandemic.metaculus.com/questions/6712/new-us-covid-hospital-admissions-21-27-march/>
- GJO question URL: <https://www.gjopen.com/questions/1952>

### • Question 2

- What will be the total number of new incident deaths due to COVID-19 in the US for the week beginning on 2021-03-21 and ending on 2021-03-27 (inclusive)?
- Resolution Criteria: This question will resolve as the number of new deaths due to confirmed COVID-19 for the week beginning on 2021-03-21 and ending on 2021-03-27 (inclusive) as recorded in the Johns Hopkins University (JHU) CSSE Github data repository. This file records the daily number of deaths by county. From this file deaths are summed across all counties and aggregated by week to generate the number of new deaths per week. The number of deaths for the week beginning on 2021-03-21 will be computed by adding the number of new deaths from the 2021-03-21 up to, and including, 2021-03-27. The report will be accessed no sooner than (2021-04-04).
- Range:[0-25k]
- Metaculus question URL: <https://pandemic.metaculus.com/questions/6713/new-us-covid-deaths-21-27-march/>
- GJO question URL: <https://www.gjopen.com/questions/1953>

- Question 3

- Question: What will be the number of new incident confirmed positive COVID-19 cases in the US beginning on 2021-03-21 and ending on 2021-03-27 (inclusive)?
- Resolution Criteria: This question will resolve as the number of new confirmed cases beginning on 2021-03-21 and ending on 2021-03-27 (inclusive) recorded in the Johns Hopkins University (JHU) CSSE Github data repository. This file records the daily number of cases by county. From this file cases are summed across all counties and aggregated by week to generate the number of new cases per week. The report will be accessed no sooner than 2021-04-04.
- Range:[0-2M]
- Metaculus question URL: <https://pandemic.metaculus.com/questions/6714/new-us-covid-cases-21-27-march/>
- GJO question URL: <https://www.gjopen.com/questions/1954>

- Question 4

- Question: What will be the 7-day rolling average of the % B.1.1.7 in the US on 27 March 2021 (between 21 March 2021 and 27 March 2021)?
- Resolution Criteria: This question will resolve as the 7 day rolling average of % sequences that are B.1.1.7 in the U.S. on 27 March 2021 (i.e. the average percentage between 21 March 2021 and 27 March 2021) at [this website](#). This percentage will be accessed no sooner than 6 April 2021.
- Range:[0-100]
- Metaculus question URL <https://pandemic.metaculus.com/questions/6717/-b117-in-the-us-21-27-march/>

- Question 5

- Question: In the context of community transmission, what will be the recommended minimum % of positive COVID-19 cases that should be sequenced?
- Resolution Criteria: This question will resolve as the minimum CDC recommended percent of confirmed positive COVID-19 cases that should be sequenced that assumes community transmission. If the CDC does not release such guidance before the end of 2021, then the most-cited paper that provides a recommendation on the minimum recommended percent of positive COVID-19 cases that should be sequenced in the context of community transmission will be consulted on 1 January 2022.
- Range:[0-100]
- Metaculus question URL: <https://pandemic.metaculus.com/questions/6718/-covid-cases-that-should-be-sequenced/>

- Question 6

- Question: How many variants of concern will be monitored by the US CDC as of 4 April?



- Resolution Criteria: This question will resolve as the number of variants of concern [monitored by the CDC](#) as of Sunday, 2021-04-04. For example, as of 2021-03-02 this page shows that there are three variants: B.1.1.7, B.1.351, and P.1. This page is updated on Sundays, Tuesdays, and Thursdays by 7pm ET and will be accessed at approximately 10pm ET on 2021-04-04 (a Sunday).
- Range:[0-8]
- Metaculus question URL: <https://pandemic.metaculus.com/questions/6719/-variants-monitored-by-cdc-on-4-april/>

- Question 7

- Question: What will be the cumulative number of people who receive one or more doses of a COVID-19 vaccine in the US on 2021-03-31?
- Resolution Criteria: This question will resolve as the cumulative number of people who receive one or more doses of a COVID-19 vaccine on 2021-03-31 as recorded by the Centers for Disease Control COVID-19 Data tracker. The radio buttons "People Receiving 1 or More Doses" and "Cumulative" will be selected and the bar corresponding to 2021-03-31 will be accessed. Data is updated daily by 8pm ET and will be accessed no sooner than 2021-04-04. If the CDC changes how it reports vaccination data, we will provide clarifying language as necessary. For purposes of this question, a person receiving a single-dose vaccine would count as a person having received one or more doses of a COVID-19 vaccine.
- Range:[0-140M]
- Metaculus question URL: <https://pandemic.metaculus.com/questions/6768/cumulative-1st-dose-us-vaccinations-31-march/>
- GJO question URL: <https://www.gjopen.com/questions/1955>

- Question 8

- Question: What will be the cumulative number of people who receive two doses of a COVID-19 vaccine in the US on 2021-03-31?
- Resolution Criteria: This question will resolve as the cumulative number of people who receive 2 doses of a COVID-19 vaccine on 2021-03-31 as recorded by the Centers for Disease Control COVID-19 Data tracker. The radio buttons "People Receiving 2 Doses" and "Cumulative" will be selected and the bar corresponding to 2021-03-31 will be accessed. Data is updated daily by 8pm ET and will be accessed no sooner than 2021-04-04. If the CDC changes how it reports vaccination data, we will provide clarifying language as necessary. For purposes of this question, a person receiving a single-dose vaccine would count as a person having received one or more doses of a COVID-19 vaccine.

**Mar 8 edit:** On 2021-03-08, the CDC's vaccine tracker at <https://covid.cdc.gov/covid-data-tracker/#vaccinations> changed the "receiving 2 doses" figure to "fully vaccinated" to account for people who receive one dose of the Johnson & Johnson vaccine, which has been authorized as a single-dose regimen (by contrast, Pfizer/BioNTech and Moderna are authorized as two-dose vaccines). This question will resolve on the basis of the new "fully vaccinated" figure reported by the CDC.

- Range:[0-70M]
- Metaculus question URL: <https://pandemic.metaculus.com/questions/6769/cumulative-two-dose-us-vaccinations-31-march/>