COVID-19
ForecastHub

The US COVID-19 Forecast Hub: operations, research, and a few anecdotes

Nicholas G. Reich

Presentation for the European CDC Forecast Hub
23 February 2021

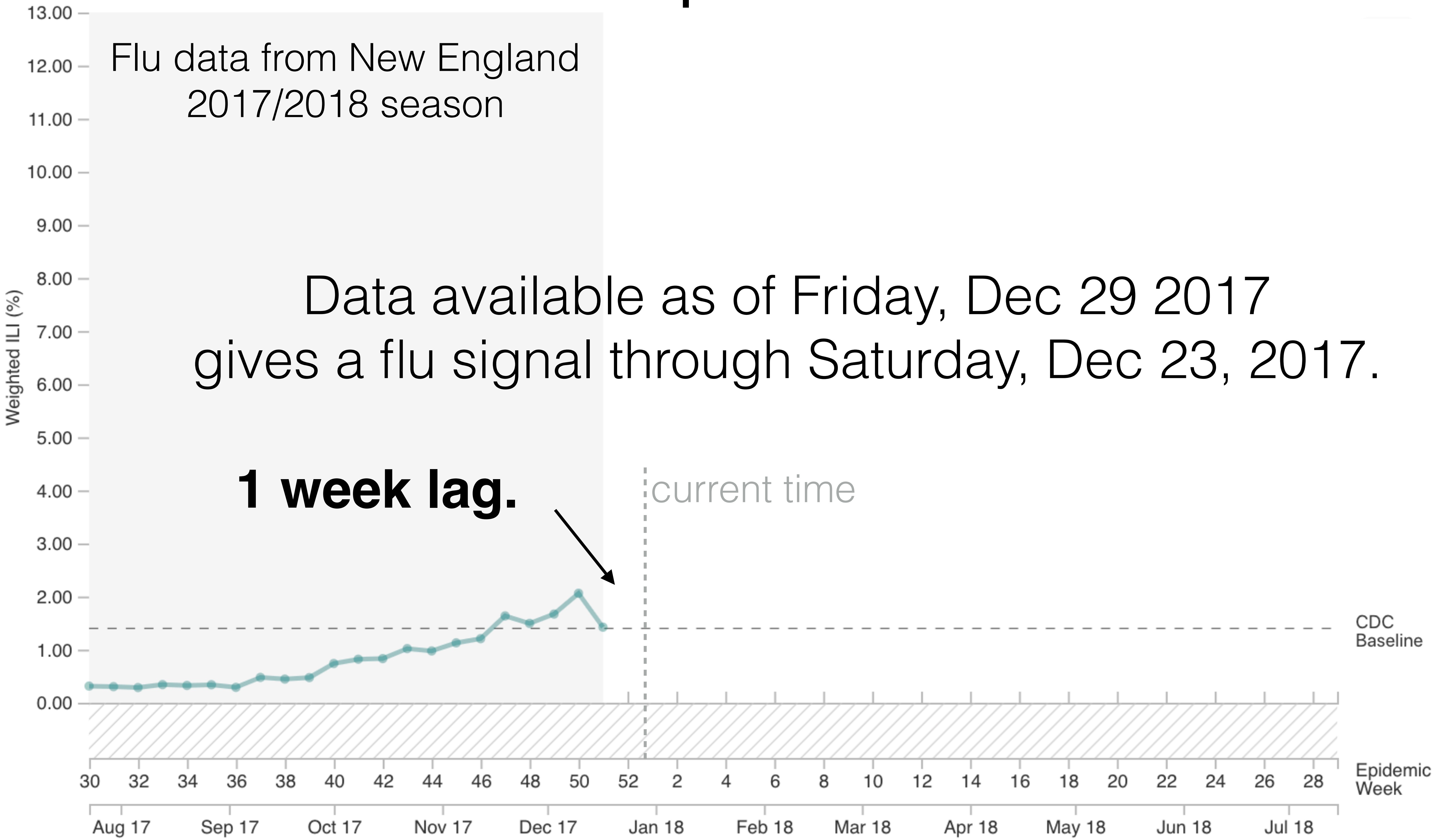
covid19forecasthub.org
reichlab.io
reichlab

Reich Lab | AT UMASS AMHERST

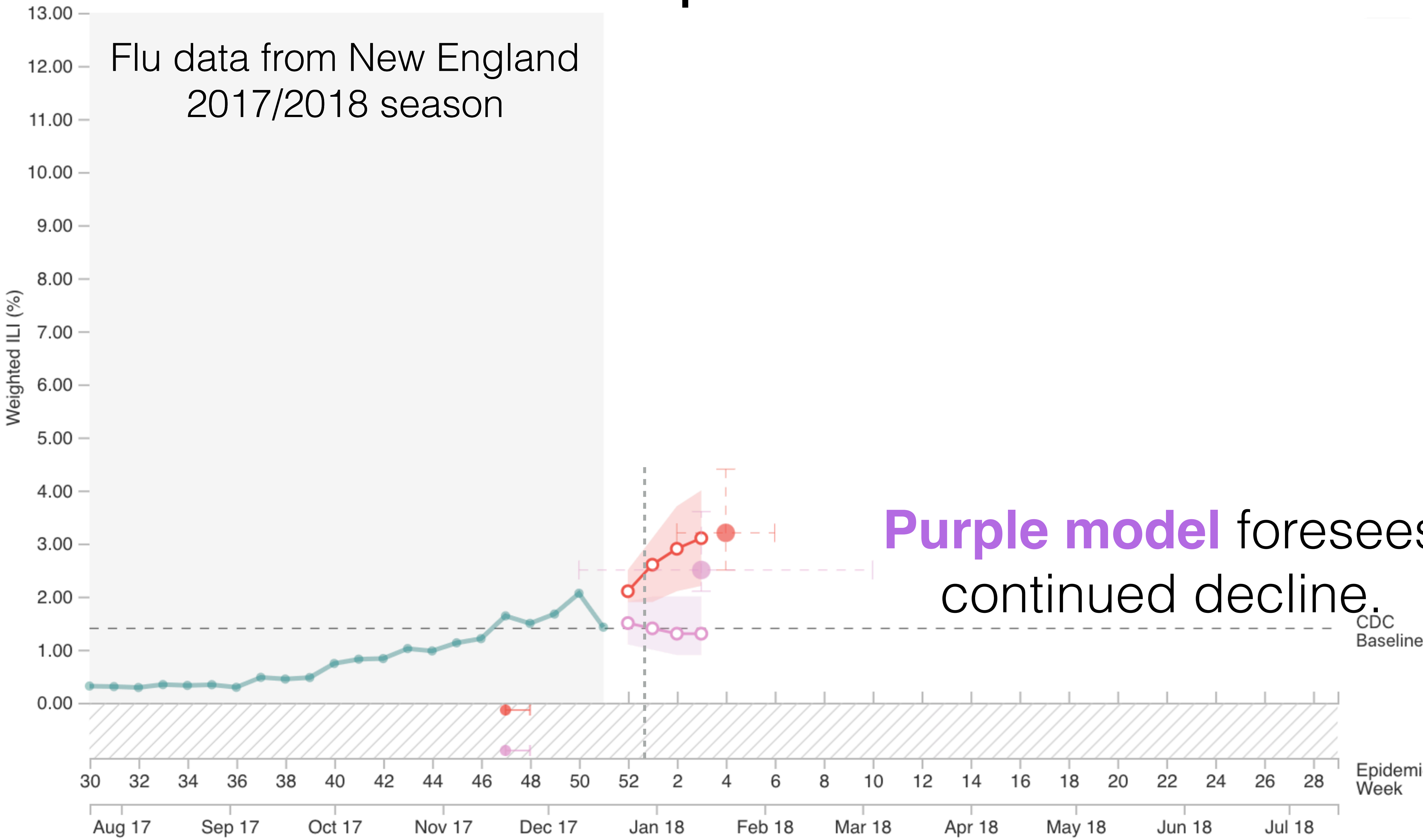
The COVID-19 Forecast Hub has been supported by the National Institutes of General Medical Sciences (R35GM119582) and the US Centers for Disease Control and Prevention (1U01IP001122). The content is solely the responsibility of the authors and does not necessarily represent the official views of NIGMS, the National Institutes of Health, or US CDC.

Why model outbreaks?

Real-time public health data is imperfect

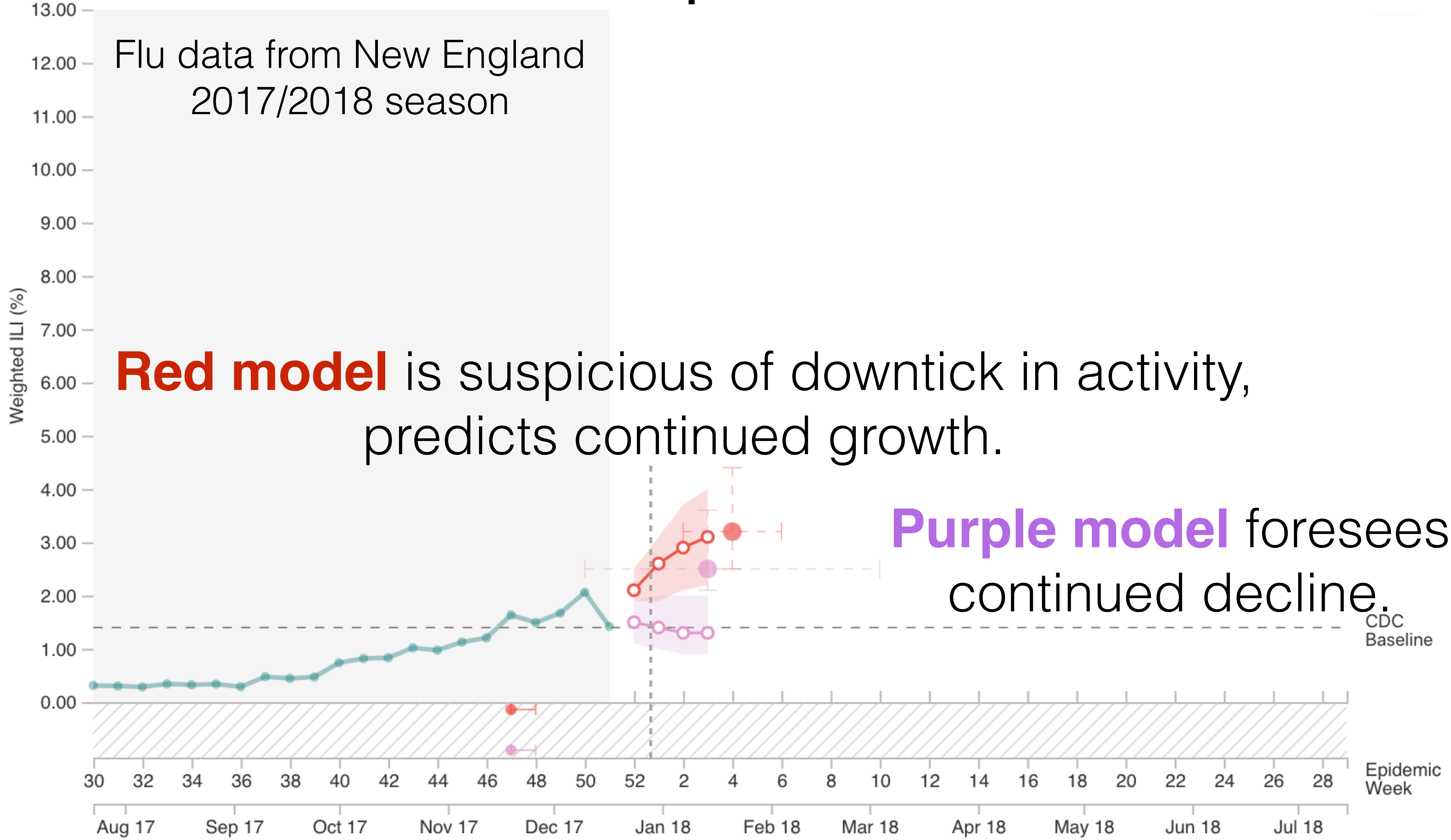


Real-time public health data is imperfect



Purple model foresees continued decline.

Real-time public health data is imperfect



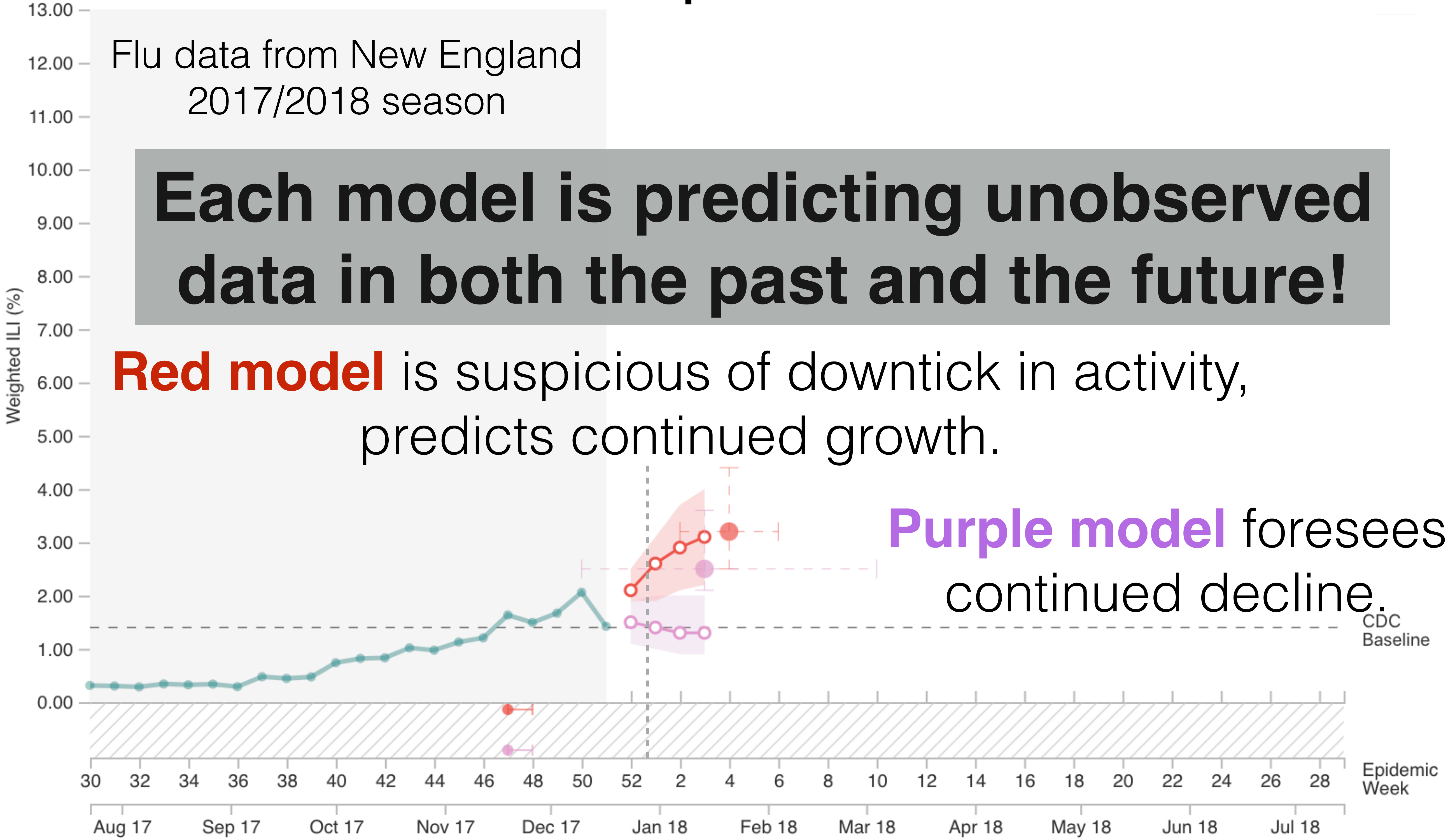
Real-time public health data is imperfect

Flu data from New England
2017/2018 season

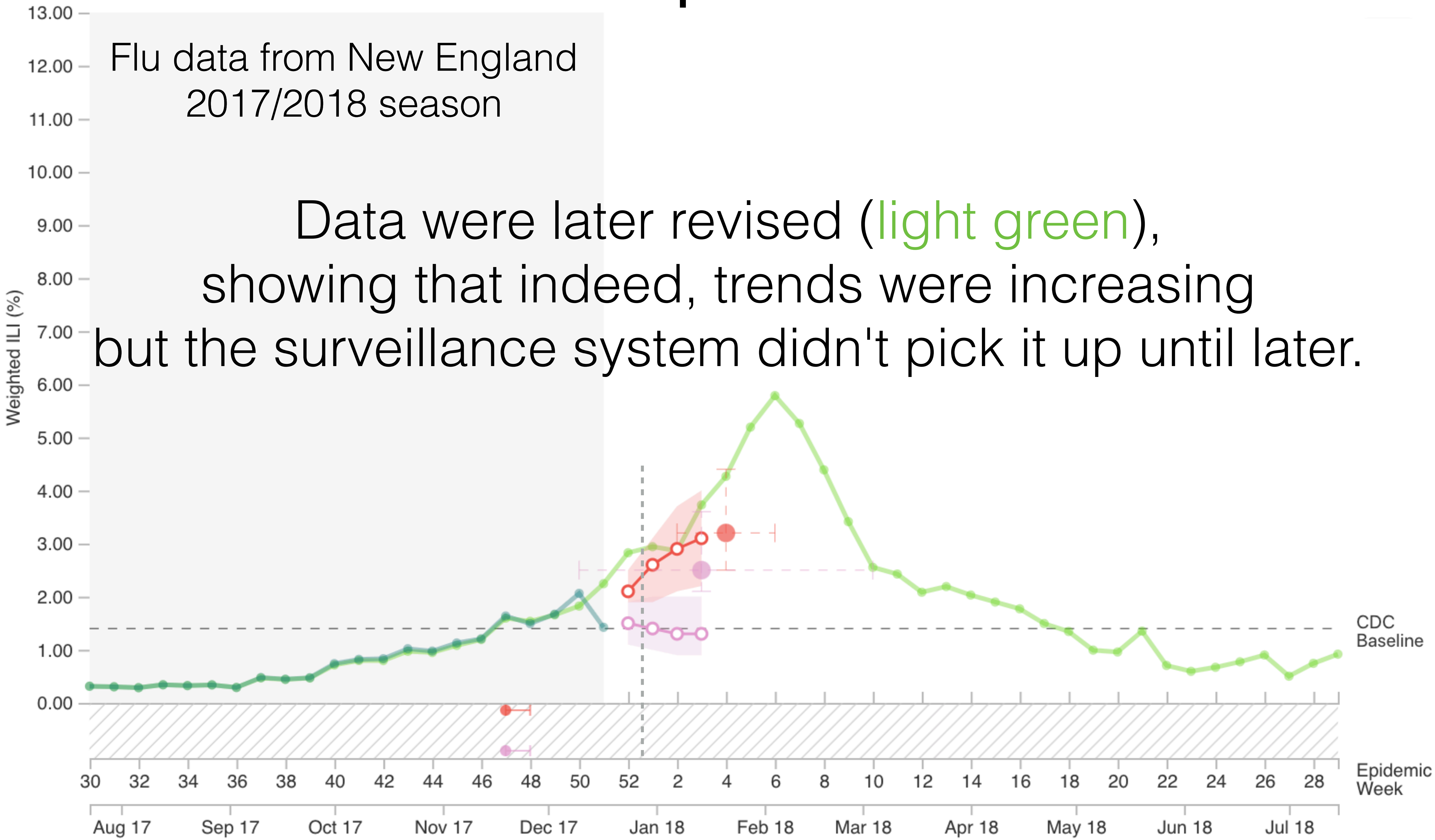
Each model is predicting unobserved data in both the past and the future!

Red model is suspicious of downtick in activity, predicts continued growth.

Purple model foresees continued decline.



Real-time public health data is imperfect



Good models might...

- Anticipate and adjust for data quality issues.
- Infer what is happening right now.
- Forecast what will be observed in the near future.
- Project hypothetical outcomes in the distant future.

Don't expect a single model to do all of these things well!

COVID-19 example

California COVID Assessment Tool

https://calcat.covid19.ca.gov/cacovidmodels/

Nowcasts

How fast is COVID-19 spreading right now?

Forecasts

What can we expect in the next 2-4 weeks?

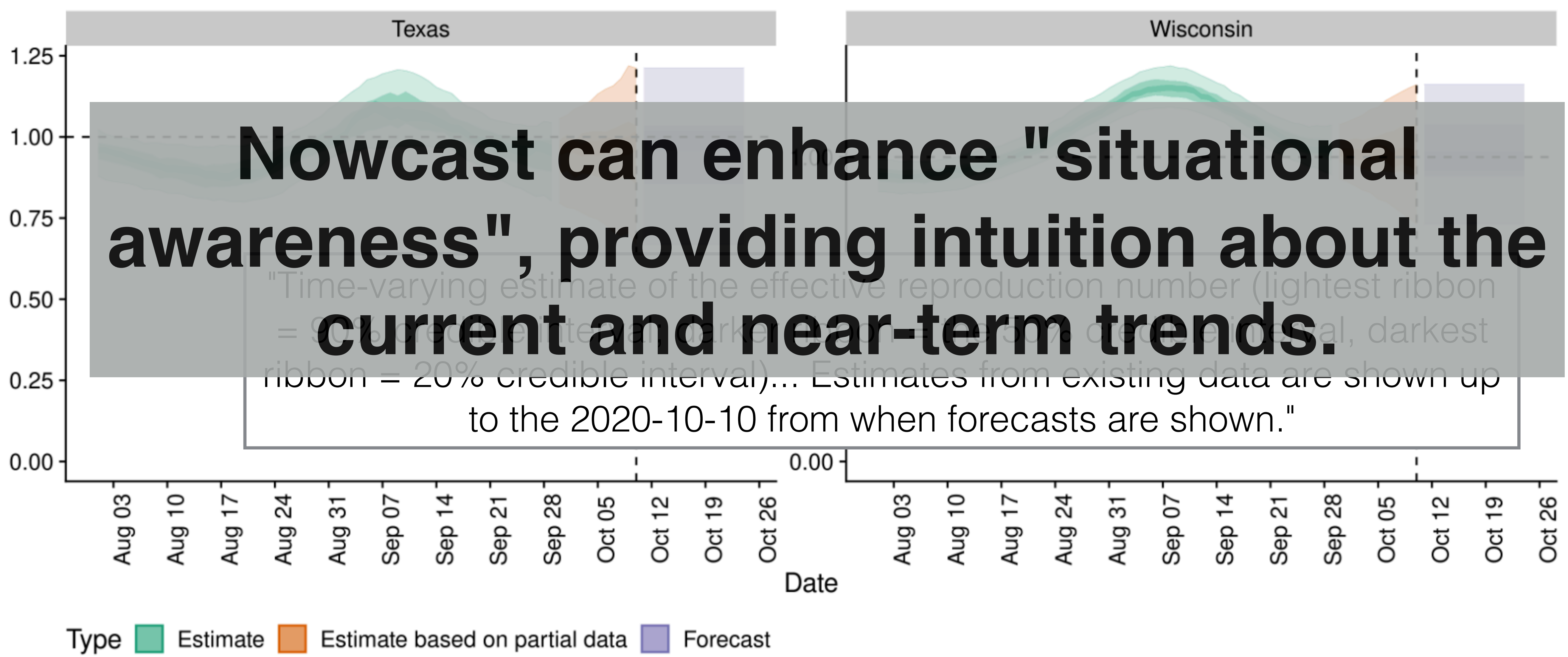
Scenarios

What are the long-term impacts under different scenarios?

How fast is COVID-19 spreading right now?

Nowcasting

Not as agreed upon definition, but I'd vote for "building a model that draws inference about trends the recent past."



Nowcast can enhance "situational awareness", providing intuition about the current and near-term trends.

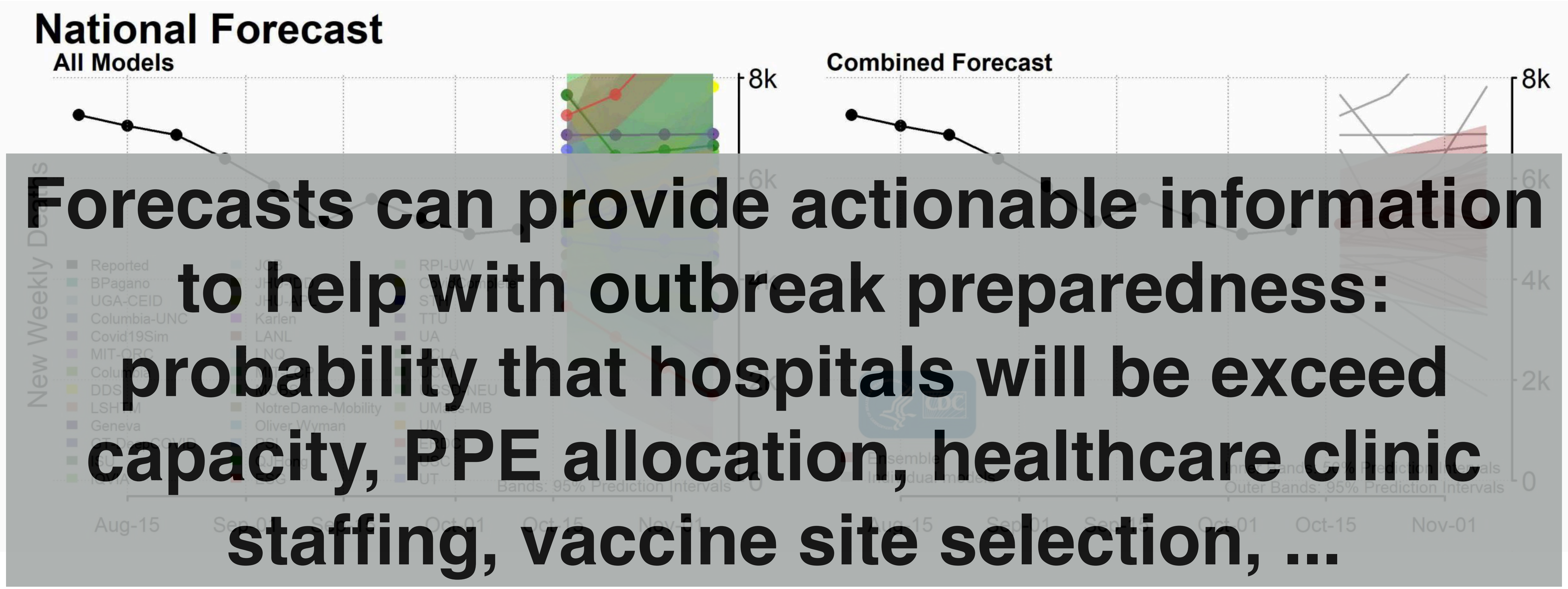
"Time-varying estimate of the effective reproduction number (lightest ribbon = 90% credible interval; darker ribbon = the 50% credible interval, darkest ribbon = 20% credible interval)... Estimates from existing data are shown up to the 2020-10-10 from when forecasts are shown."

https://epiforecasts.io/covid/posts/national/united-states/

Short-term Forecasting

What can we expect in the next 2-4 weeks?

Making **falsifiable, evaluable** predictions of observable future quantities.



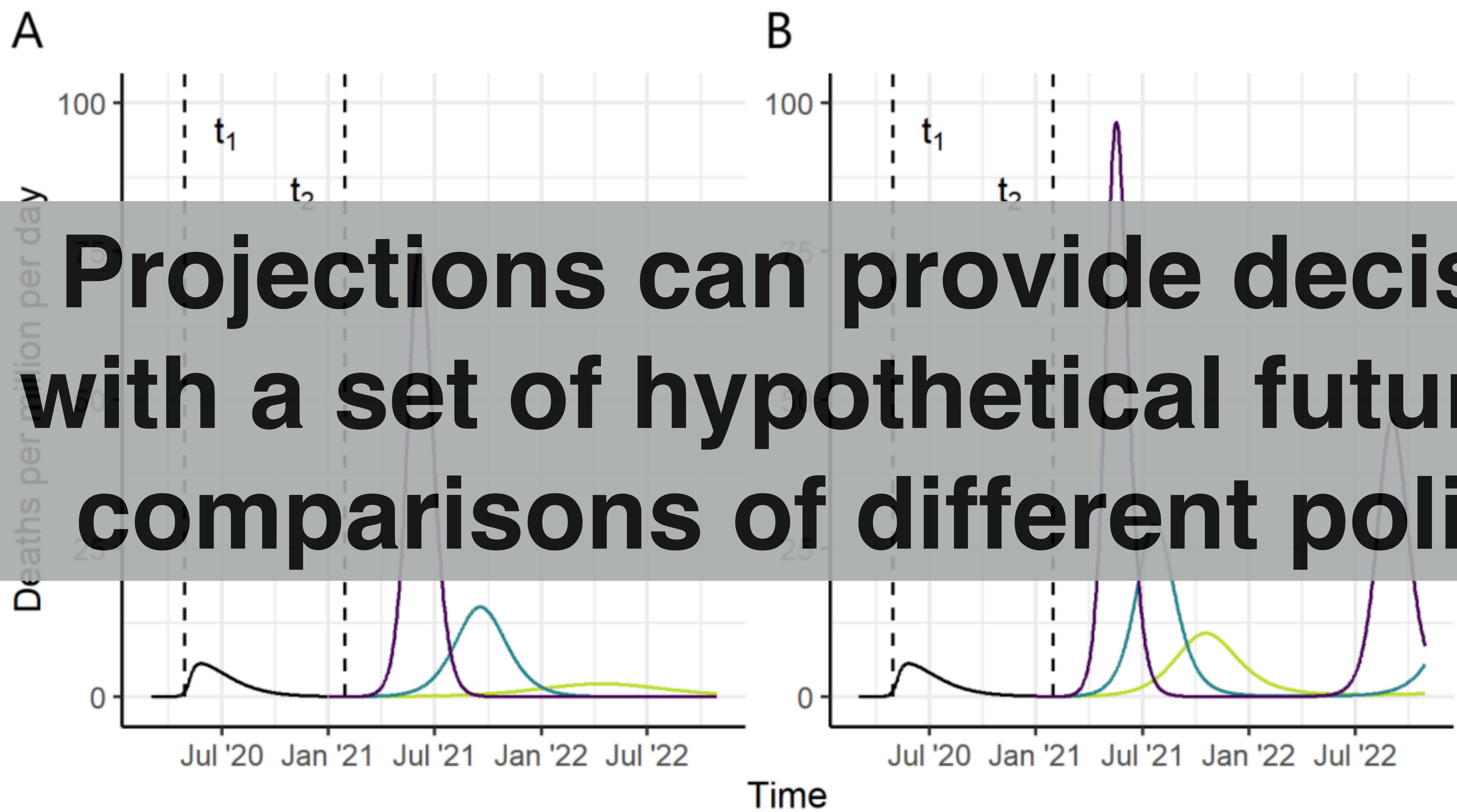
Forecasts can provide actionable information to help with outbreak preparedness: probability that hospitals will be exceed capacity, PPE allocation, healthcare clinic staffing, vaccine site selection, ...

https://www.cdc.gov/coronavirus/2019-ncov/covid-data/forecasting-us.html

Long-term Scenarios

What are the long-term impacts under different scenarios?

Projections based on specific assumptions.



Projections can provide decision-makers with a set of hypothetical futures based on comparisons of different policy choices.

Figure 1: Scenarios for the Course of the Epidemic from 2020–2022, for a High-Income Country Setting, in the Absence of a Vaccine (counterfactual scenarios). (A) Assuming "long immunity" and (B) assuming an average duration of naturally acquired immunity of 1 year. We assume that R_0=2.5 up to time t_1 (May 2020) and that R_{t1}

https://www.imperial.ac.uk/media/imperial-college/medicine/mrc-gida/2020-09-25-COVID19-Report-33.pdf

Central goals of US Hub

1. Provide decision-makers and general public with reliable information about where the pandemic is headed in the next month.

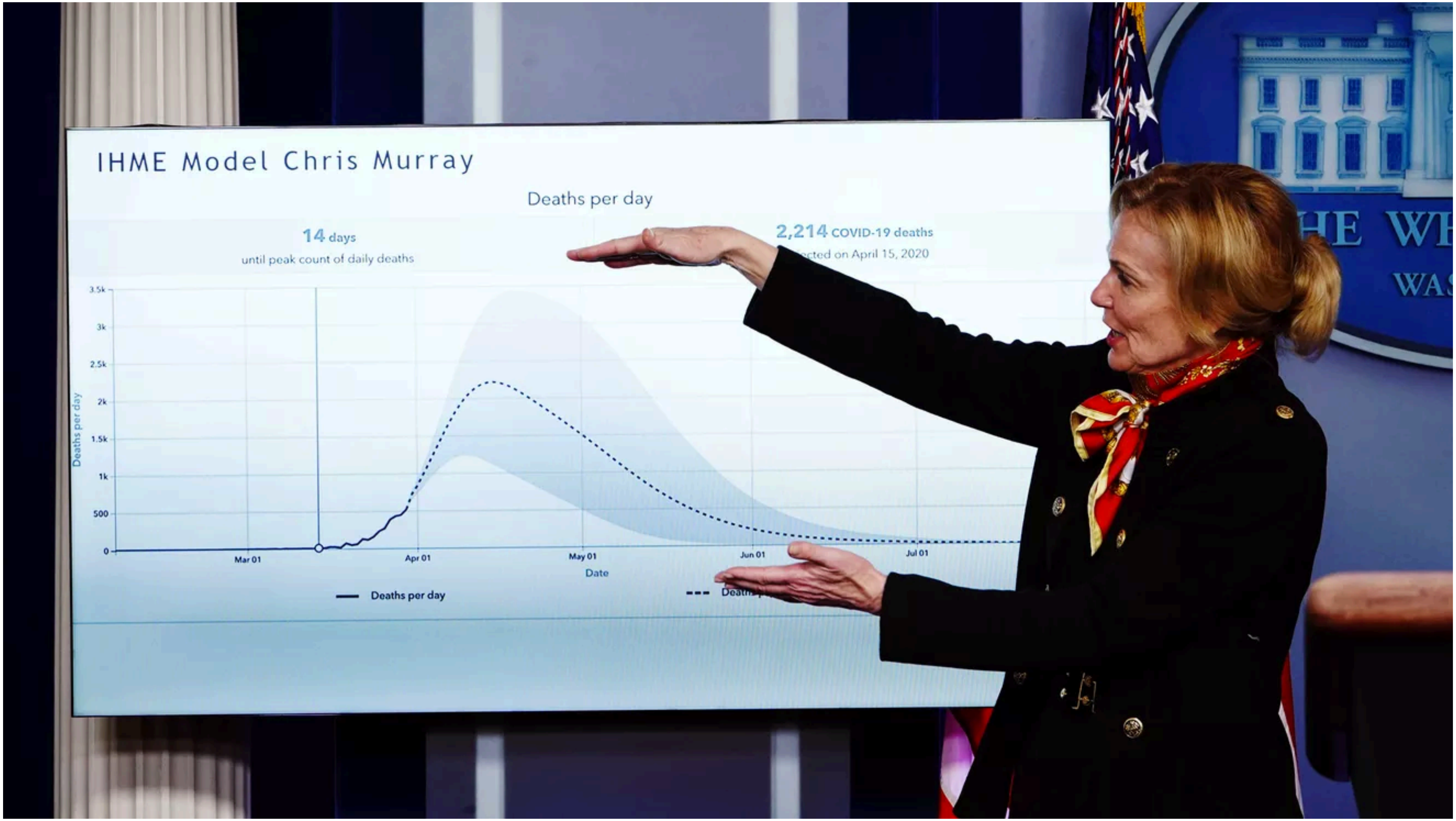
2. Gain insight into which modeling approaches do well. (Secondarily, hold models "accountable".)

3. Assess the reliability of forecasts for different measures of disease severity.

4. Create a community of infectious disease modelers underpinned by an open-science ethos.

Why a Hub?

Policy makers need >1 model



Model coordination for outbreaks

- Ensemble forecasting in established as the gold-standard in many fields: weather, economics, etc...

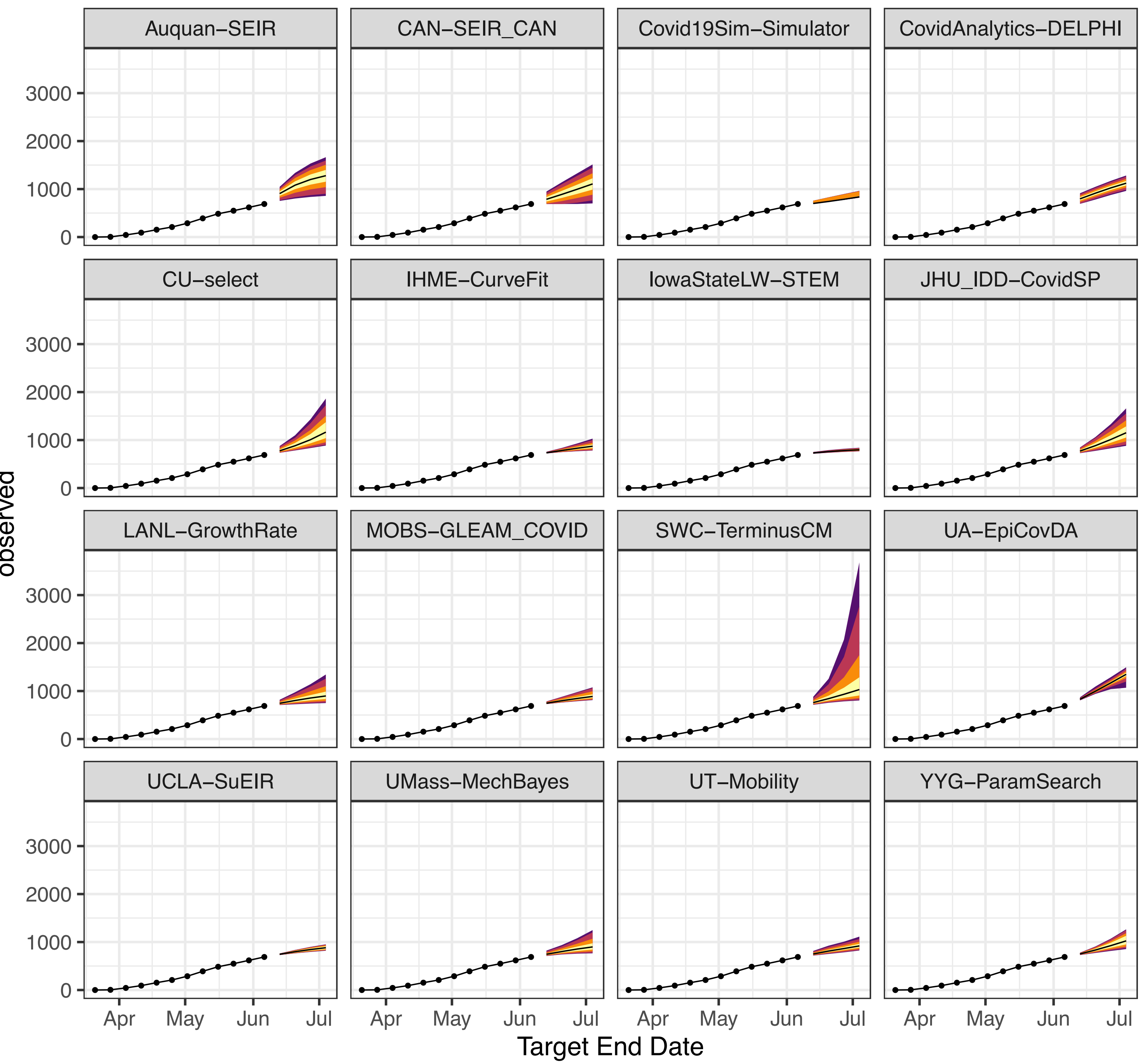
- There have been numerous government-coordinated outbreak forecasting efforts (flu, Ebola, chikungunya, Zika, dengue, etc...).

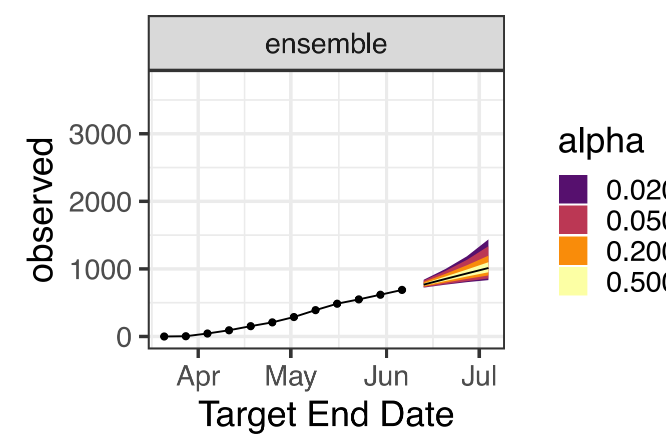
- One consistent finding across all efforts:

Combining models into an "ensemble" provides more consistent forecasts than any single model.

Flu: Reich et al. 2019, *PLOS Comp Bio*. https://doi.org/10.1371/journal.pcbi.1007486
Flu: McGowan et al. 2019, *Sci Rep*. https://doi.org/10.1038/s41598-018-36361-9
Dengue: Johansson et al. 2019, *PNAS*.
Ebola: Viboud et al. 2018, *Epidemics*.

A "Hub" enables model synthesis





COVID-19 ForecastHub

Background

- Each week the US Hub receives forecasts of weekly incident cases, hospitalizations and deaths in the US due to COVID-19 from over 50 research groups.

- The US Hub builds an ensemble that combines predictions from these models for 1 through 4 week ahead forecasts for the following targets and spatial scales.

target variable	scale	county	state	national
new cases	weekly	x	x	x
new hospitalizations	daily		x	x
new deaths	weekly		x	x
cumulative deaths	weekly		x	x

COVID-19 **ForecastHub**

- Forecast data from the COVID-19 Forecast Hub is shared directly with the CDC, and published on the CDC website weekly.

CDC Centers for Disease Control and Prevention
CDC 24/7: Saving Lives, Protecting People™

COVID-19 Forecasts: Deaths

Updated Nov. 19, 2020 Print

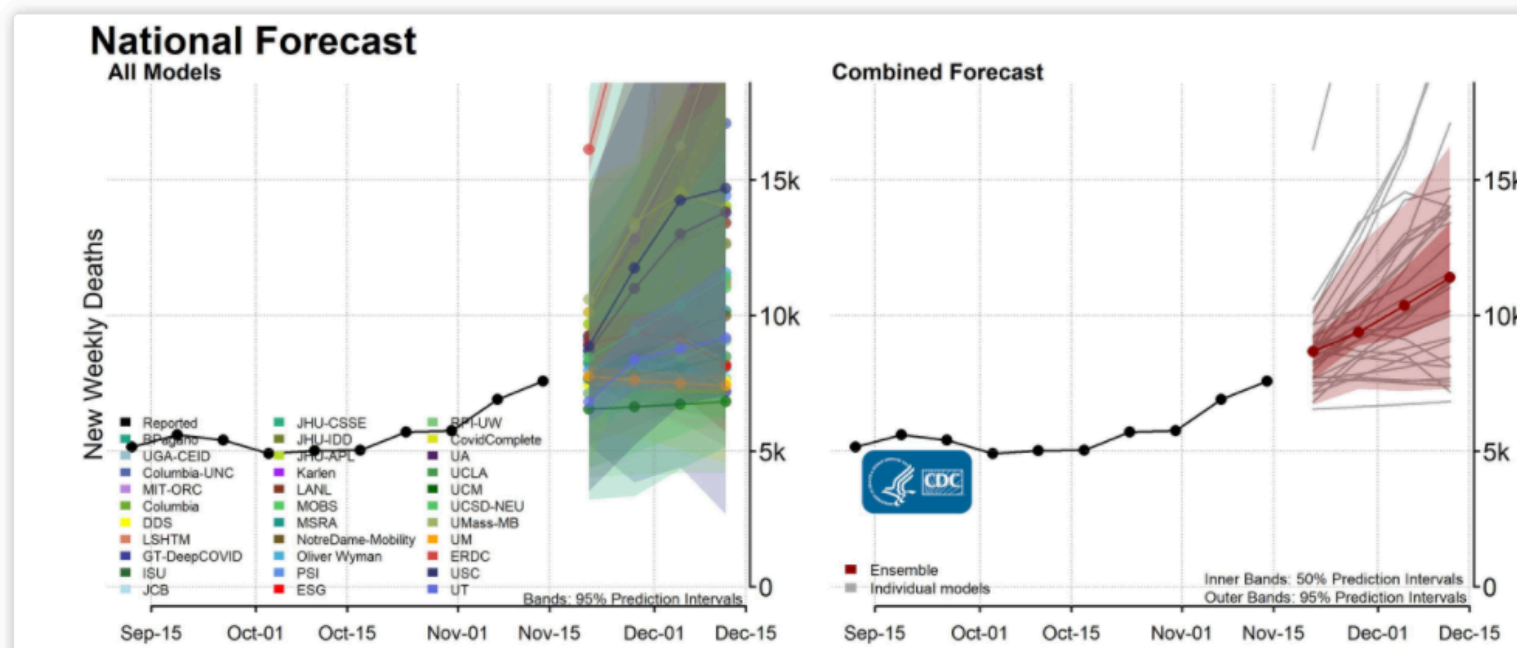
Observed and forecasted new and total reported COVID-19 deaths as of November 16, 2020.

Interpretation of Forecasts of New and Total Deaths

- This week CDC received forecasts of COVID-19 deaths over the next 4 weeks from 36 modeling groups that were included in the ensemble forecast. Of the 36 groups, 33 provided forecasts for both new and total deaths, two groups forecasted total deaths only, and one forecasted new death only.
- This week's national ensemble forecast predicts that the number of newly reported COVID-19 deaths will likely increase over the next four weeks, with 7,300 to 16,000 new deaths likely to be reported in the week ending December 12, 2020. The national ensemble predicts that a total of 276,000 to 298,000 COVID-19 deaths will be reported by this date.
- The state- and territory-level ensemble forecasts predict that over the next 4 weeks, the number of newly reported deaths per week will likely increase in 36 jurisdictions, which are indicated in the forecast plots below. Trends in numbers of future reported deaths are uncertain or predicted to remain stable in the other states and territories.

On This Page

National Forecast

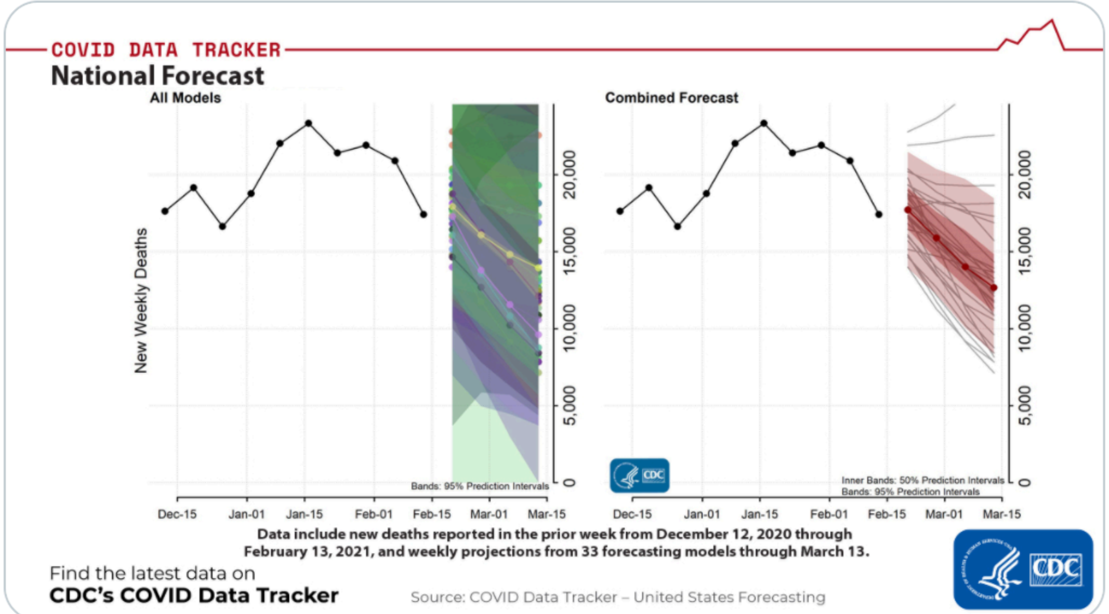


https://www.cdc.gov/coronavirus/2019-ncov/covid-data/mathematical-modeling.html

Current weekly public communications from US CDC

CDC @CDCgov · Feb 18

As of February 15, national forecasts predict that 8,400–18,500 new #COVID19 deaths and 530,000–559,000 total deaths will be reported during the week ending March 13. More: bit.ly/3cKQIl4.

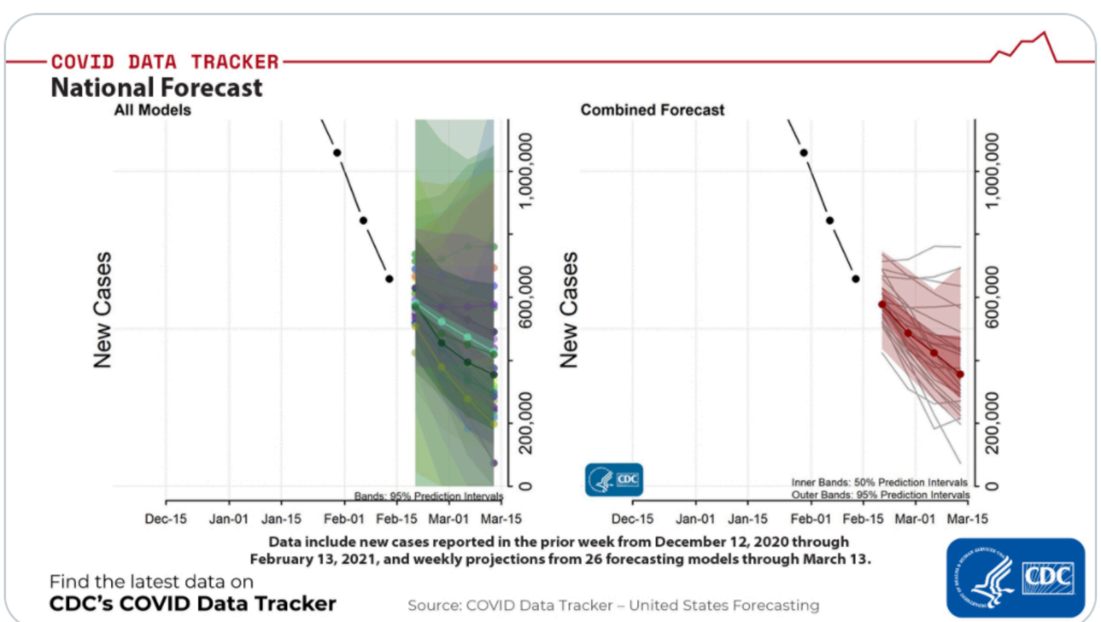


28 113 137

CDC @CDCgov · Feb 18

Replying to @CDCgov

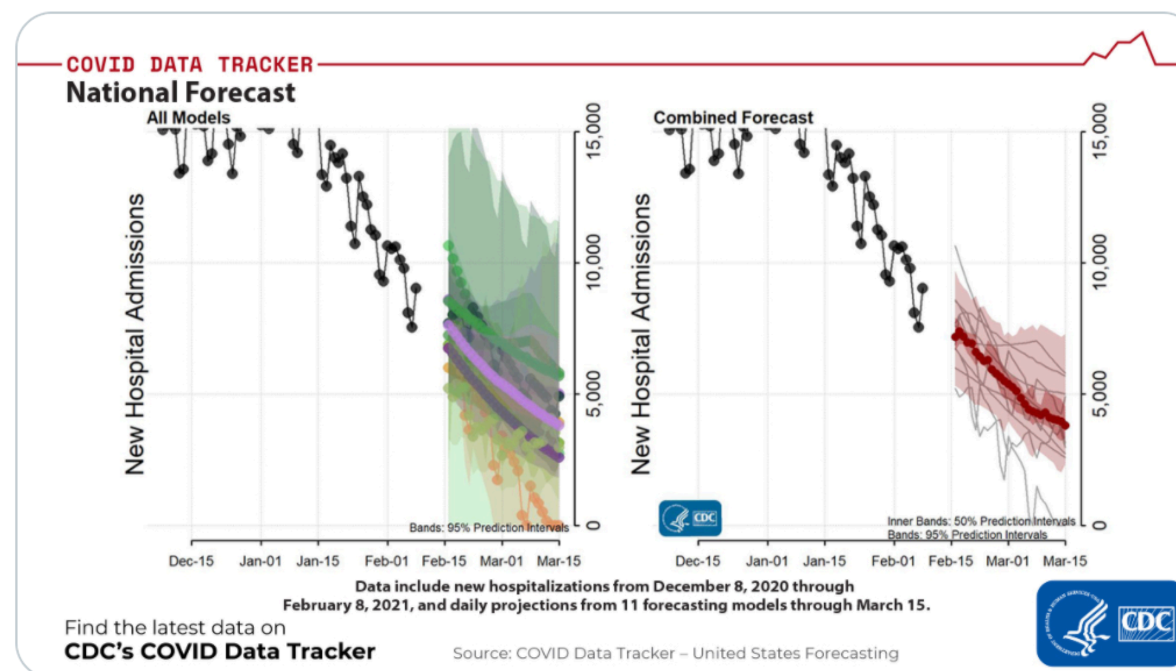
As of February 15, national forecasts predict that 206,000 to 699,000 new #COVID19 cases will likely be reported during the week ending March 13. More: bit.ly/2MRNLlv.



9 39 49

CDC @CDCgov · Feb 18

As of February 15, national forecasts predict that 2,300 to 7,300 new #COVID19 hospitalizations will likely be reported on March 15. More: bit.ly/2MUhi4i.

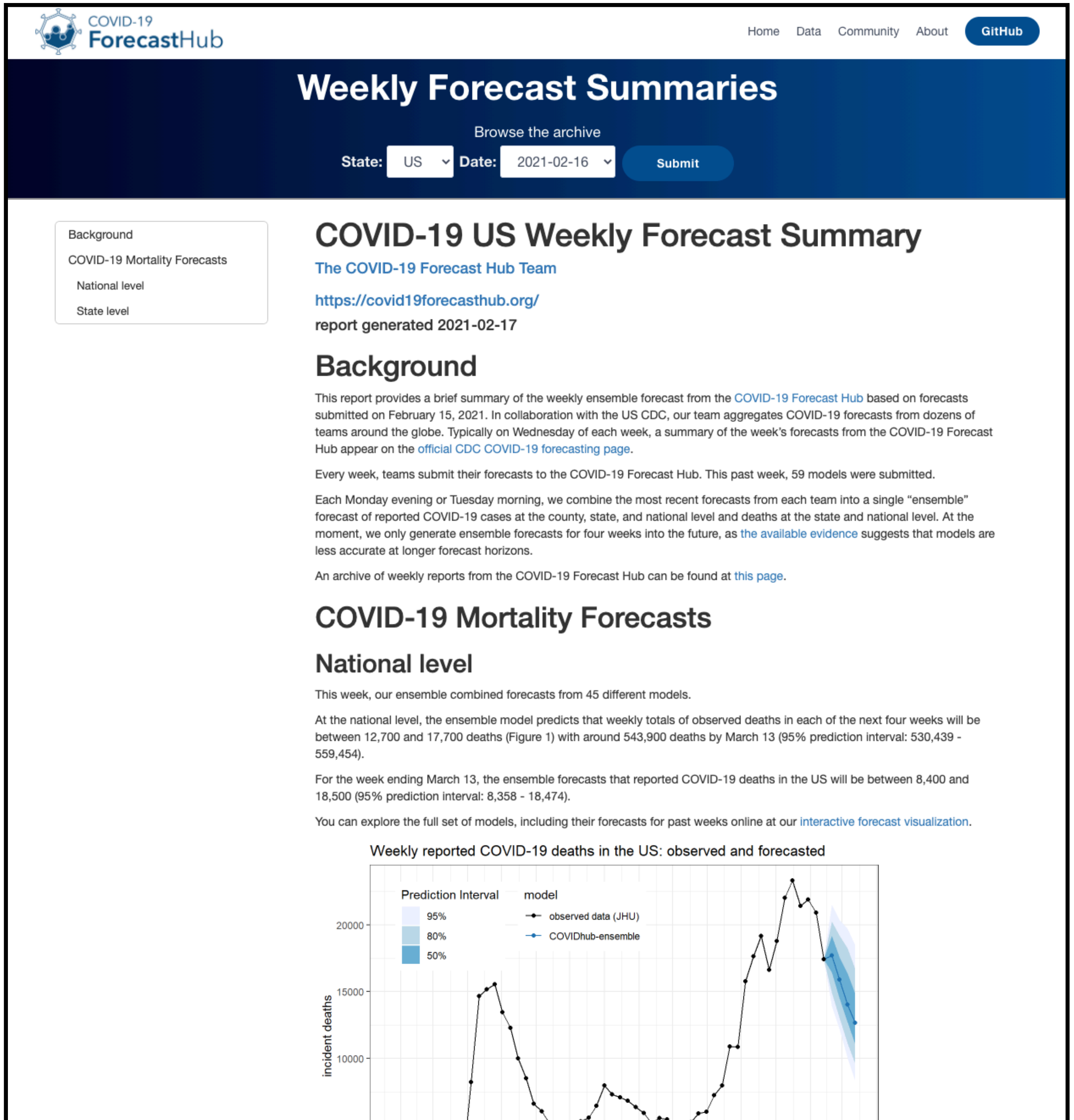


5 40 49

https://twitter.com/CDCgov/status/1362534701198671873

Weekly reports from the Hub

Summary


COVID-19 ForecastHub

Home Data Community About GitHub

Weekly Forecast Summaries

Browse the archive

State: US Date: 2021-02-16 Submit

Background
COVID-19 Mortality Forecasts
National level
State level

COVID-19 US Weekly Forecast Summary

The COVID-19 Forecast Hub Team

https://covid19forecasthub.org/

report generated 2021-02-17

Background

This report provides a brief summary of the weekly ensemble forecast from the COVID-19 Forecast Hub based on forecasts submitted on February 15, 2021. In collaboration with the US CDC, our team aggregates COVID-19 forecasts from dozens of teams around the globe. Typically on Wednesday of each week, a summary of the week's forecasts from the COVID-19 Forecast Hub appear on the official CDC COVID-19 forecasting page.

Every week, teams submit their forecasts to the COVID-19 Forecast Hub. This past week, 59 models were submitted.

Each Monday evening or Tuesday morning, we combine the most recent forecasts from each team into a single "ensemble" forecast of reported COVID-19 cases at the county, state, and national level and deaths at the state and national level. At the moment, we only generate ensemble forecasts for four weeks into the future, as the available evidence suggests that models are less accurate at longer forecast horizons.

An archive of weekly reports from the COVID-19 Forecast Hub can be found at this page.

COVID-19 Mortality Forecasts

National level

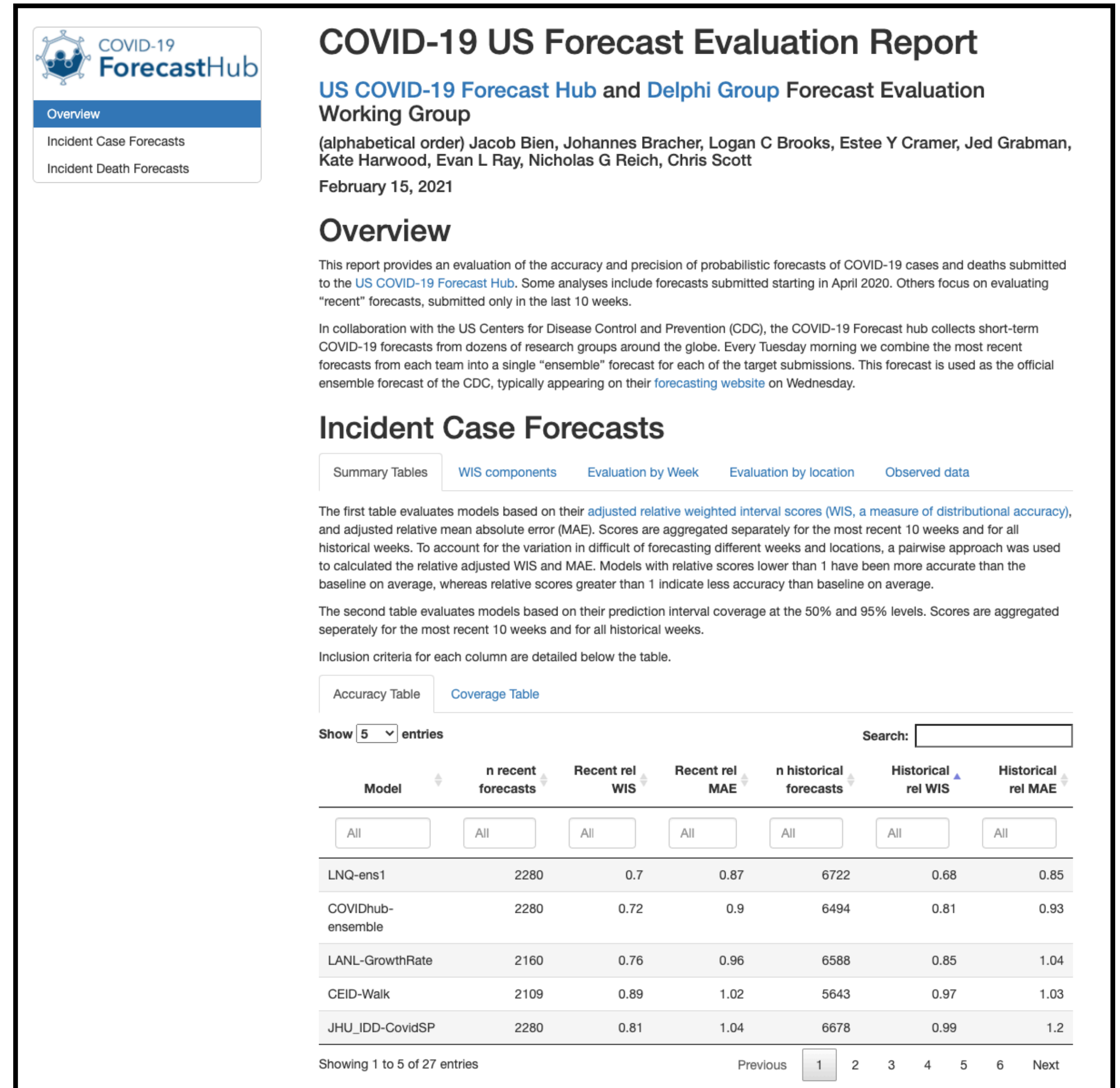
This week, our ensemble combined forecasts from 45 different models.

At the national level, the ensemble model predicts that weekly totals of observed deaths in each of the next four weeks will be between 12,700 and 17,700 deaths (Figure 1) with around 543,900 deaths by March 13 (95% prediction interval: 530,439 - 559,454).

For the week ending March 13, the ensemble forecasts that reported COVID-19 deaths in the US will be between 8,400 and 18,500 (95% prediction interval: 8,358 - 18,474).

You can explore the full set of models, including their forecasts for past weeks online at our interactive forecast visualization.

Evaluation


COVID-19 ForecastHub

Overview
Incident Case Forecasts
Incident Death Forecasts

COVID-19 US Forecast Evaluation Report

US COVID-19 Forecast Hub and Delphi Group Forecast Evaluation Working Group

(alphabetical order) Jacob Bien, Johannes Bracher, Logan C Brooks, Estee Y Cramer, Jed Grabman, Kate Harwood, Evan L Ray, Nicholas G Reich, Chris Scott

February 15, 2021

Overview

This report provides an evaluation of the accuracy and precision of probabilistic forecasts of COVID-19 cases and deaths submitted to the US COVID-19 Forecast Hub. Some analyses include forecasts submitted starting in April 2020. Others focus on evaluating "recent" forecasts, submitted only in the last 10 weeks.

In collaboration with the US Centers for Disease Control and Prevention (CDC), the COVID-19 Forecast hub collects short-term COVID-19 forecasts from dozens of research groups around the globe. Every Tuesday morning we combine the most recent forecasts from each team into a single "ensemble" forecast for each of the target submissions. This forecast is used as the official ensemble forecast of the CDC, typically appearing on their forecasting website on Wednesday.

Incident Case Forecasts

Summary Tables | WIS components | Evaluation by Week | Evaluation by location | Observed data

The first table evaluates models based on their adjusted relative weighted interval scores (WIS, a measure of distributional accuracy), and adjusted relative mean absolute error (MAE). Scores are aggregated separately for the most recent 10 weeks and for all historical weeks. To account for the variation in difficult of forecasting different weeks and locations, a pairwise approach was used to calculated the relative adjusted WIS and MAE. Models with relative scores lower than 1 have been more accurate than the baseline on average, whereas relative scores greater than 1 indicate less accuracy than baseline on average.

The second table evaluates models based on their prediction interval coverage at the 50% and 95% levels. Scores are aggregated seperately for the most recent 10 weeks and for all historical weeks.

Inclusion criteria for each column are detailed below the table.

Accuracy Table | Coverage Table

Show 5 entries   Search:

Model	n recent forecasts	Recent rel WIS	Recent rel MAE	n historical forecasts	Historical rel WIS	Historical rel MAE
All	All	All	All	All	All	All
LNQ-ens1	2280	0.7	0.87	6722	0.68	0.85
COVIDhub-ensemble	2280	0.72	0.9	6494	0.81	0.93
LANL-GrowthRate	2160	0.76	0.96	6588	0.85	1.04
CEID-Walk	2109	0.89	1.02	5643	0.97	1.03
JHU_IDD-CovidSP	2280	0.81	1.04	6678	0.99	1.2

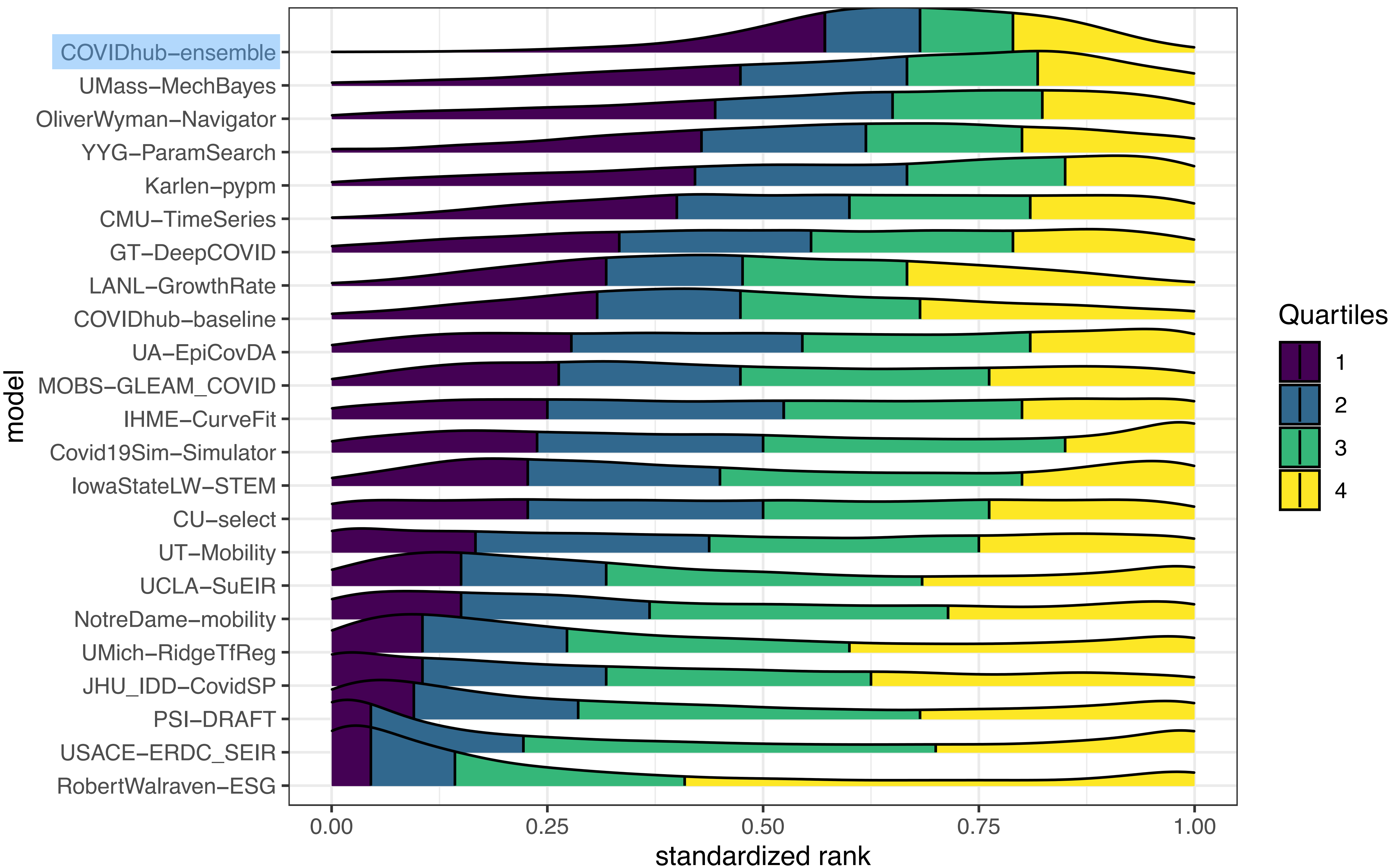
Showing 1 to 5 of 27 entries   Previous 1 2 3 4 5 6 Next

https://covid19forecasthub.org/doc/reports/

A few "results"

Hub ensemble is most consistent

Across 5,296 predictions it made, the ensemble is ranked in the top half of all forecasts for incident deaths over 75% of the time. No other model achieves this level of consistency.



Predicting cases is harder than deaths

We can look at an average "weighted interval score" relative to a naive baseline model (rel WIS) as a measure of predictability.
"95% cov" refers to observed coverage rates of 95% prediction intervals.

Case forecast accuracy

model	rel. WIS	95% cov.
LNQ-ens1	0.68	0.96
COVIDhub-ensemble	0.81	0.77
LANL-GrowthRate	0.85	0.85
CEID-Walk	0.97	0.68
JHU_IDD-CovidSP	0.99	0.82
COVIDhub-baseline	1.00	0.68

Death forecast accuracy

model	rel. WIS	95% cov.
COVIDhub-ensemble	0.68	0.87
UMass-MechBayes	0.72	0.95
OliverWyman-Navigator	0.74	0.85
Karlen-pypm	0.80	0.83
GT-DeepCOVID	0.83	0.82
CMU-TimeSeries	0.88	0.69
IHME-CurveFit	0.90	0.68
LANL-GrowthRate	0.93	0.90
CEID-Walk	0.95	0.79
MOBS-GLEAM_COVID	0.98	0.67
COVIDhub-baseline	1.00	0.82

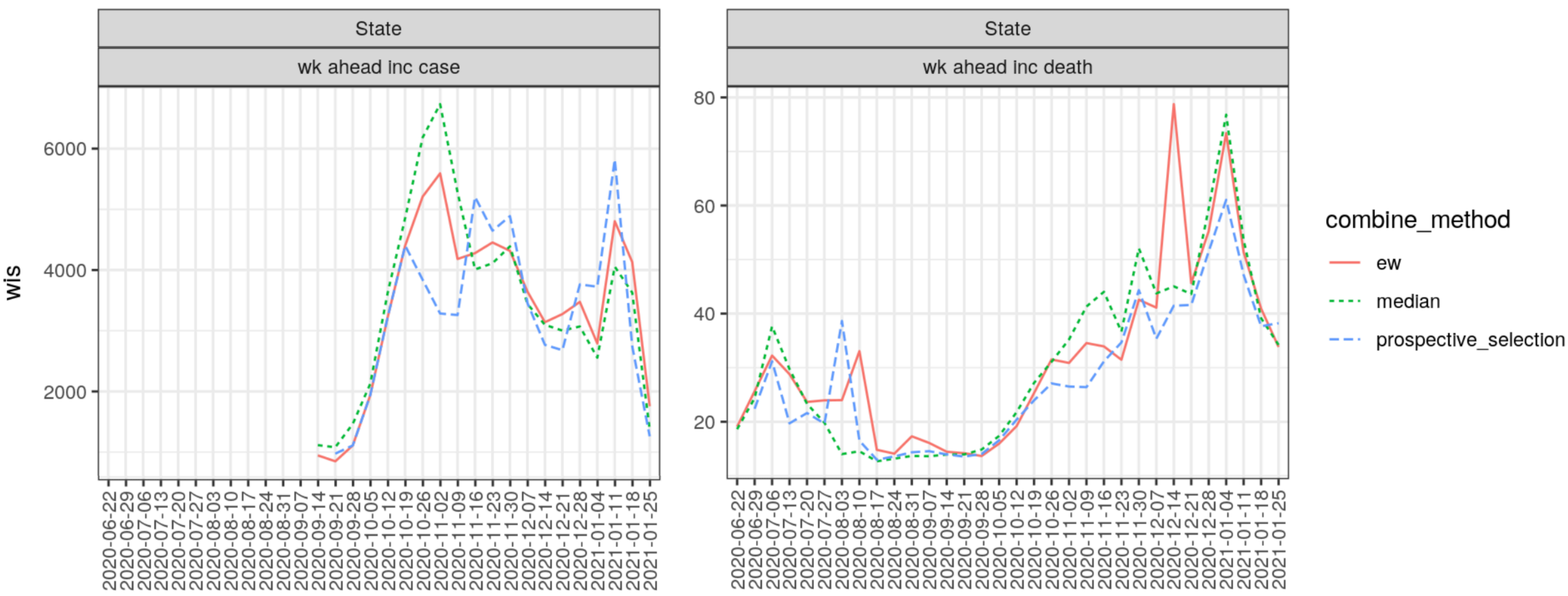
1. 5 models beat baseline for case forecasts, 10 for deaths.

2. 3 models are in both tables.

Simple ensemble is hard to beat

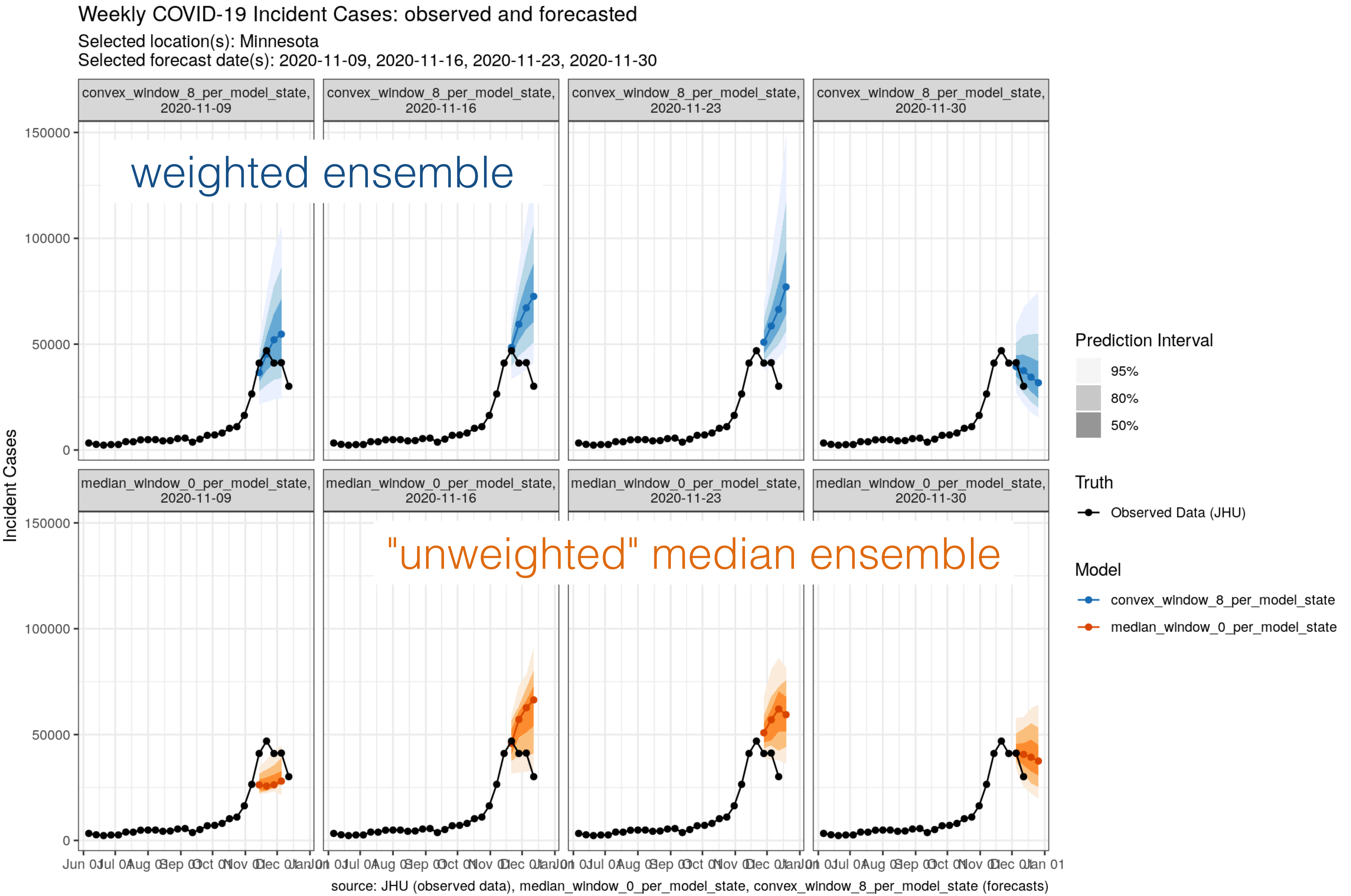
- We have looked at whether an ensemble that weights models differently would perform better.

- There is increasing evidence that, for incident deaths, weighting the models improves accuracy. Less so for cases.



https://forecasters.org/blog/2020/10/28/comparing-ensemble-approaches-for-short-term-probabilistic-covid-19-forecasts-in-the-u-s/

weighted ensembles sometimes are more pessimistic at the peaks



Individual models vary

roughly ordered by date of first submission

- IHME-CurveFit: "**hybrid modeling approach** to generate our forecasts, which incorporates elements of statistical and disease transmission models."
- YYG-ParamSearch: "**machine learning** techniques on top of a **classic infectious disease model** to make projections for infections and deaths."
- MOBS-GLEAM_COVID: "The GLEAM framework is based on **a metapopulation approach** in which the world is divided into geographical subpopulations. Human **mobility between subpopulations is represented on a network**."
- UMass-MechBayes: "**classical compartmental models from epidemiology**, prior distributions on parameters, models for time-varying dynamics, models for partial/noisy observations of confirmed cases and deaths."
- UT-Mobility: "For each US state, **we use local data from mobile-phone GPS traces** made available by [SafeGraph] to quantify the changing impact of social-distancing measures on 'flattening the curve.' "
- GT-DeepCOVID: "This **data-driven deep learning model** learns the dependence of hospitalization and mortality rate on various detailed syndromic, demographic, mobility and clinical data."
- Google Cloud AI: "a novel approach that integrates **machine learning** into **compartmental disease modeling** to predict the progression of COVID-19"
- Facebook AI: "**recurrent neural networks** with a vector autoregressive model and train the joint model with a specific regularization scheme that increases the **coupling between regions**"

How did we get here?

US COVID-19 Forecast Hub timeline

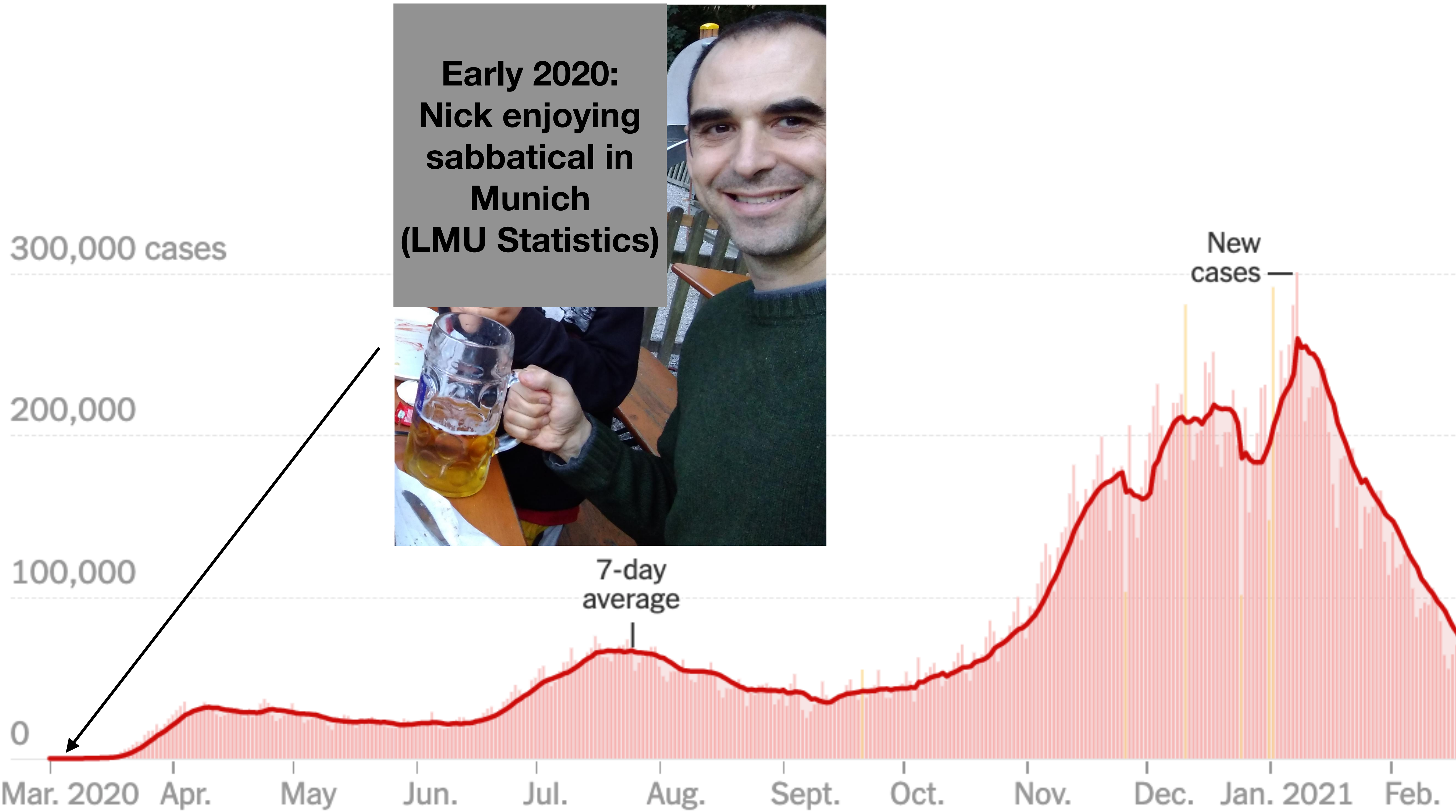


image credit: NY Times

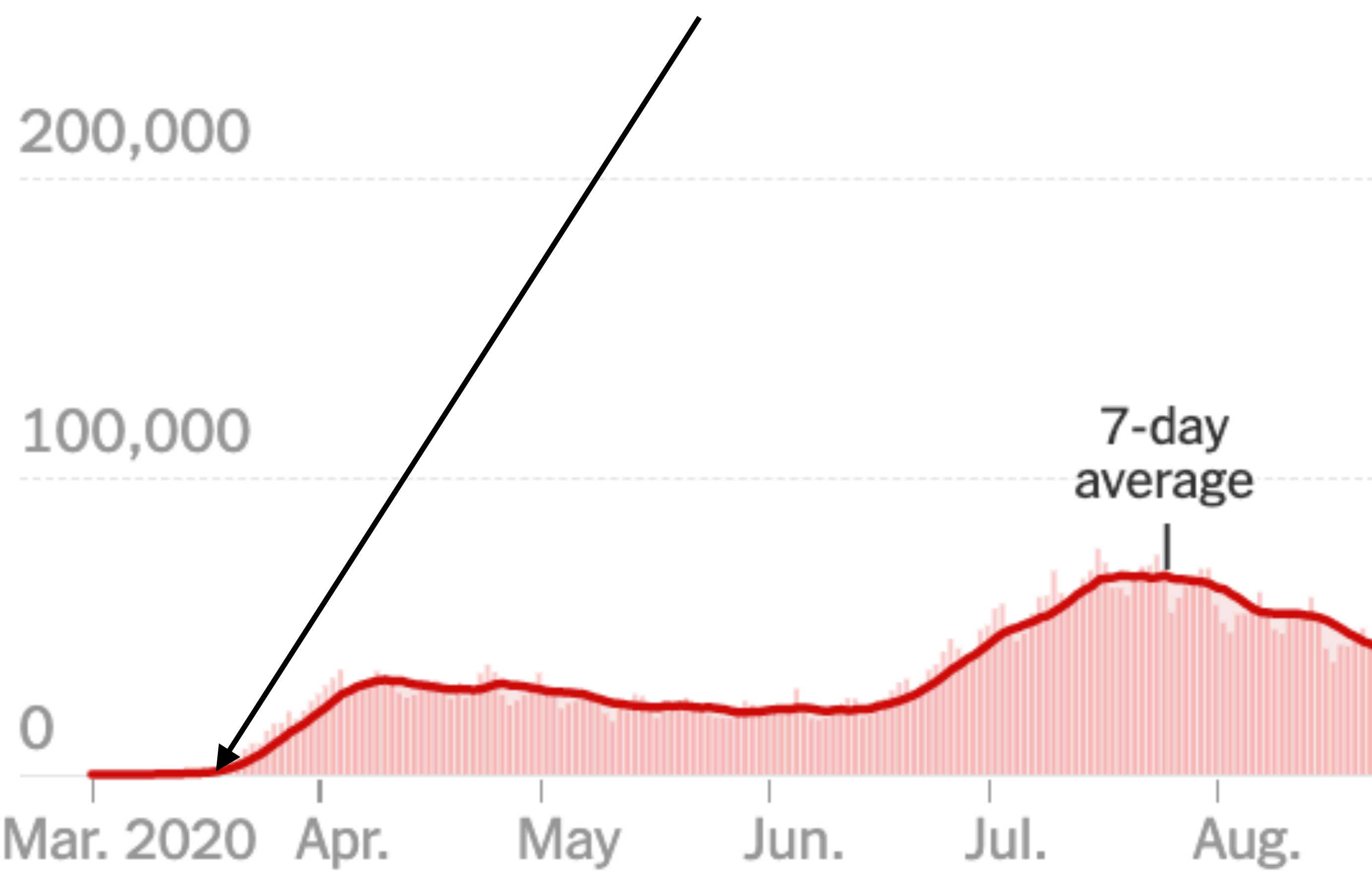
US COVID-19 Forecast Hub timeline


FiveThirtyEight

Politics Sports **Science** Podcasts Video

MAR. 20, 2020, AT 5:01 PM

Infectious Disease Experts Don't Know How Bad The Coronavirus Is Going To Get, Either



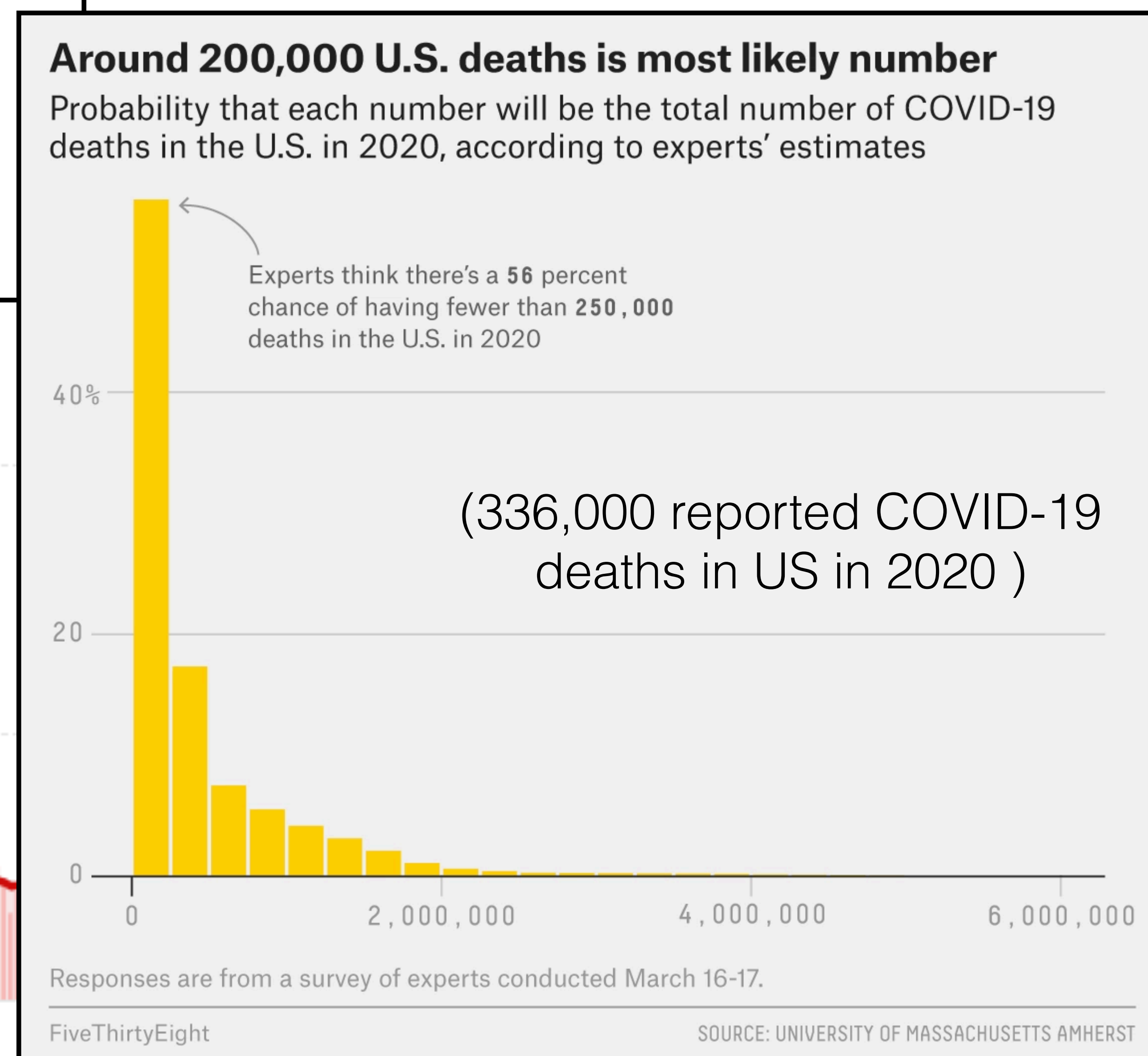
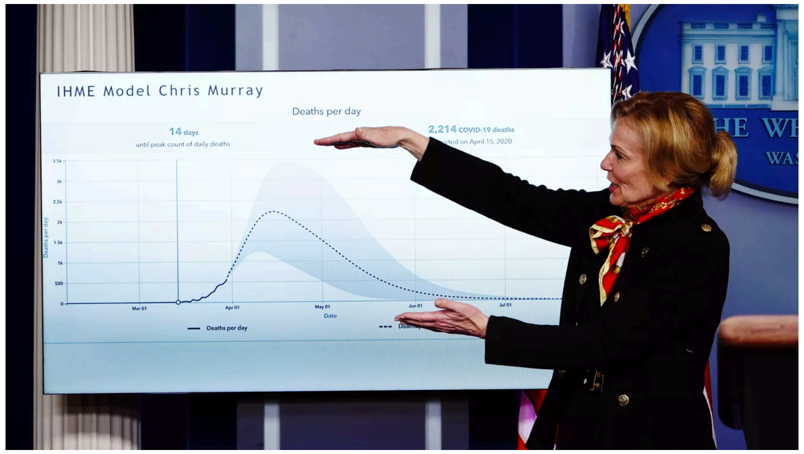
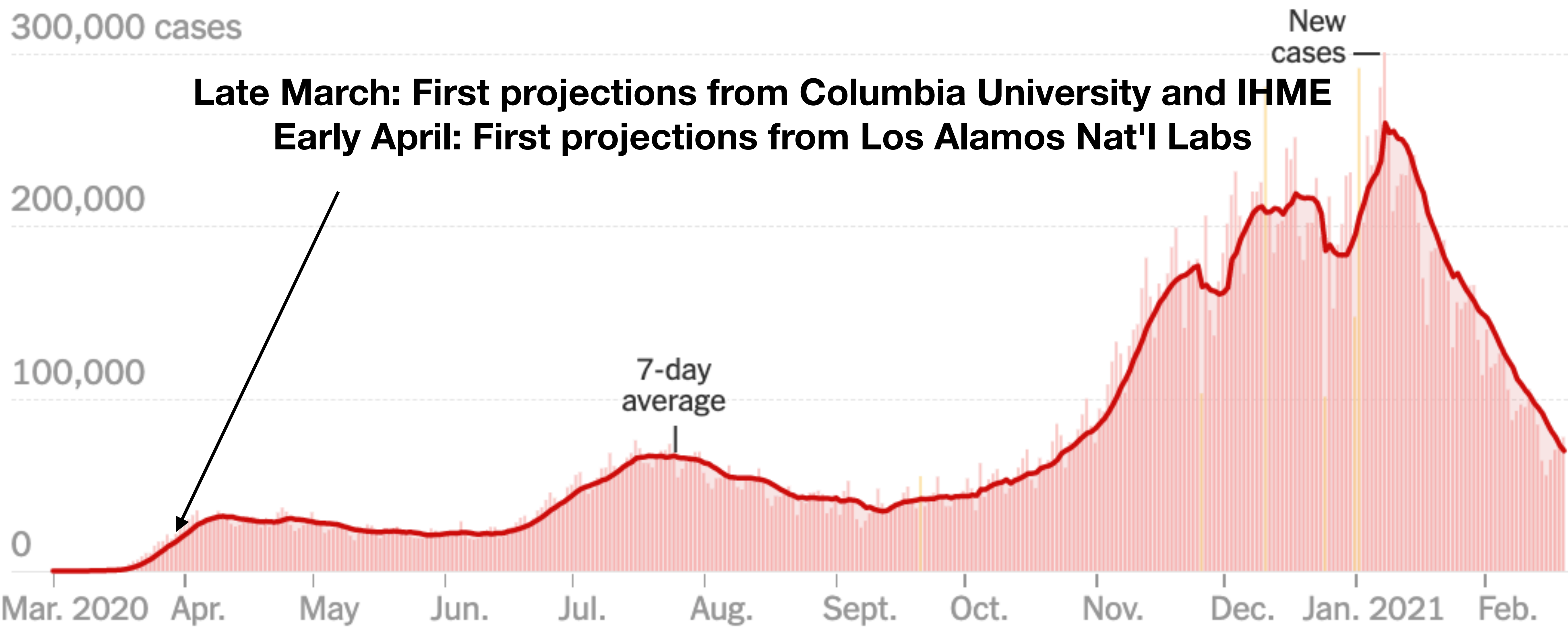


image credit: NY Times

US COVID-19 Forecast Hub timeline





Late March: First projections from Columbia University and IHME
Early April: First projections from Los Alamos Nat'l Labs

image credit: NY Times

US COVID-19 Forecast Hub timeline

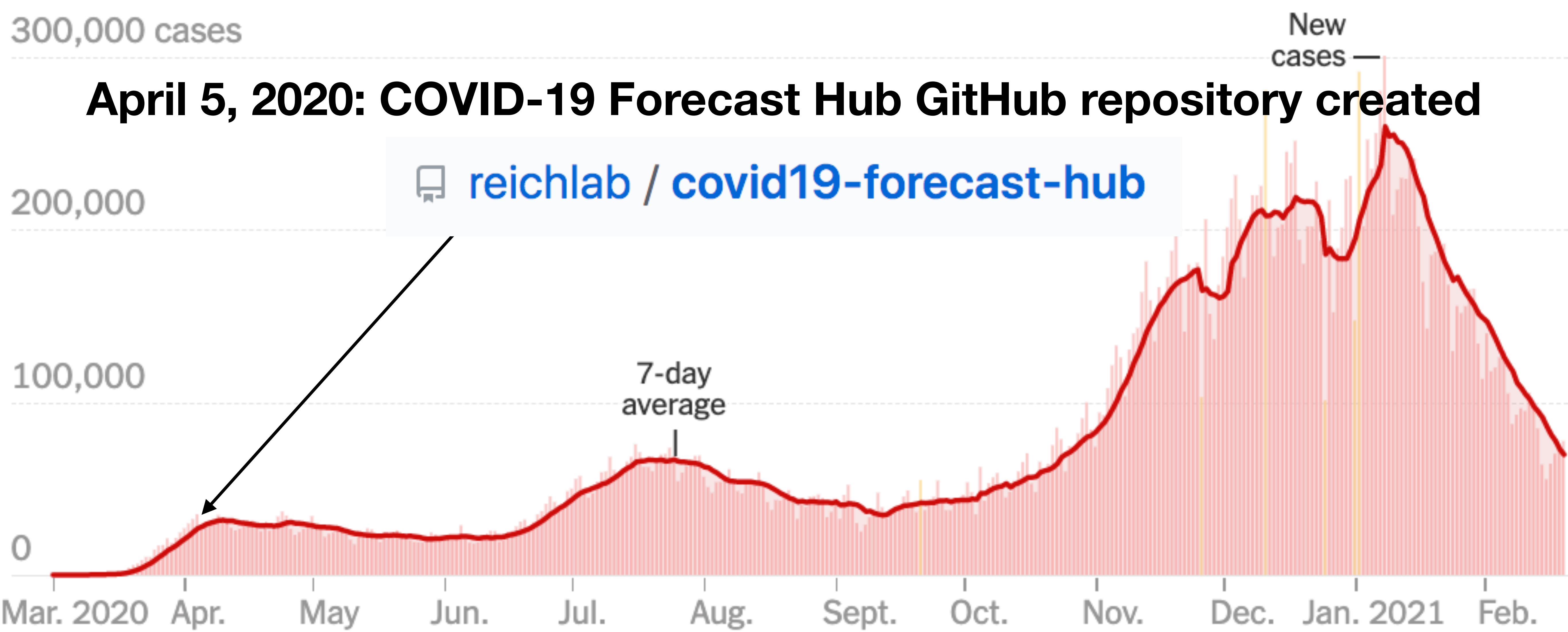


image credit: NY Times

US COVID-19 Forecast Hub timeline

April 13, 2020:
first Hub forecasts
displayed on CDC website.
3 teams.

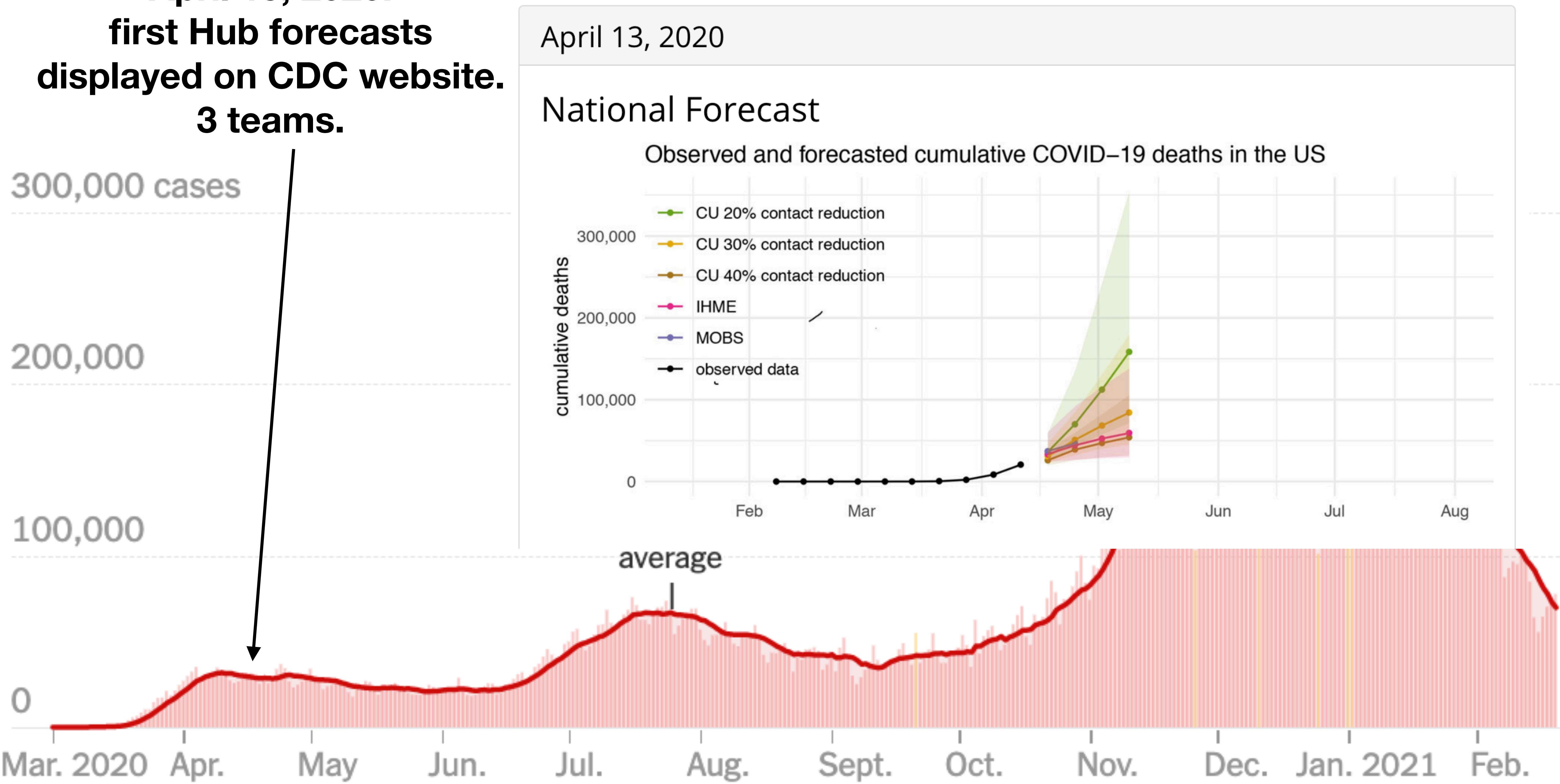


image credit: NY Times

US COVID-19 Forecast Hub timeline

April 27, 2020: first Hub ensemble forecast on CDC website. 6 teams.

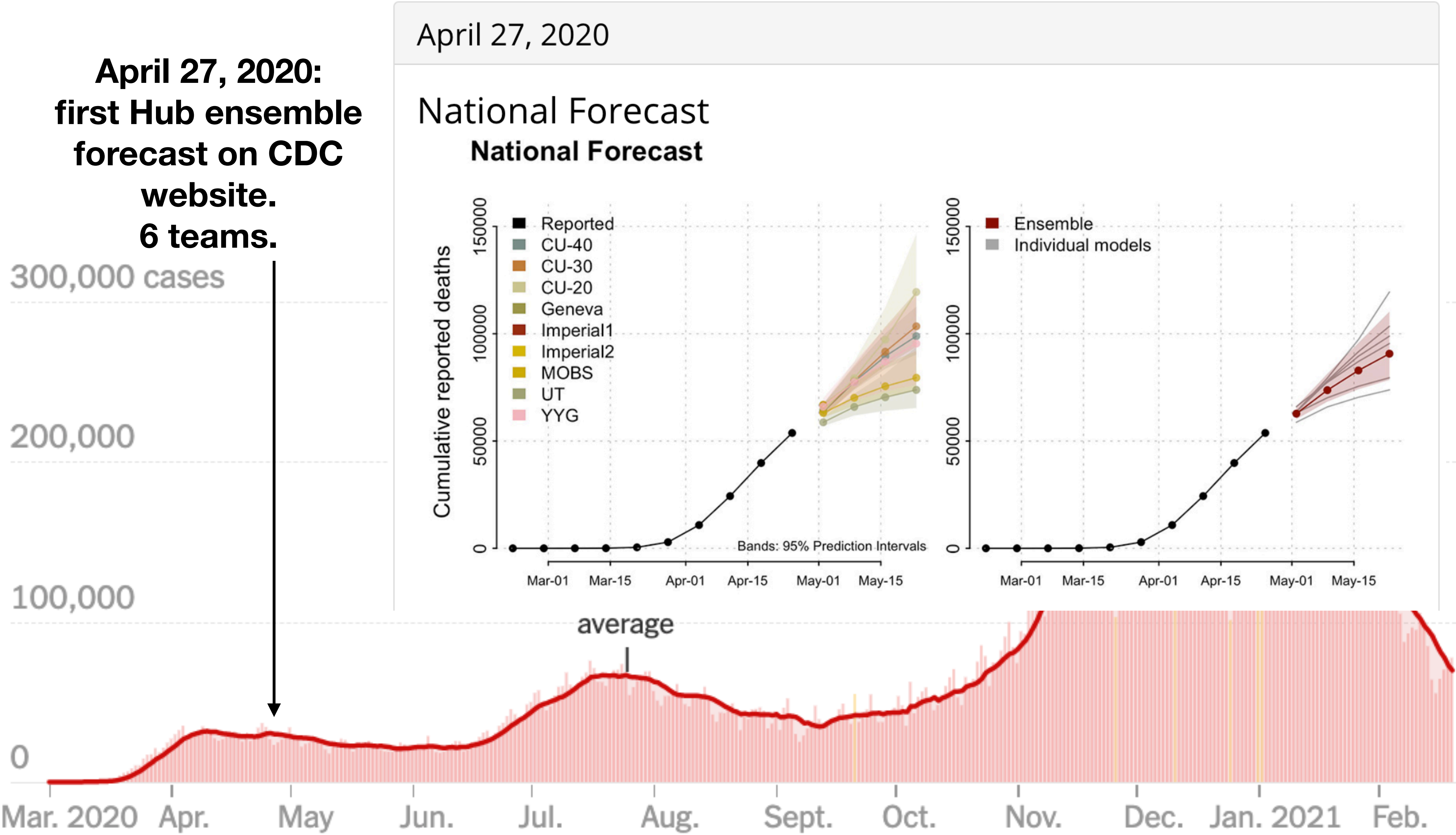


image credit: NY Times

US COVID-19 Forecast Hub timeline

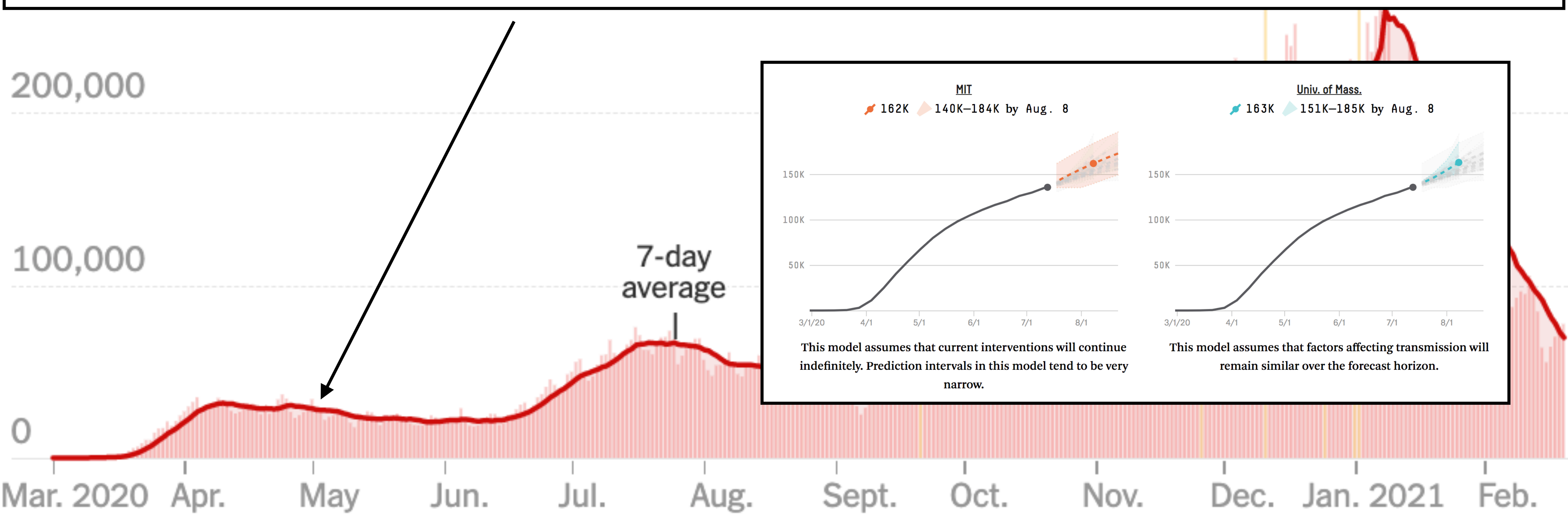
May 1, 2020 (updated daily ever since)

FiveThirtyEight

UPDATED JUL. 14, 2020, AT 8:00 AM

Where The Latest COVID-19 Models Think We're Headed — And Why They Disagree

By Ryan Best and Jay Boice

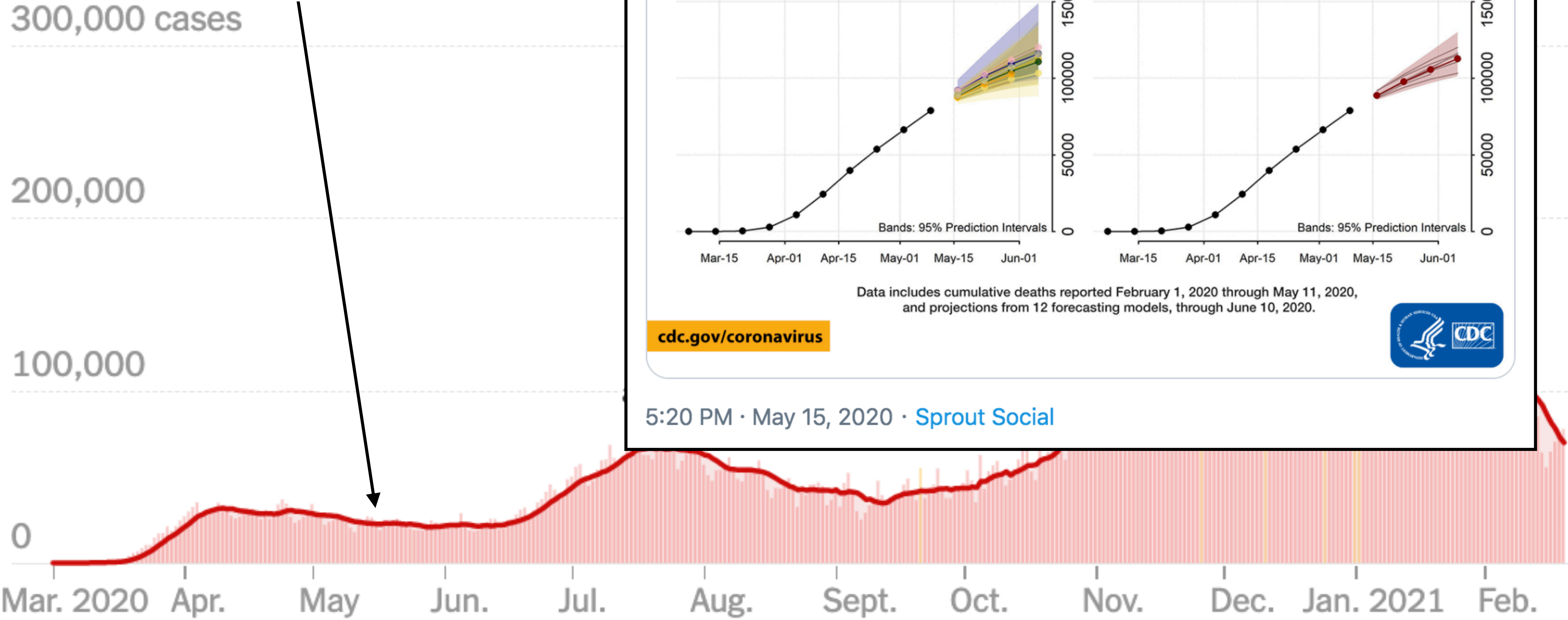


https://projects.fivethirtyeight.com/covid-forecasts/

image credit: NY Times

US COVID-19

May 15, 2020:
CDC Director Redfield tweets out results. CDC forecasting website viewed by >1m visitors.
11 teams.



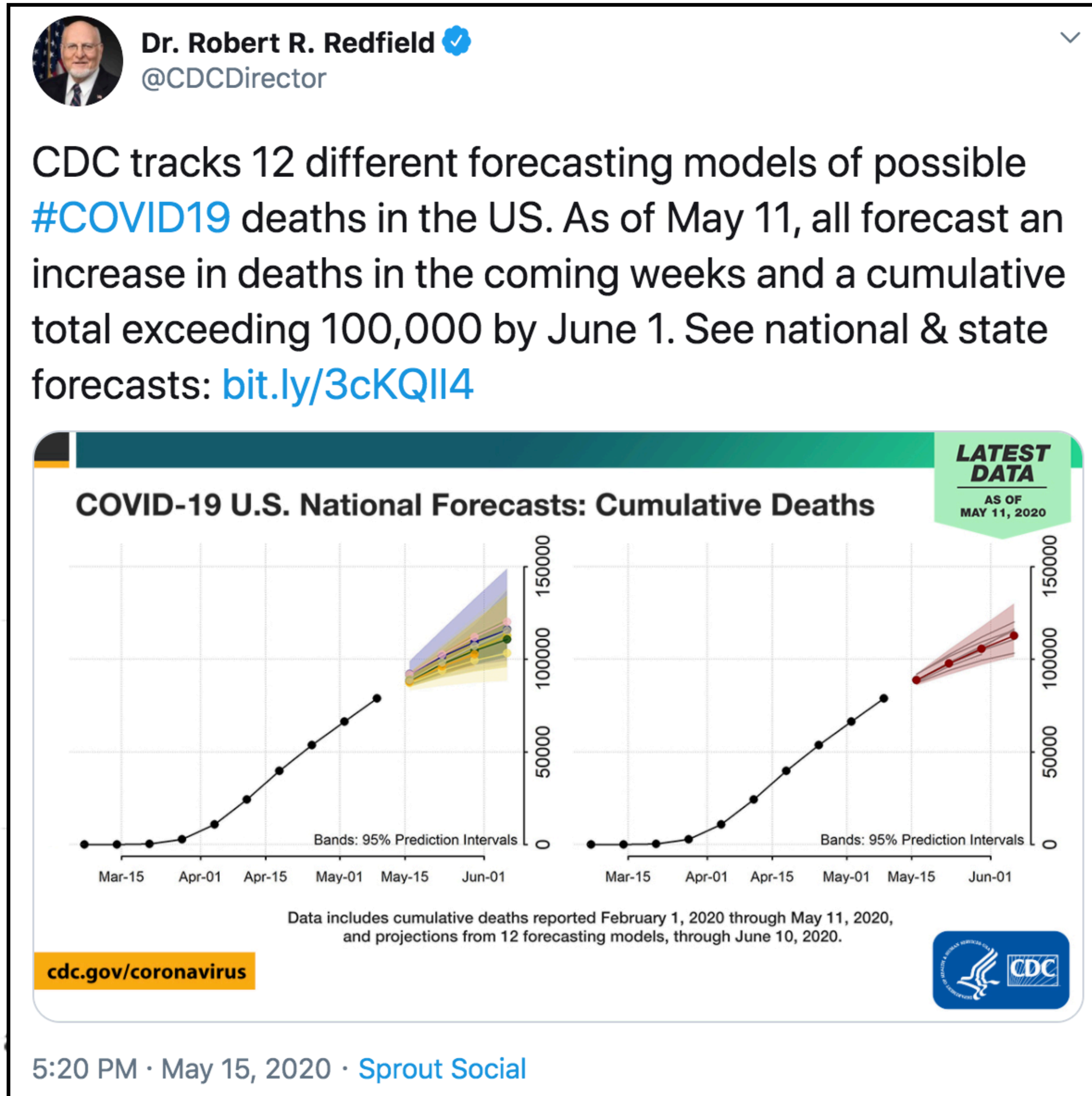


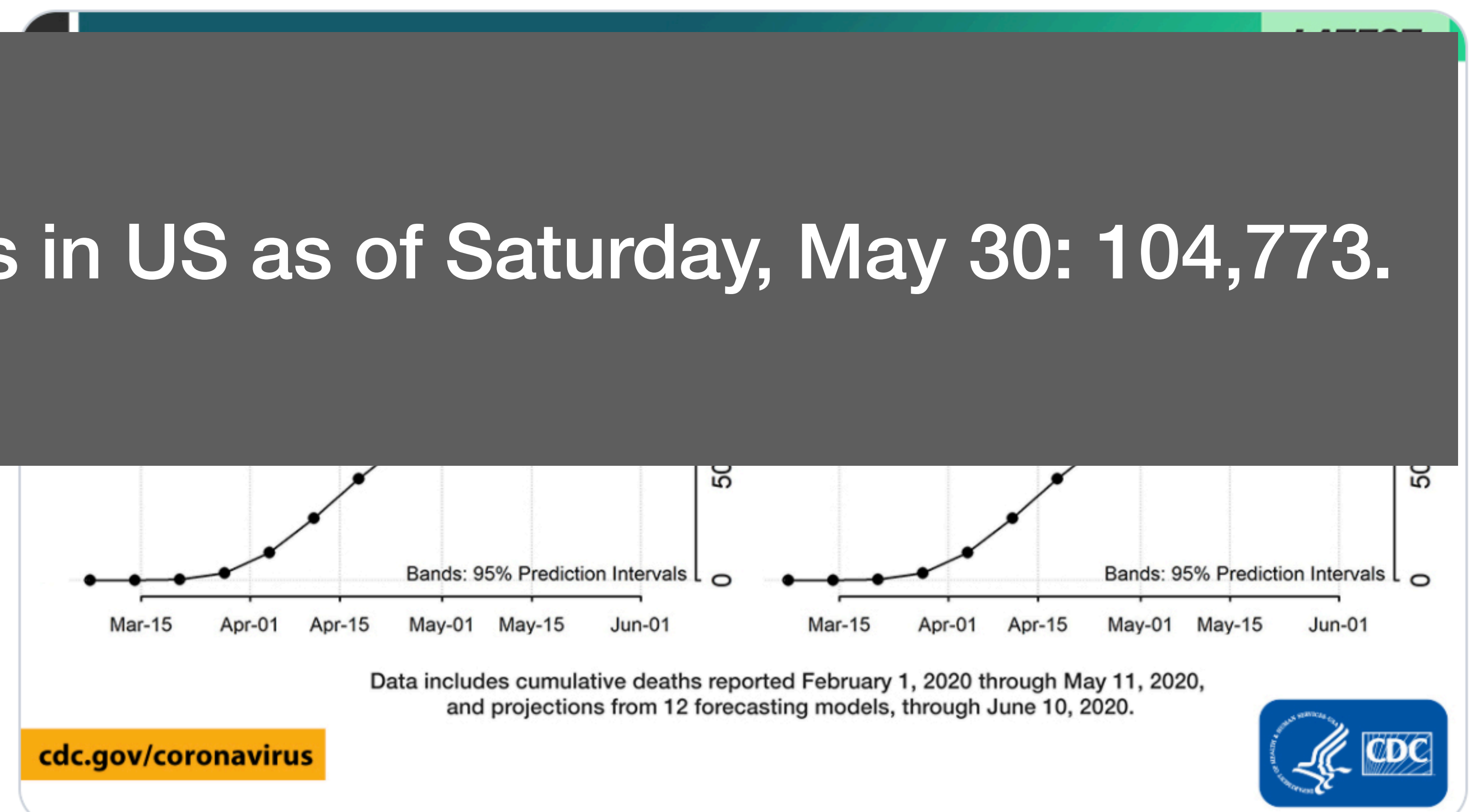
image credit: NY Times

US COVID-19

May 15, 2020:
CDC Director Redfield tweets out results. CDC forecasting website

Dr. Robert R. Redfield
@CDCDirector

CDC tracks 12 different forecasting models of possible #COVID19 deaths in the US. As of May 11, all forecast an increase in deaths in the coming weeks and a cumulative total exceeding 100,000 by June 1. See national & state forecasts: bit.ly/3cKQlI4



5:20 PM · May 15, 2020 · Sprout Social

Total cumulative deaths in US as of Saturday, May 30: 104,773.

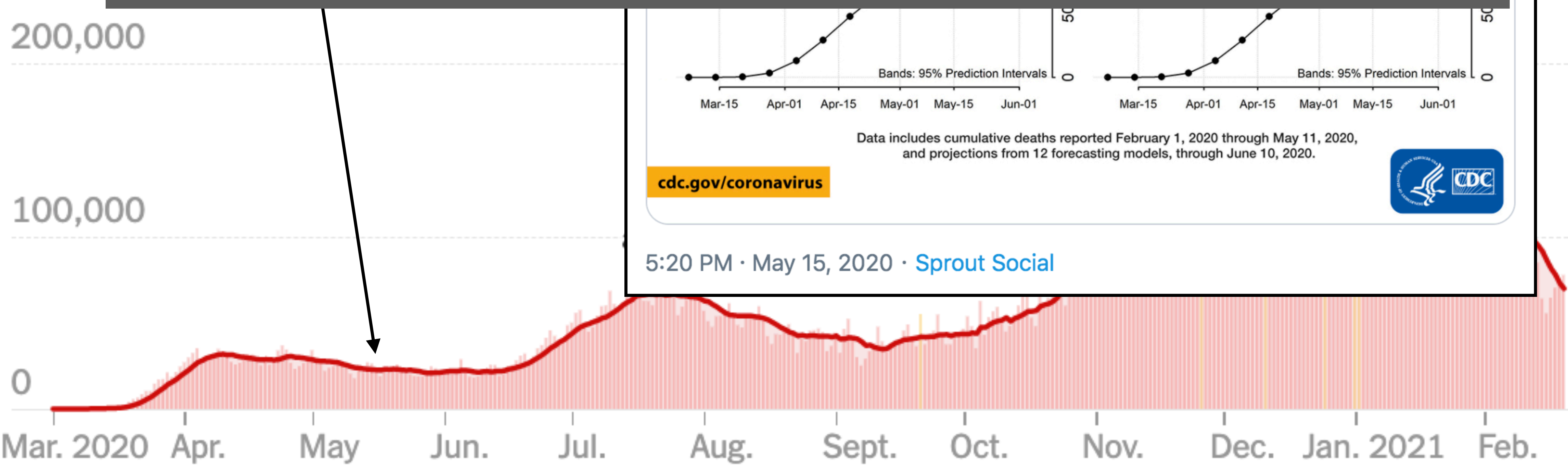


image credit: NY Times

US COVID-19 Forecast Hub timeline

June 4, 2020: 20 teams

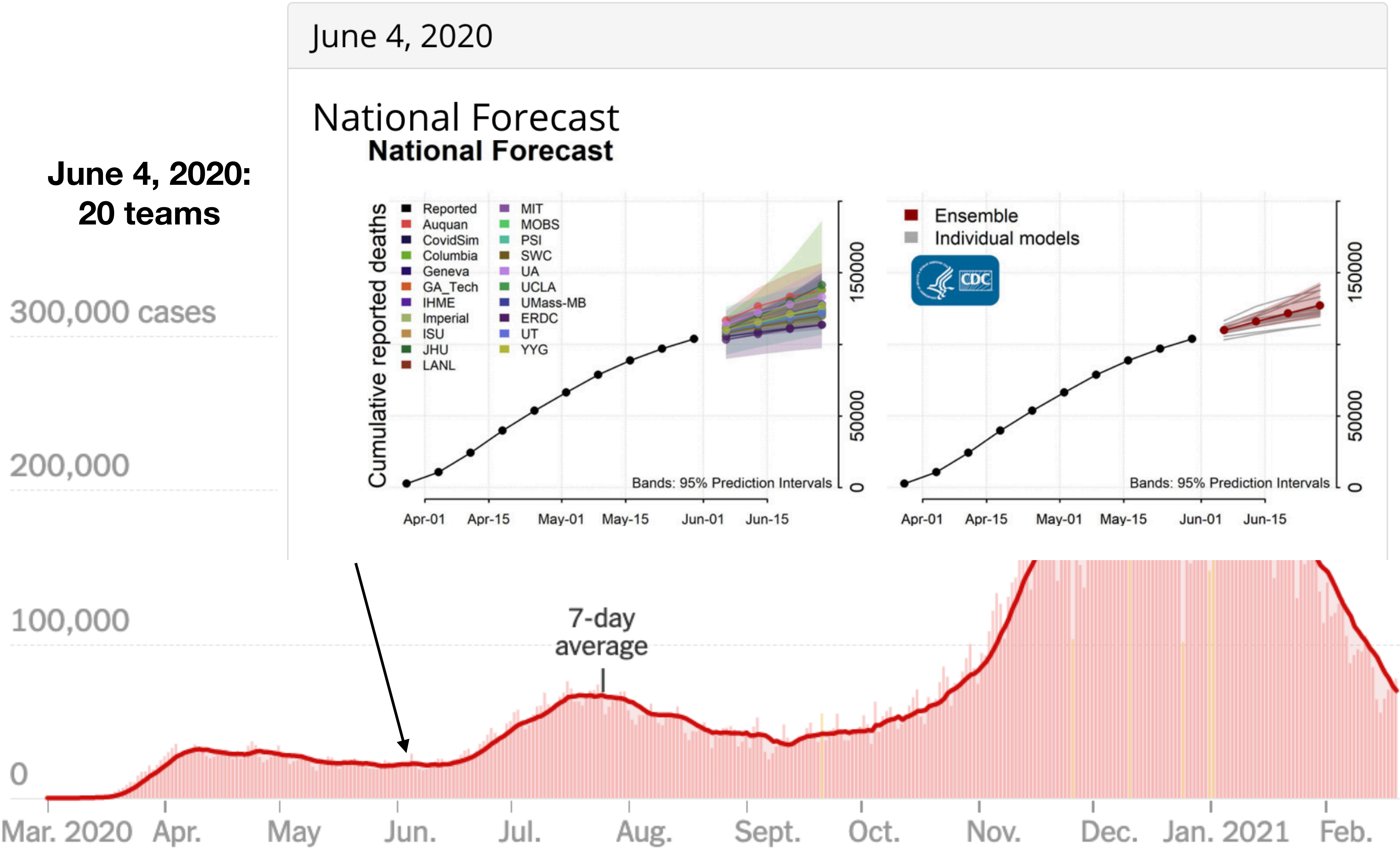
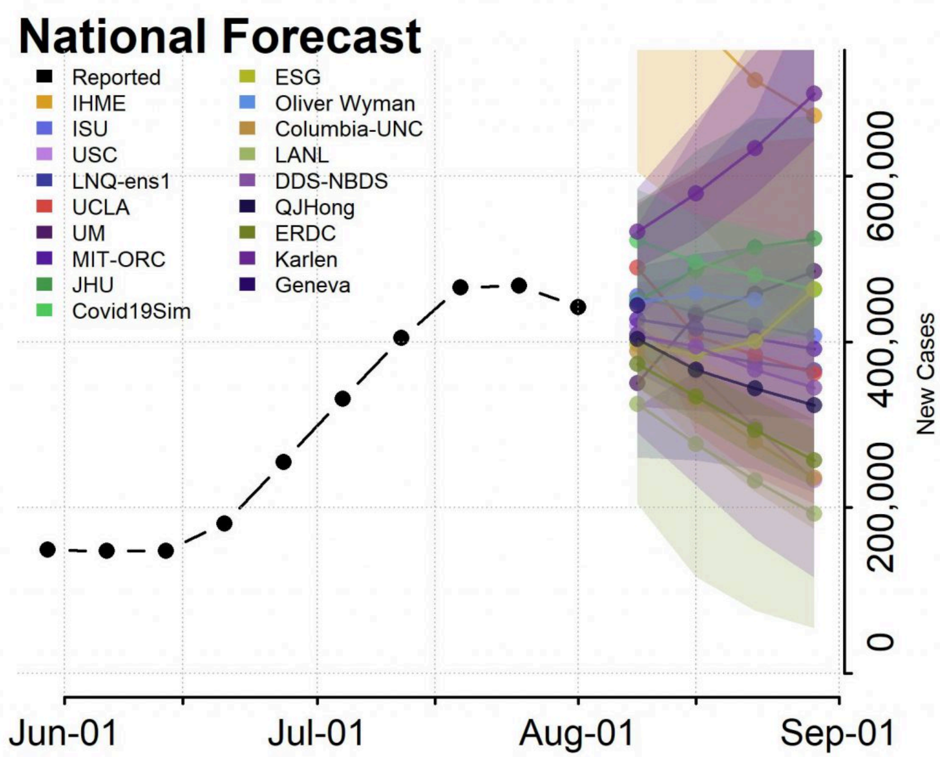
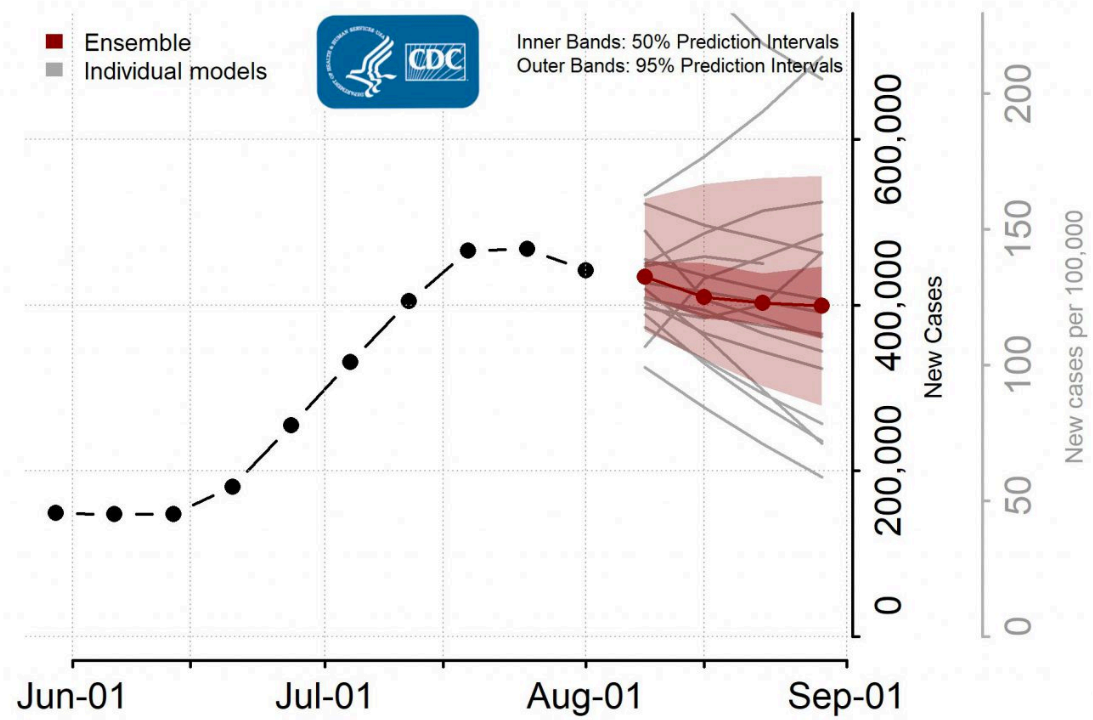


image credit: NY Times

US COVID-19 Forecast Hub timeline

August 6, 2020: 18 teams submit in first week of case forecasts. 53 models submit in all.





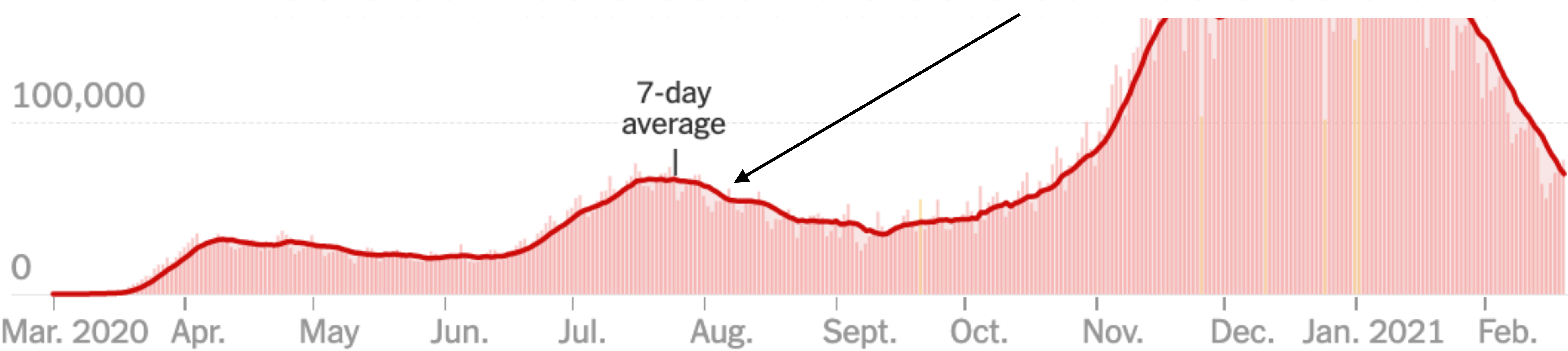


image credit: NY Times

US COVID-19 Forecast Hub timeline

December 9, 2020:
8 teams submit in first week of hospitalization forecasts with data.
58 models submit in all.

image credit: NY Times

US COVID-19 Forecast Hub timeline

Friday, December 18, 2020

PENN STATE NEWS

HOME | RESEARCH | ACADEMICS | IMPACT | CAMPUS LIFE | ATHLETICS | ADMINISTRATION | ARTS AND ENTERTAINMENT

Penn State to begin spring semester remotely, delay in-person classes

In-person classes will begin Feb. 15 following remote period; students must complete COVID-19 testing if returning to campus community

December 18, 2020:
23 campuses delay opening after university president looked at ensemble forecasts with faculty epidemiologist.
(personal communication from Matt Ferrari)

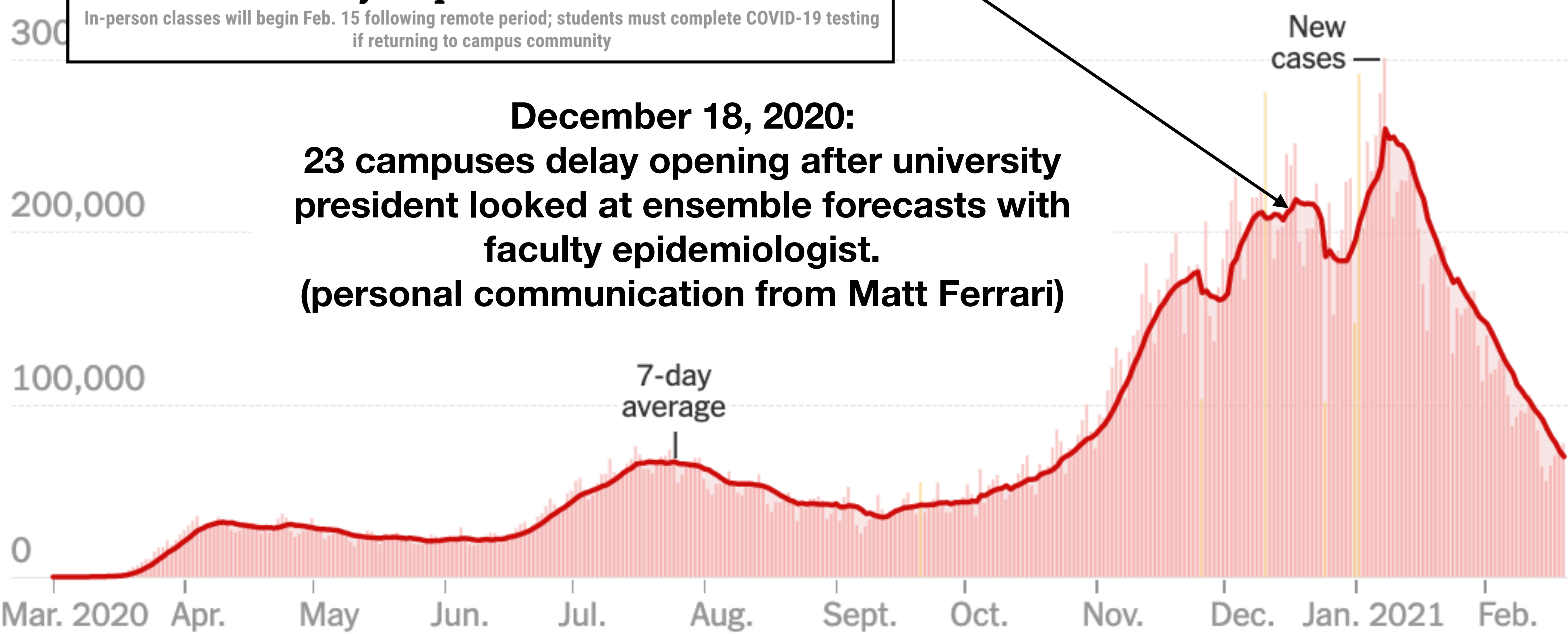


image credit: NY Times

US COVID-19 Forecast Hub timeline

CNN health | Food | Fitness | Wellness | Parenting | Vital Signs | • LIVE TV | Edition

Biden says Covid-19 death toll will likely top 500,000 next month, asks Americans to wear mask for his first 100 days

By **Dakin Andone** and **Madeline Holcombe**, CNN

Updated 4:14 PM ET, Thu January 21, 2021

Biden's warning of tens of thousands more deaths in just a few weeks echoes the latest ensemble forecast by the US Centers for Disease Control and Prevention, which projects the death toll could reach up to 508,000 by February 13.

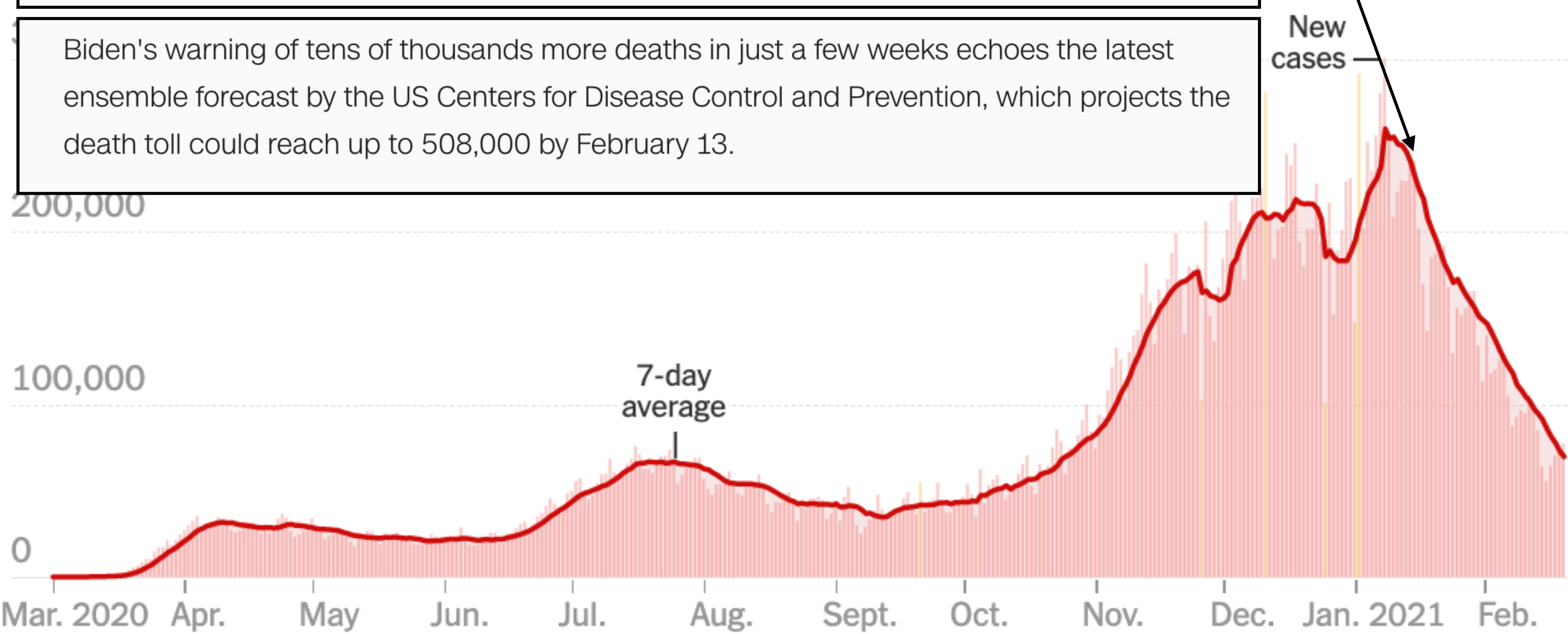


image credit: NY Times

What is the day-to-day?

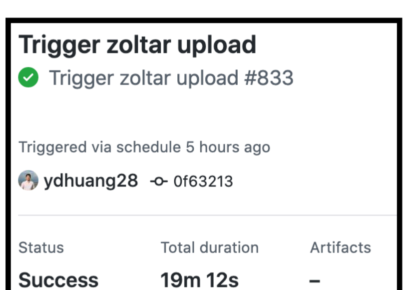
Forecast submission process

Research group submits Pull Request (PR) on GitHub

Automated checks and validations triggered

If needed, resubmit until any issues resolved

PR merged by Hub team member



forecasts batch uploaded to Zoltar SQL database.

Ensemble creation process

1. 6pm ET Monday: all forecasts submitted before the deadline are merged.

2. Current version of repository is "tagged" for retrospective analysis.

3. Ensemble script is triggered by a Hub team member: 30 minute run-time. (A separate longer-running script generates an experimental "trained" ensemble.)

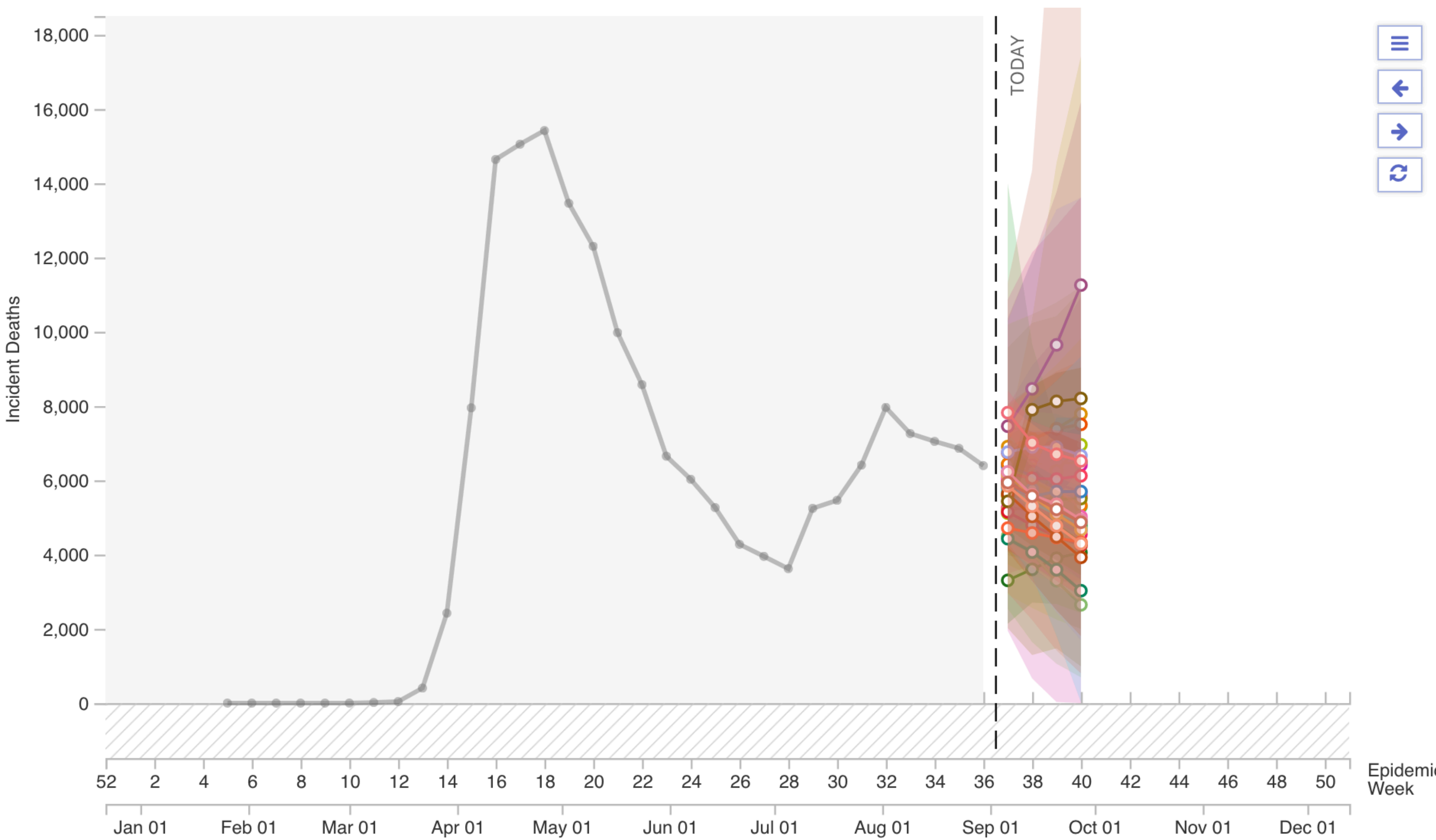
4. Resulting plots are examined by humans.

5. PR created with ensemble forecast to GitHub repo, merged after checks pass.

6. Interactive visualization and weekly reports are updated and deployed.

Demo Visualization

https://viz.covid19forecasthub.org/



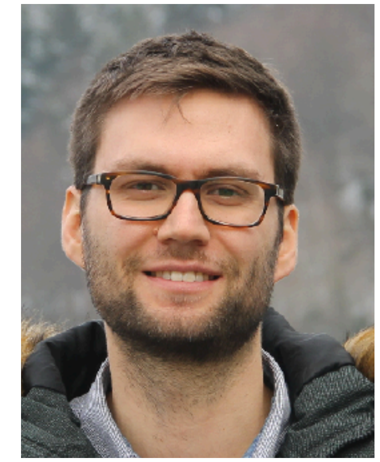
faculty co-leads

Nick

Evan

COVID-19 ForecastHub

external collaborators

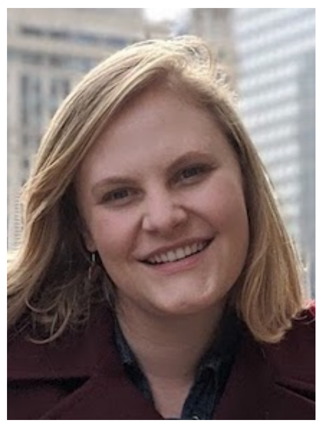


Johannes

Jarad

Students

Nutcha

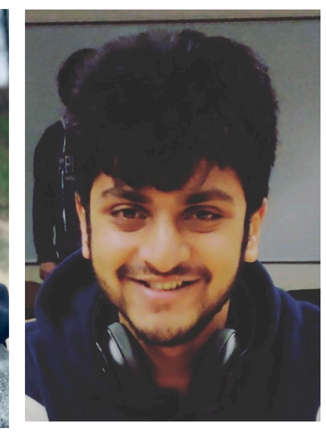


Estee

Ariane

Alvaro

Aaron

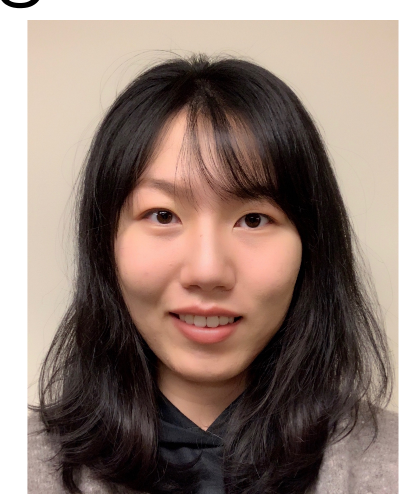


Apurv

Dasuni

Programmers and data analysts (* = full time)

Matt*



Serena*

Yuxin*

Khoa

Abdul

Martha

COVID-19 ForecastHub

https://covid19forecasthub.org/

Team: Martha Zorn, Nutcha Wattanachit, Serena Wang, Ariane Stark, Apurv Shah, Nicholas Reich, Evan Ray, Jarad Niemi, Khoa Le, Abdul Kanji, Dasuni Jayawardena, Yuxin Huang, Katie House, Aaron Gerding, Estee Cramer, Matt Cornell, Alvaro J. Castro Rivadeneira, Andrea Brennen, Johannes Bracher

* underline denotes ensemble contributor

US CDC Collaborators: Matthew Biggerstaff, Michael Johansson, Velma Lopez, Rachel Slayton, Jo Walker

Ensemble “advisors": Jacob Bien, Logan Brooks, Sebastian Funk, Tilmann Gneiting, Anja Muhlemann, Aaron Rumack, Ryan Tibshirani

Modeling groups: Over 80 (!!) groups at various institutions have contributed forecasts to the hub

COVID-19 ForecastHub

https://covid19forecasthub.org/

Hub Citations (https://covid19forecasthub.org/doc/research/)

- Cramer EY, Ray EL, Lopez VK, et al. “Evaluation of individual and ensemble probabilistic forecasts of COVID-19 mortality in the US.” 2021. *medRxiv*. (preprint). https://www.medrxiv.org/content/10.1101/2021.02.03.21250974v1

- Bracher J, Ray EL, Gneiting T, Reich NG. “Evaluating epidemic forecasts in an interval format.” 2021. *PLOS Comp Bio:* 17 (2), e1008618. https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1008618

- NG Reich, M Cornell, EL Ray, K House, K Le. The Zoltar forecast archive, a tool to standardize and store interdisciplinary prediction research. 2021. *Scientific Data:* 8 (1), 1-11. https://www.nature.com/articles/s41597-021-00839-5

- Logan C. Brooks, Evan L. Ray, Jacob Bien, Johannes Bracher, Aaron Rumack, Ryan J. Tibshirani, Nicholas G. Reich. “Comparing ensemble approaches for short-term probabilistic COVID-19 forecasts in the U.S.” 2020. International Institute of Forecasters Blog. https://forecasters.org/blog/2020/10/28/comparing-ensemble-approaches-for-short-term-probabilistic-covid-19-forecasts-in-the-u-s/

- Evan L Ray, Nutcha Wattanachit, et al. “Ensemble Forecasts of Coronavirus Disease 2019 (COVID-19) in the U.S.” 2020. *medRxiv*. (preprint) https://www.medrxiv.org/content/10.1101/2020.08.19.20177493v1

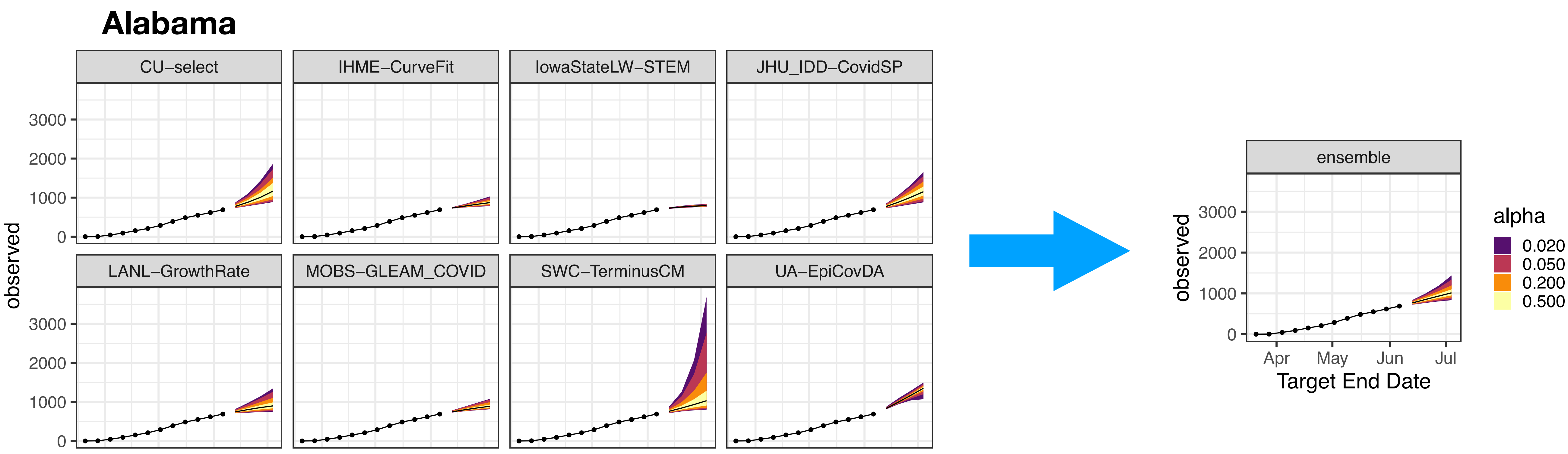
COVID-19 ForecastHub

Thank you!

Reich Lab | AT UMASS AMHERST

building the ensemble

Building the Ensemble: View 1



- For each combination of spatial unit s, time point t, and forecast horizon h, teams are required to submit K=23 quantiles of a predictive distribution:

$$\widehat{P}(Y \le q^m_{s,t,h,1}) = 0.01,\ \widehat{P}(Y \le q^m_{s,t,h,2}) = 0.025,\ \ldots,\ \widehat{P}(Y \le q^m_{s,t,h,12}) = 0.5,\ \ldots,\ \widehat{P}(Y \le q^m_{s,t,h,23}) = 0.99$$

The predictive median

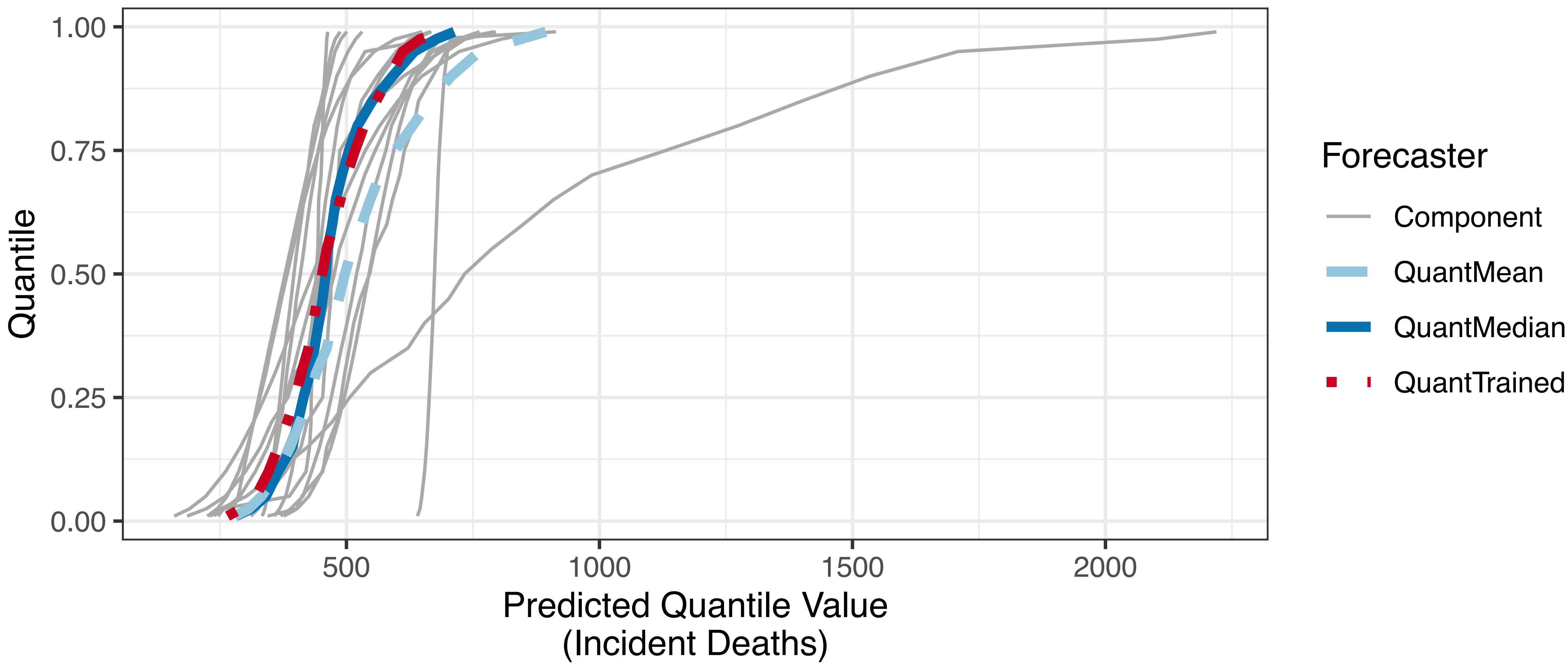
Limits of a 98% prediction interval

- The predictive quantiles for the ensemble are a combination of component predictions at each quantile level:

$q_{s,t,h,k} = f(q^1_{s,t,h,k}, \ldots, q^M_{s,t,h,k})$ **for each** $k = 1,\ldots,23$

Building an Ensemble: View 2

- The pairs $\left(q^m_{s,t,h,k}, \widehat{P}(Y^m_{s,t,h} \leq q^m_{s,t,h,k})\right)$ fall along the predictive CDF for model m



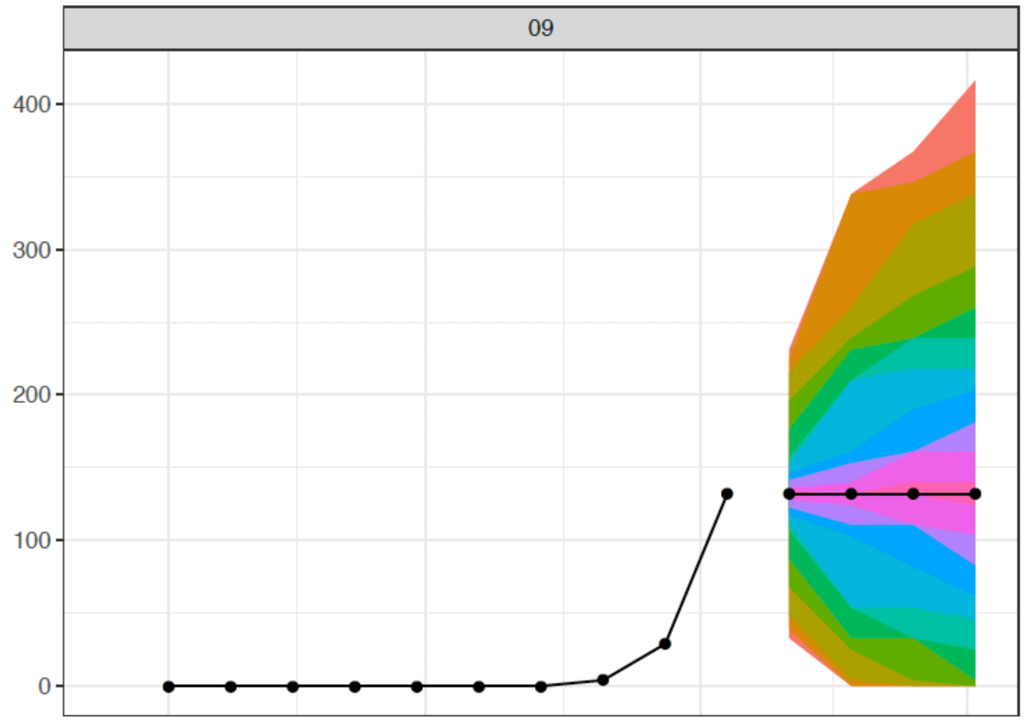
- Three options for the combination function f:
  - QuantMean: $$q_{s,t,h,k} = \frac{1}{M}\sum_{m=1}^{M} q^m_{s,t,h,k}$$ **Used through July 21, 2020**
  - QuantMedian: $$q_{s,t,h,k} = \textbf{median}(q^1_{s,t,h,k}, \ldots, q^M_{s,t,h,k})$$ **Used starting July 28, 2020**
  - QuantTrained: $$q_{s,t,h,k} = \beta^0_{t,h,k} + \sum_{m=1}^{M} \beta^m_{t,h,k} \cdot q^m_{s,t,h,k}$$ **Evaluated, not released each week**

building the baseline

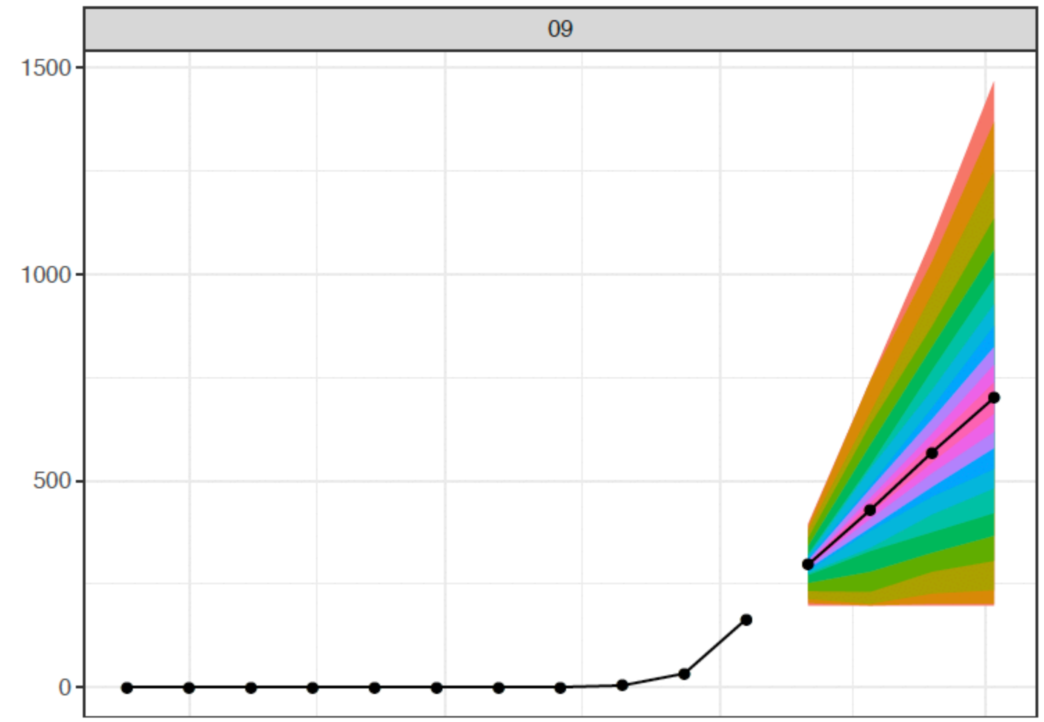
Baseline Model

- Different from flu forecasting baseline model! Not "seasonally" driven.
- Acknowledgment: idea adapted from a suggestion by Ryan Tibshirani (CMU).
- Goal: Median predicted incidence is most recent observed incidence.
- Predictions of cumulative deaths derived from predictions of incident deaths.

Incident Deaths



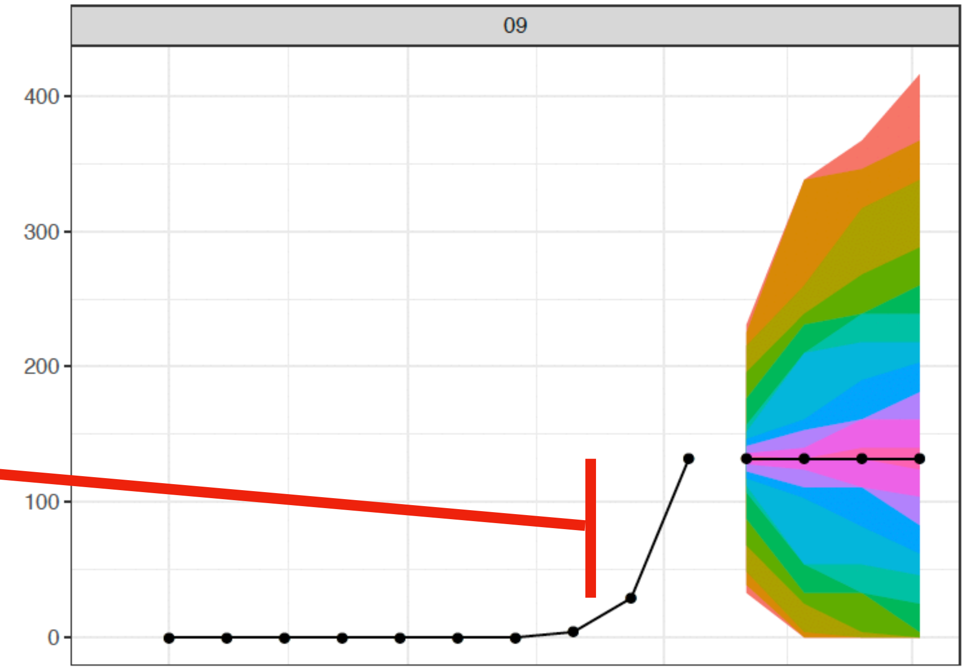
Cumulative Deaths



Baseline Model

- Procedure:
    - Compute first differences of historical incidence:

$d_t = y_t - y_{t-1}$



- Collect first differences and their negatives
- Sample first differences and add to last observed incidence; take quantiles of the resulting distribution
- Iterate for horizons > 1
- Adjustments for "niceness":
    - Force median = last observed incidence
    - Truncate at 0

measuring accuracy

Forecast Skill: Weighted Interval Score

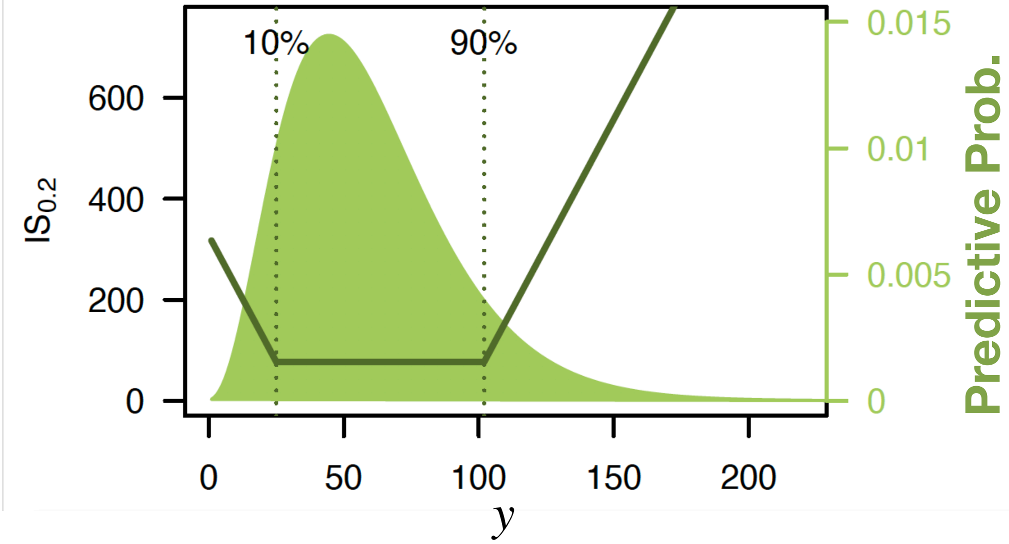
- A probabilistic version of the mean absolute error.
- A single number that measures a distance between an observed value and a predicted distribution.

Forecast Skill: Weighted Interval Score

- Consider a single $(1-\alpha)\times 100\,\%$ predictive interval [l, u] for the observed response y. The interval score is:

$$\mathbf{IS}_{\alpha}(F,y) = \underbrace{(u-l)}_{\text{Width of interval}} + \underbrace{\frac{2}{\alpha}\cdot(l-y)\cdot\mathbf{1}(y<l)}_{\text{Penalty if interval is too high}} + \underbrace{\frac{2}{\alpha}\cdot(y-u)\cdot\mathbf{1}(y>u)}_{\text{Penalty if interval is too low}},$$

Figure due to Johannes Bracher



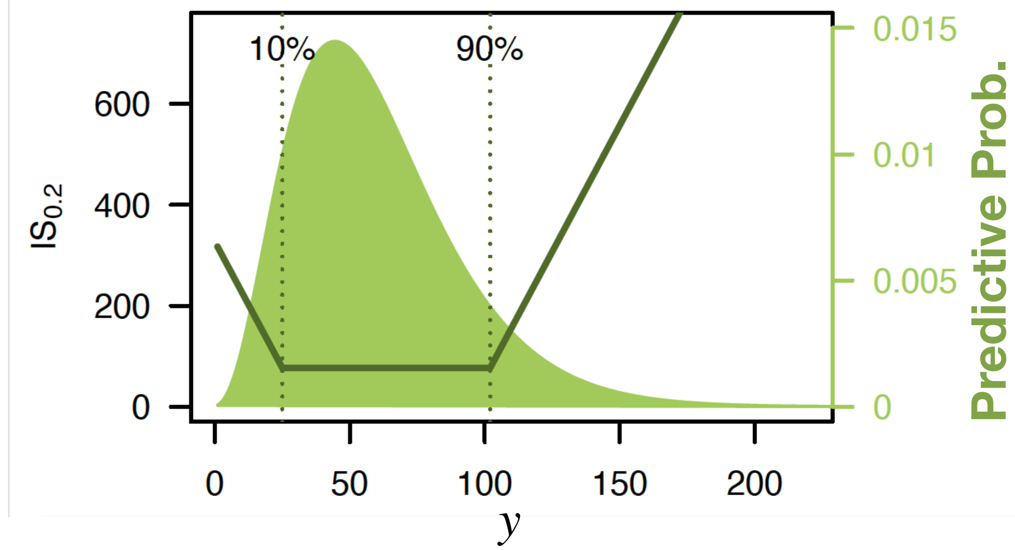
- Smaller $\mathbf{IS}_{\alpha}$ is better

Forecast Skill: Weighted Interval Score

- Consider a single $(1-\alpha) \times 100\,\%$ predictive interval [l, u] for the observed response y. The interval score is:

$$\mathbf{IS}_\alpha(F, y) = \underbrace{(u - l)}_{\text{Width of interval}} + \underbrace{\frac{2}{\alpha} \cdot (l - y) \cdot \mathbf{1}(y < l)}_{\text{Penalty if interval is too high}} + \underbrace{\frac{2}{\alpha} \cdot (y - u) \cdot \mathbf{1}(y > u)}_{\text{Penalty if interval is too low}},$$

Figure due to Johannes Bracher



- Smaller $\mathbf{IS}_\alpha$ is better
- For multiple predictive intervals, we compute a weighted average of $\mathbf{IS}_\alpha$

$$\mathbf{WIS}_{\alpha_{0:K}}(F, y) = \frac{1}{K+1} \times \left(w_0 \times 2 \times |y - m| + \sum_{k=1}^{K} \left(w_k \times \mathbf{IS}_{\alpha_k}(F, y) \right) \right).$$

- We use weights $w_i = \dfrac{\alpha_i}{2}$, in which case WIS ≈ CRPS (continuous ranked probability score)
- The resulting score is **proper**: in expectation, it is minimized by the true predictive distribution.
- See Bracher et al. (2020) for more: https://arxiv.org/abs/2005.12881