University of Massachusetts Amherst

Distinguished Faculty Lecture Series 2021

Featuring

NICHOLAS G. REICH

How to Forecast a Pandemic: Lessons from COVID-19

BE REVOLUTIONARY™

I'm telling you stories. Trust me.

from "The Passion" by Jeanette Winterson

This data-driven story is about our journey to learn more about epidemic modeling

- what approaches produce best predictions of outbreaks in general
- how reliable have models been in predicting COVID-19
- how can we use multiple models to improve predictions

I'm telling you stories. Trust me.

from "The Passion" by Jeanette Winterson

In God we trust; all others bring data.

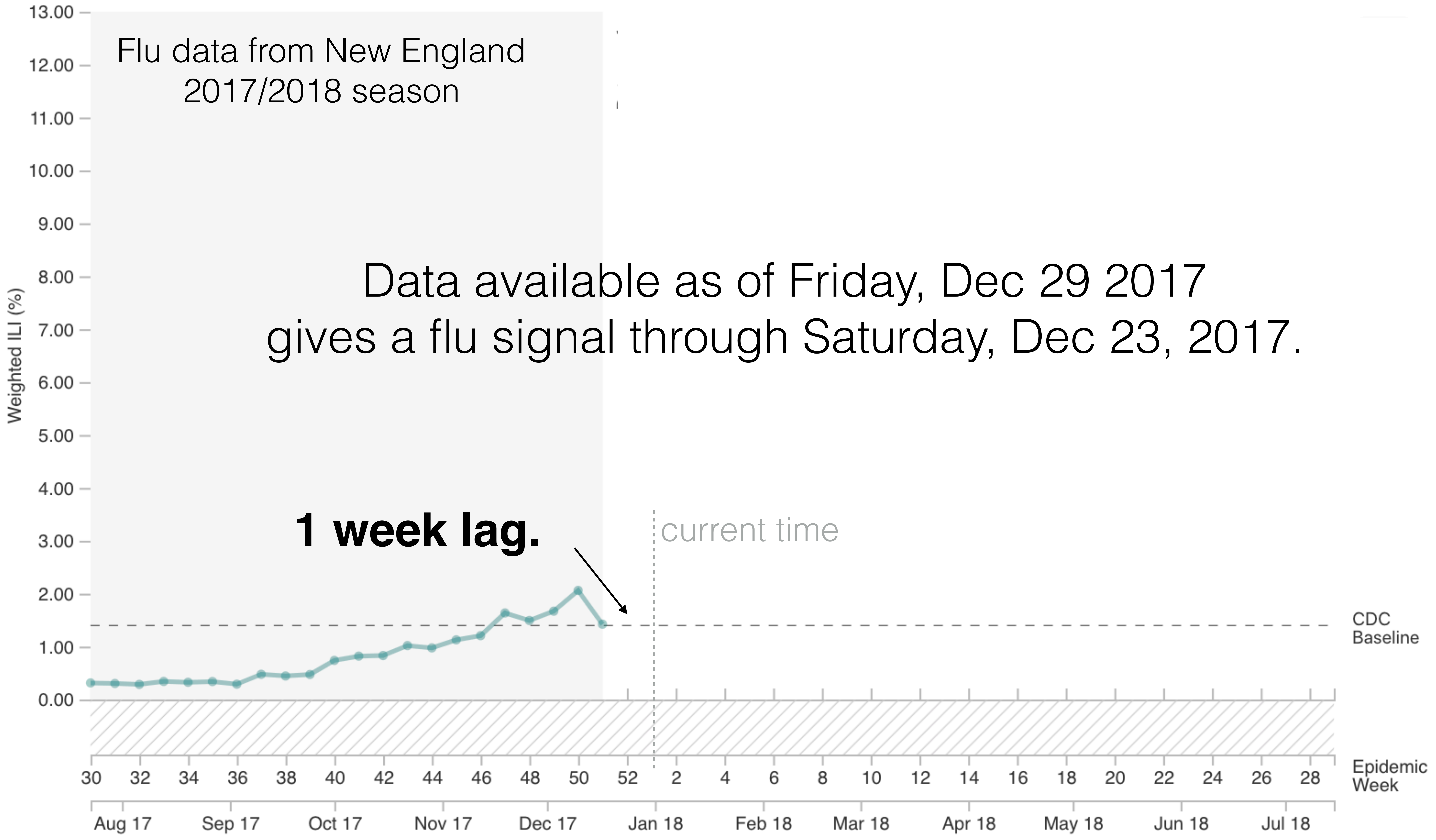
W. Edwards Deming

This data-driven story is about our journey to learn more about epidemic modeling

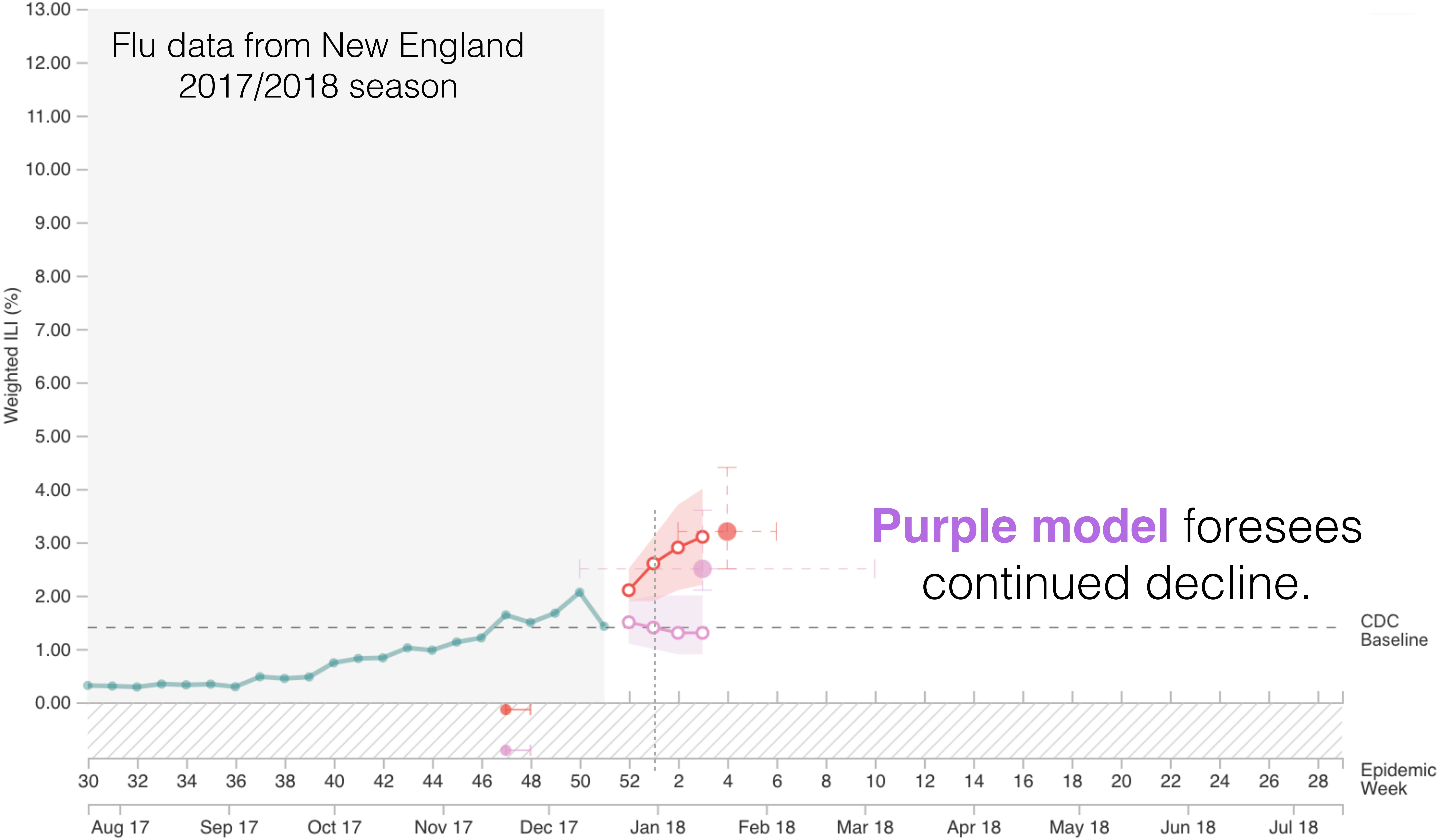
- what approaches produce best predictions of outbreaks in general
- how reliable have models been in predicting COVID-19
- how can we use multiple models to improve predictions

Preface

Modeling real-time surveillance data

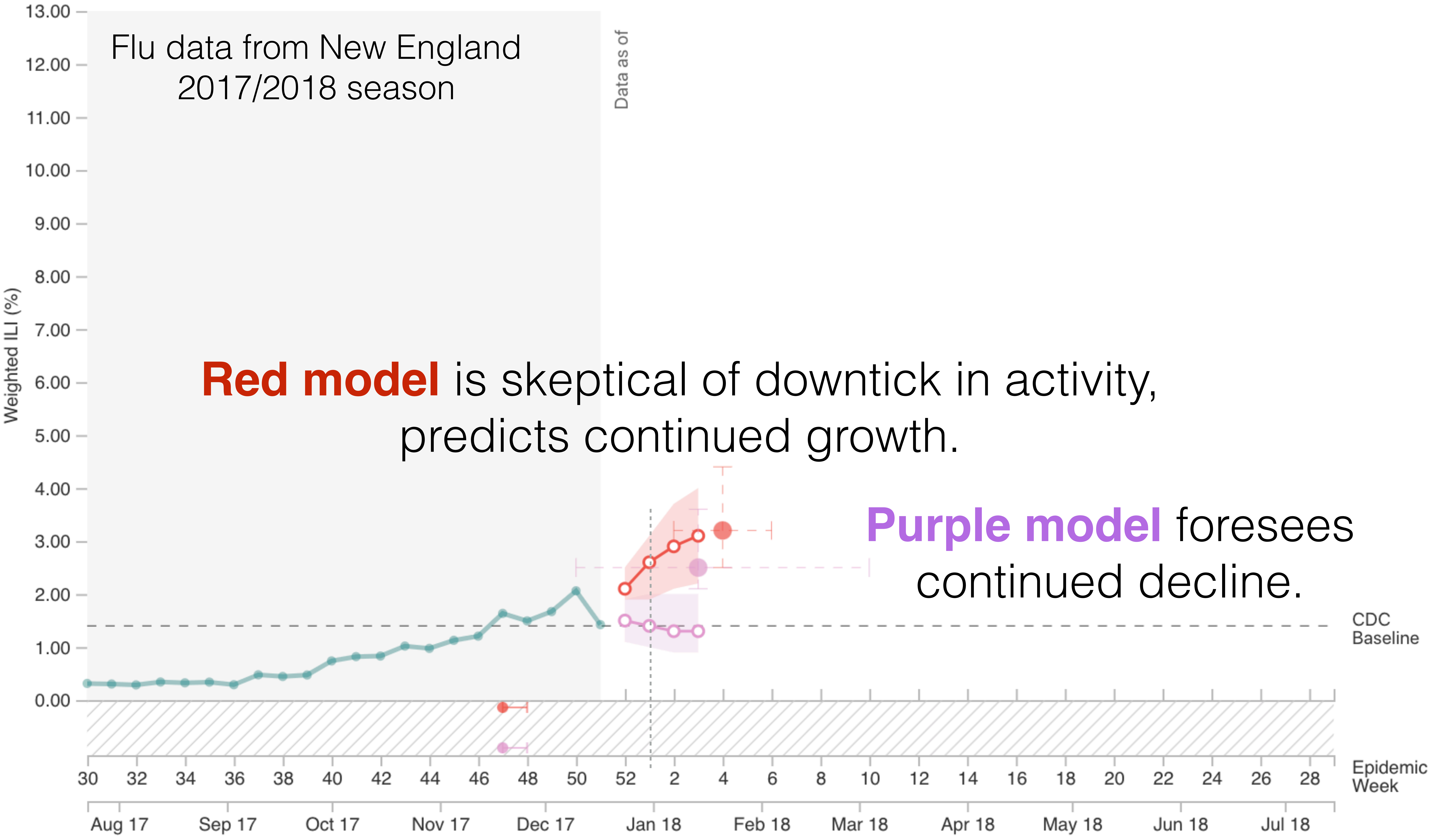


Modeling real-time surveillance data

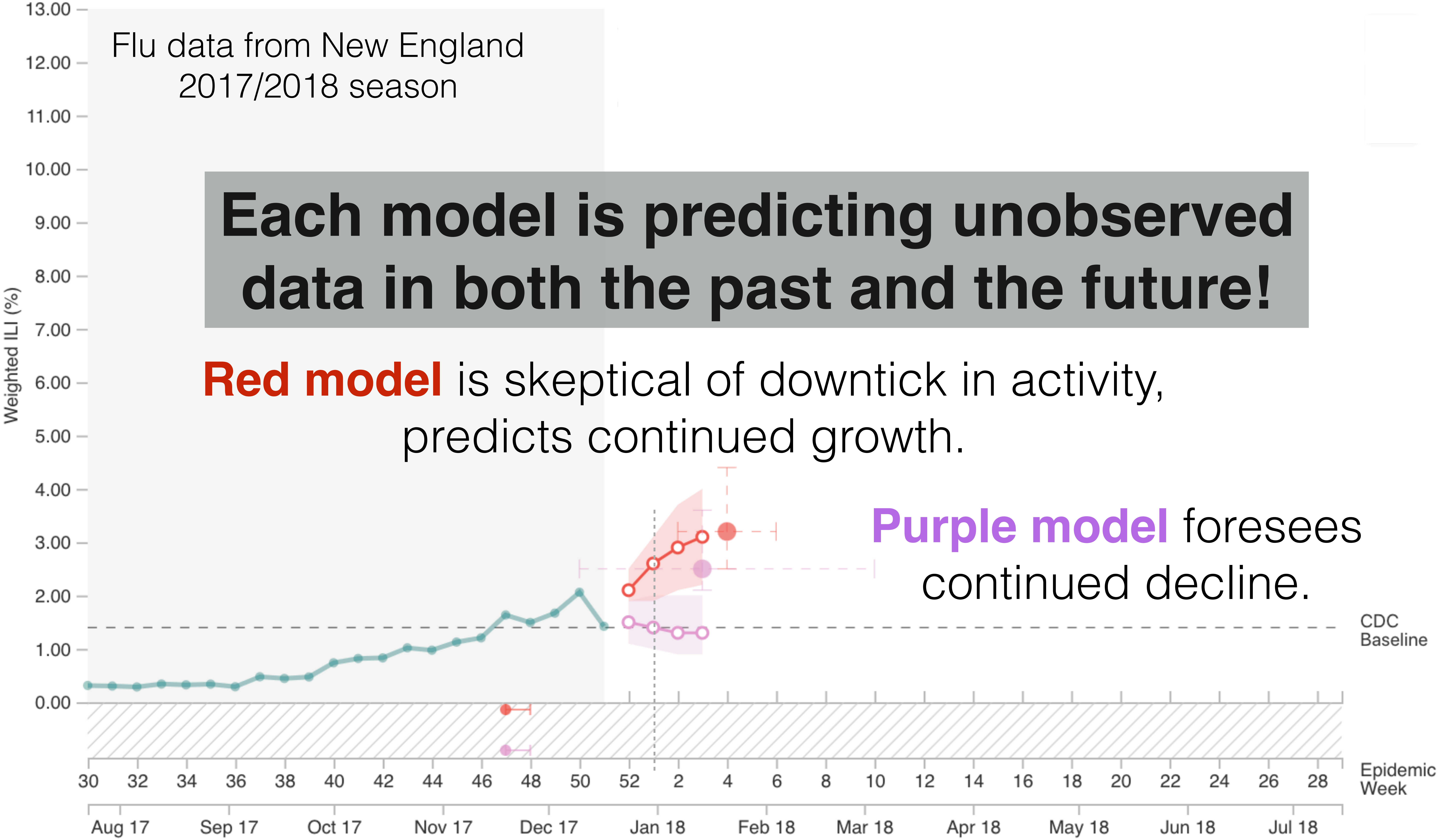


Purple model foresees continued decline.

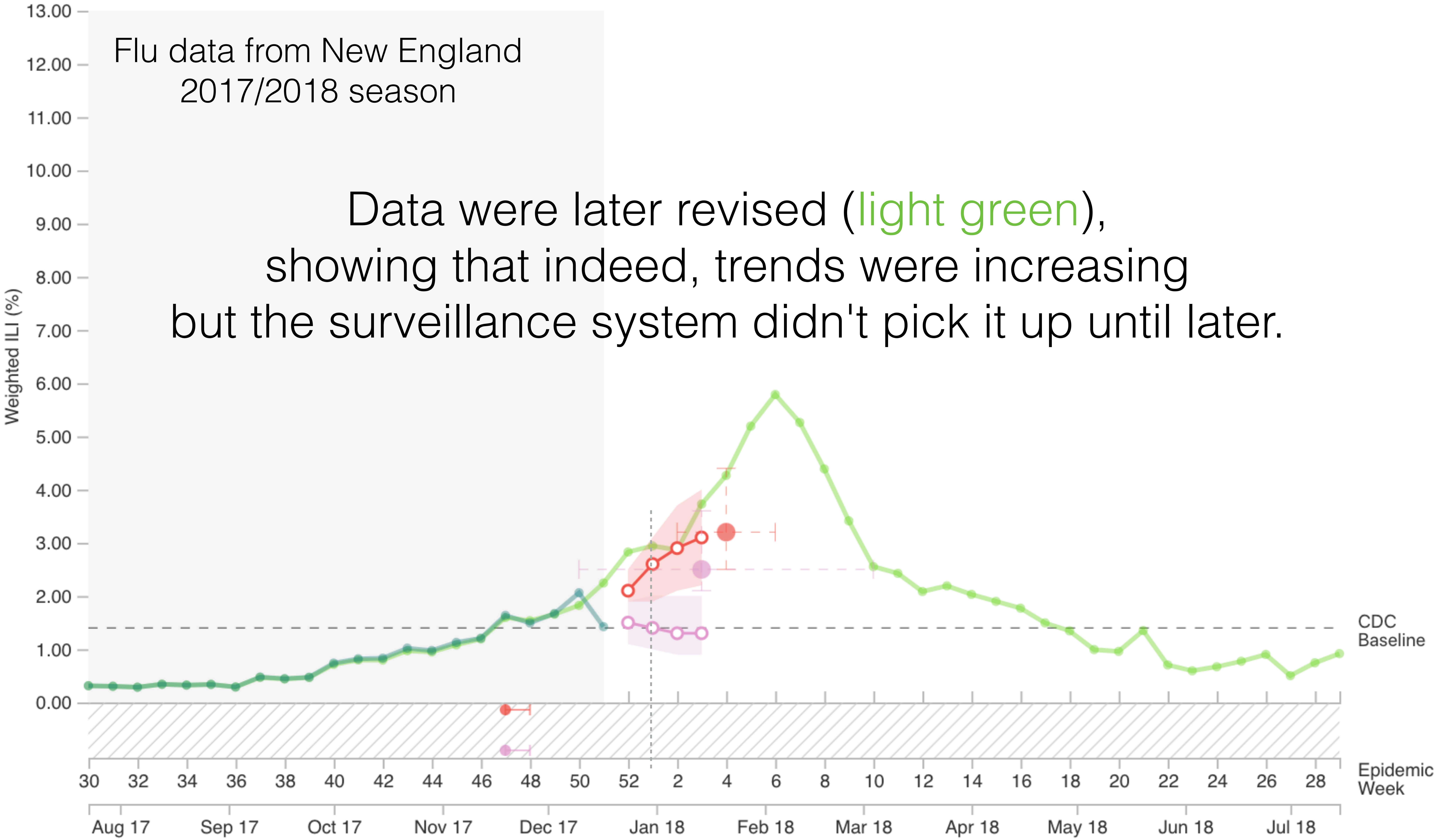
Modeling real-time surveillance data



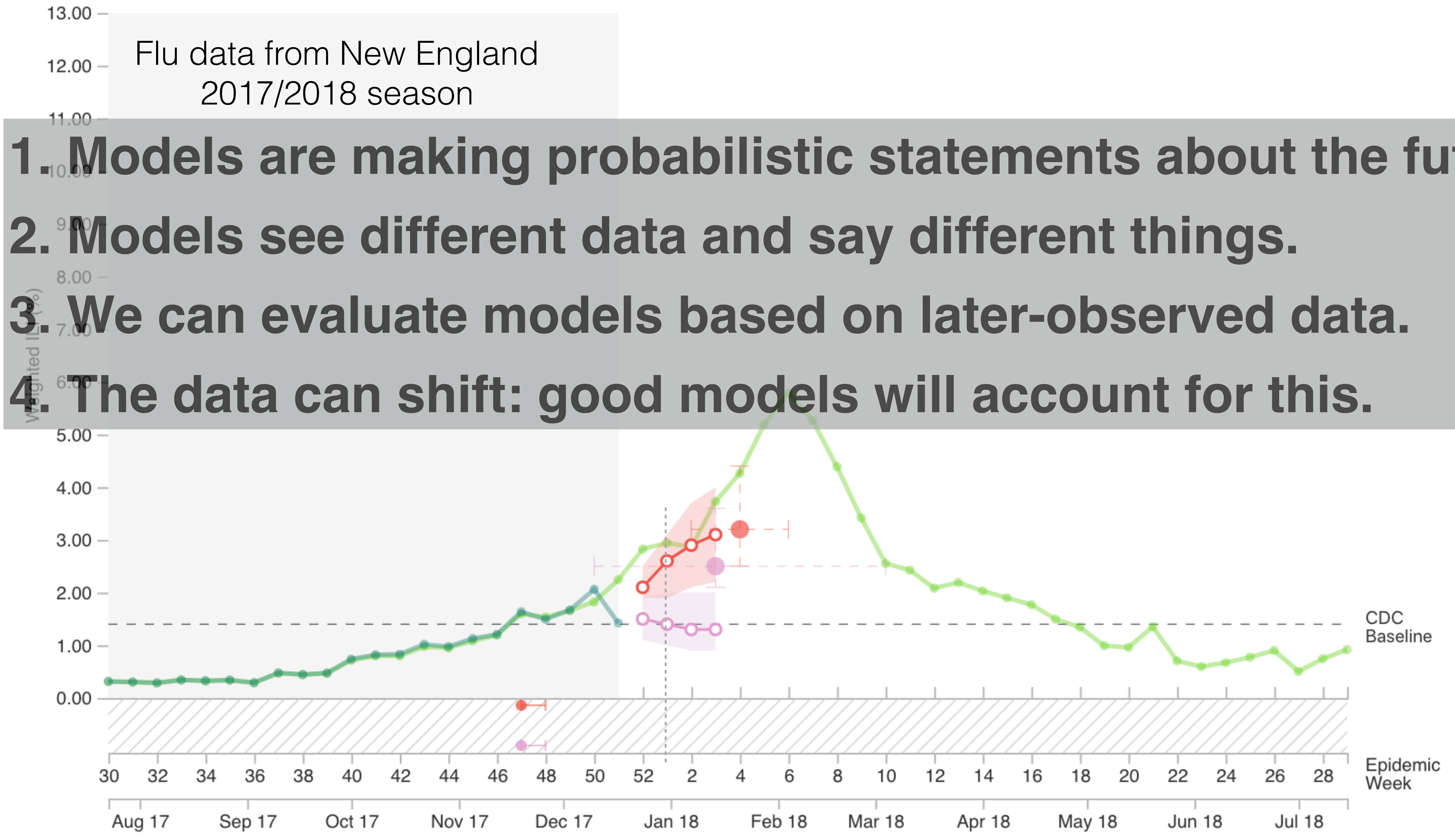
Modeling real-time surveillance data



Modeling real-time surveillance data



Modeling real-time surveillance data



1. **Models are making probabilistic statements about the future.**
2. **Models see different data and say different things.**
3. **We can evaluate models based on later-observed data.**
4. **The data can shift: good models will account for this.**

Chapter 1: An abstraction

Statistical Science
2001, Vol. 16, No. 3, 199–231

Statistical Modeling: The Two Cultures

Leo Breiman

Data arise thanks to the black box of nature.

outcome variable

predictor variables

e.g. covid cases next week in Hampshire county

covid cases this week, behavior, vaccination, cases in all past weeks, ...

Statistical Science
2001, Vol. 16, No. 3, 199–231

Statistical Modeling: The Two Cultures

Leo Breiman

"To extract some information about how nature is associating the response variables to the input variables."

One goal: **infer** something about nature from data.

We want to learn something about the "true" state of nature, but we will never be able to observe what the black box relationships are between all the $\mathbf{x}$ and y.

How do population structure, human behavior, biological features of a pathogen, etc... interact to cause an outbreak?

Statistical Science
2001, Vol. 16, No. 3, 199–231

Statistical Modeling: The Two Cultures

Leo Breiman

"To be able to predict what the responses are going to be to future input variables."

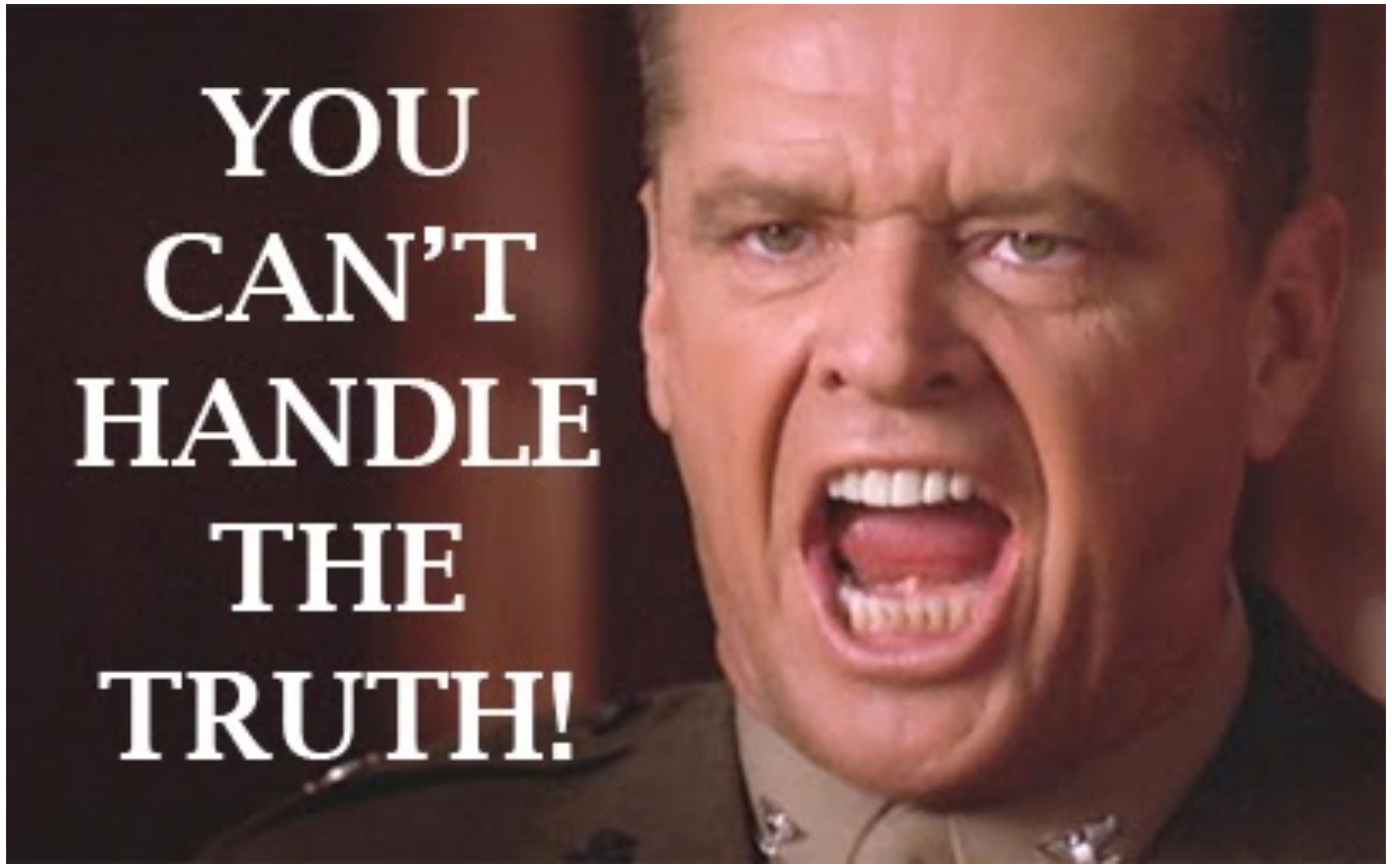
Another goal: **predict** new data.

In prediction, we might be less concerned learning about nature, and more with what the the outcome y will be. If we are careful, we can pick problems and settings where we can (eventually) know the truth about what y will be given some $\mathbf{x}$.

How many cases will be observed next week?

Philosophical differences btw. prediction and inference

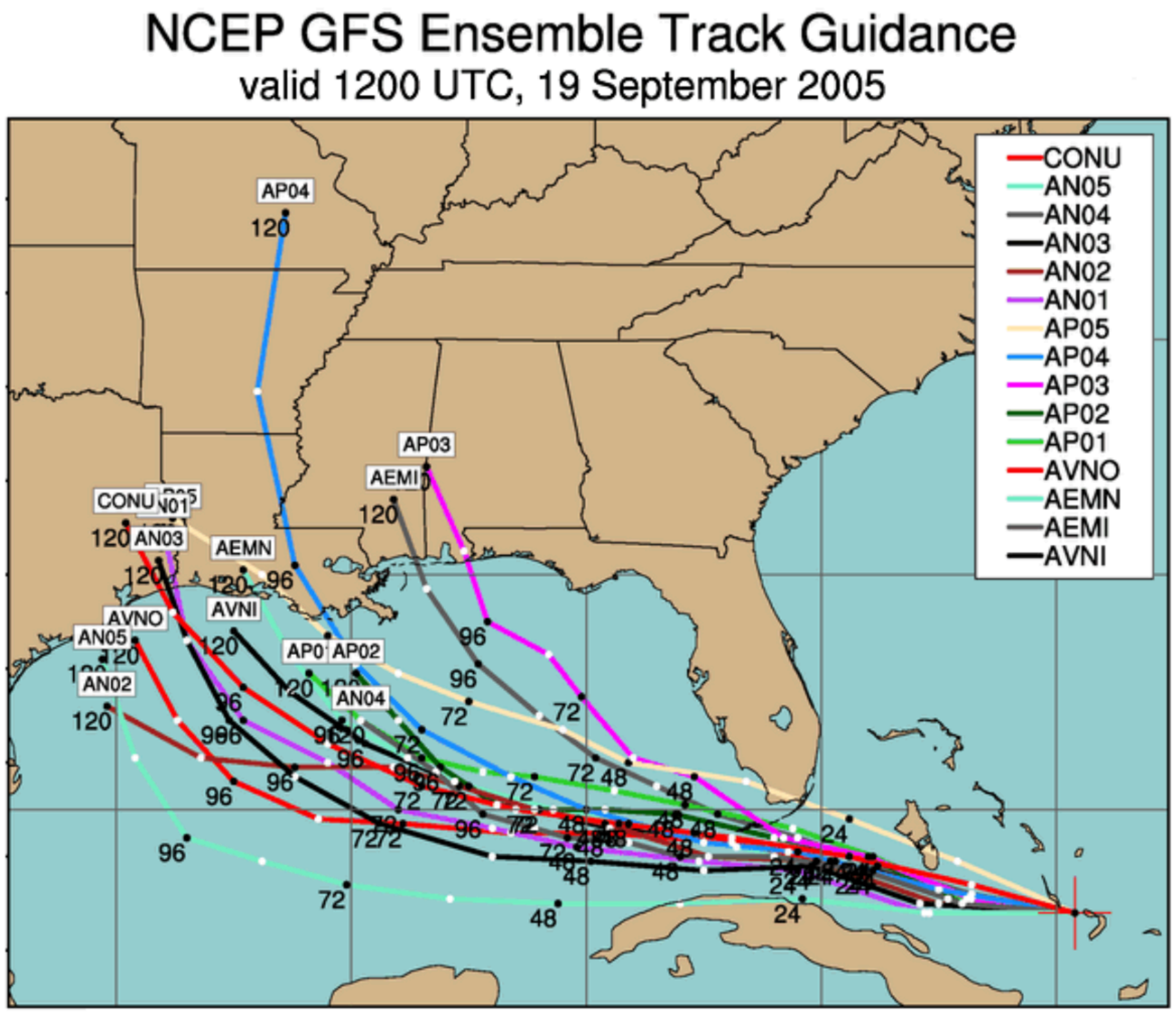
- With predictive modeling, in many cases, we will eventually see the "truth" and we will have a direct measure of the quality of our model.
- With inferential modeling, the "truth" remains forever elusive.
- This talk focuses on scientific challenges where we eventually get to see the "truth" (or a reasonable approximation of it).



The best approximations of truth involve multiple perspectives

- "Ensemble" forecasts, or combinations of predictions from different models, are a gold-standard in many fields focused on predictive modeling: climate, weather, economics, sports, etc...

- There are formal mathematical reasons for why this approach works, but the concepts behind the "wisdom of the crowd" approach are intuitive.



The GFS Ensemble track forecast for Hurricane Rita, initialized at 12Z on September 19, 2005, showed good agreement among the ensemble members in the short term, but less agreement after 24 hours.

Credit: NCAR Tropical Cyclone Guidance Project

https://learningweather.psu.edu/node/60

Chapter 2:
background on epidemic forecasting

What do epidemic models predict?

An epidemiological modeler's view on predicting the past, the near future, and the far future.

Image credit: Nicole Samay, Alex Vespignani,
via the Scenario Modeling Hub, https://covid19scenariomodelinghub.org/

What do epidemic models predict?

An epidemiological modeler's view on predicting the past, the near future, and the far future.

Image credit: Nicole Samay, Alex Vespignani,
via the Scenario Modeling Hub, https://covid19scenariomodelinghub.org/

What do epidemic models predict?

An epidemiological modeler's view on predicting the past, the near future, and the far future.

Image credit: Nicole Samay, Alex Vespignani,
via the Scenario Modeling Hub, https://covid19scenariomodelinghub.org/

What do epidemic models predict?

An epidemiological modeler's view on predicting the past, the near future, and the far future.

Image credit: Nicole Samay, Alex Vespignani,
via the Scenario Modeling Hub, https://covid19scenariomodelinghub.org/

What do epidemic models predict?

An epidemiological modeler's view on predicting the past, the near future, and the far future.

Image credit: Nicole Samay, Alex Vespignani,
via the Scenario Modeling Hub, https://covid19scenariomodelinghub.org/

What do epidemic models predict?

An epidemiological modeler's view on predicting the past, the near future, and the far future.

Scenario projections can provide decision-makers and the general public with a set of hypothetical futures based on comparisons of different policy choices or other contingencies.

Image credit: Nicole Samay, Alex Vespignani, via the Scenario Modeling Hub, https://covid19scenariomodelinghub.org/

What do epidemic models predict?

An epidemiological modeler's view on predicting the past, the near future, and the far future.

Image credit: Nicole Samay, Alex Vespignani,
via the Scenario Modeling Hub, https://covid19scenariomodelinghub.org/

What do epidemic models predict?

An epidemiological modeler's view on predicting the past, the near future, and the far future.

Image credit: Nicole Samay, Alex Vespignani,
via the Scenario Modeling Hub, https://covid19scenariomodelinghub.org/

Challenge: data sparsity

(infectious disease dynamics cannot be observed like the weather)

image credit: https://goo.gl/images/CSSQRv

Each dot represents a weather station whose data was used to create the WorldClim dataset.

image credit: https://databasin.org/datasets/15a31dec689b4c958ee491ff30fcce75

Challenge: epidemic forecast feedback loop

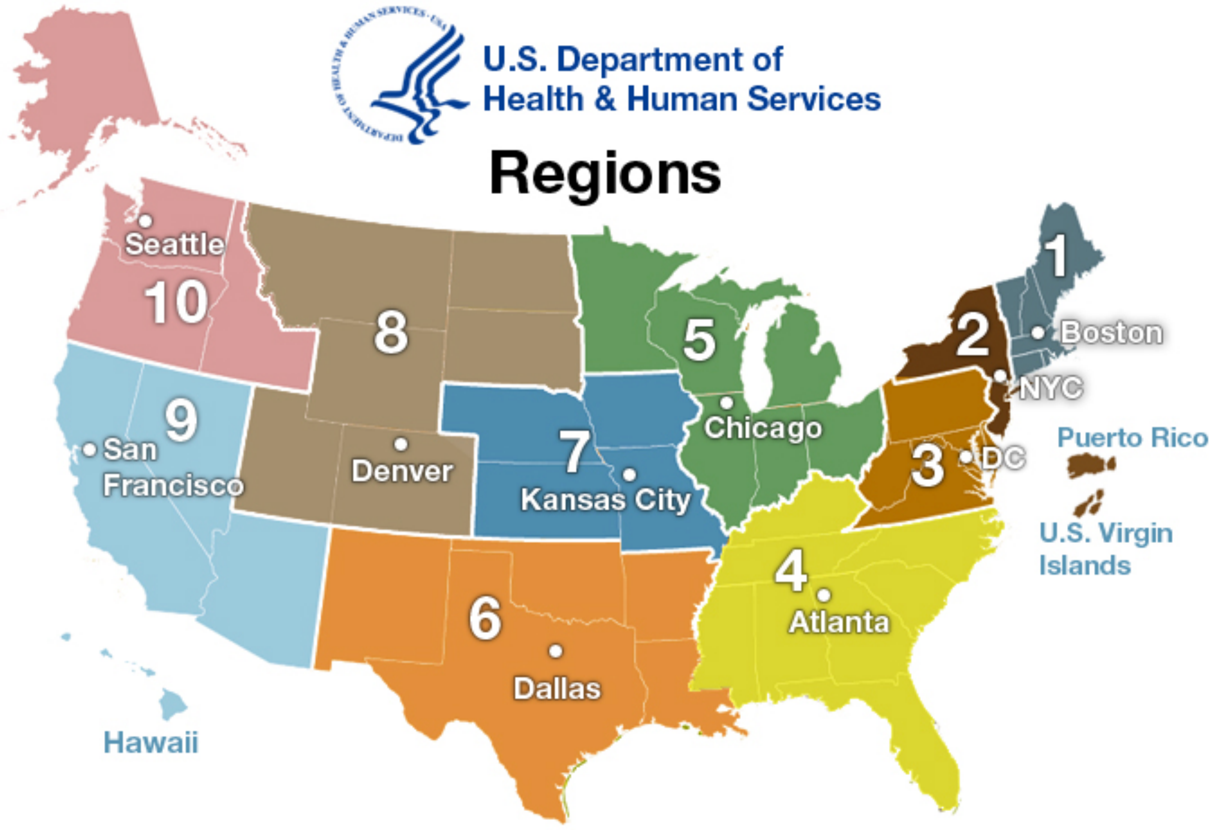
- Weather forecasts don't impact the weather.
- An outbreak forecast or projection could impact an outbreak.

2014: US military troops heading to Liberia to assist with Ebola outbreak.
image: defense.gov

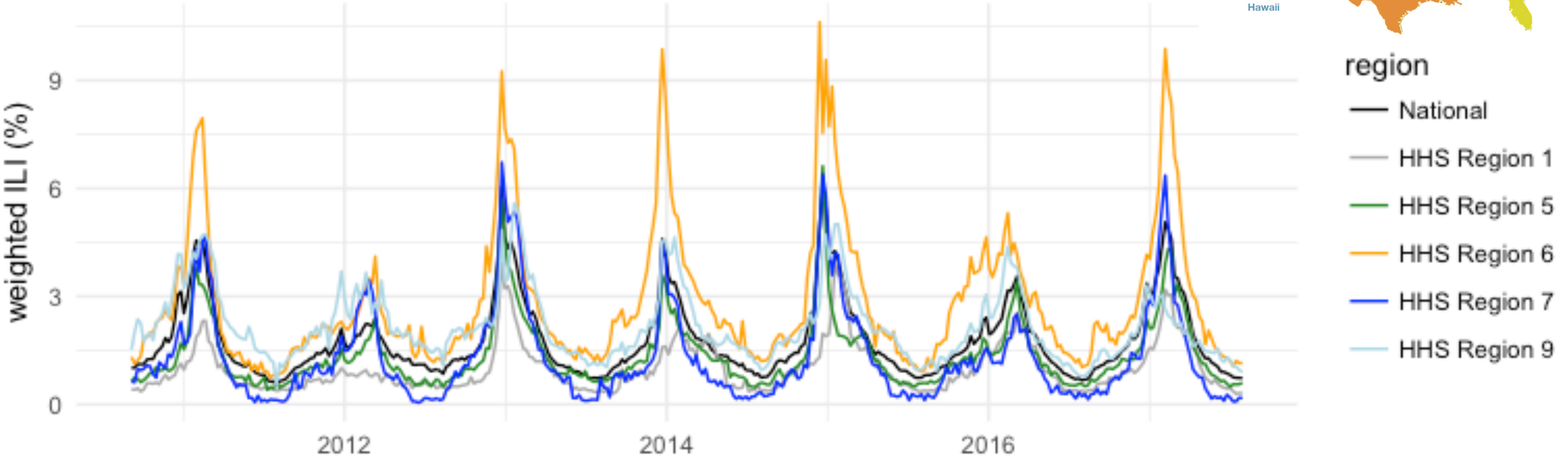
2018: vector-control activities to prevent dengue in Thailand
courtesy of Sopon Iamsirithaworn, Thailand Department of Disease Control

Typical epidemic forecasting setup

e.g. CDC FluSight challenges: U.S. national, regional, state level. Running annually since 2013.



Target variable "weighted ILI":
The % of all outpatient visits with primary complaint of influenza-like illness (ILI), weighted by state population.



Targets with public health relevance

from annual CDC FluSight forecasting challenge

Biggerstaff et al. 2016, *BMC Inf Dis*. https://doi.org/10.1186/s12879-016-1669-x
McGowan et al. 2019, *Sci Rep*. https://doi.org/10.1038/s41598-018-36361-9
Lutz et al. 2019. *BMC Pub Hlth*. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6902553/

The Washington Post

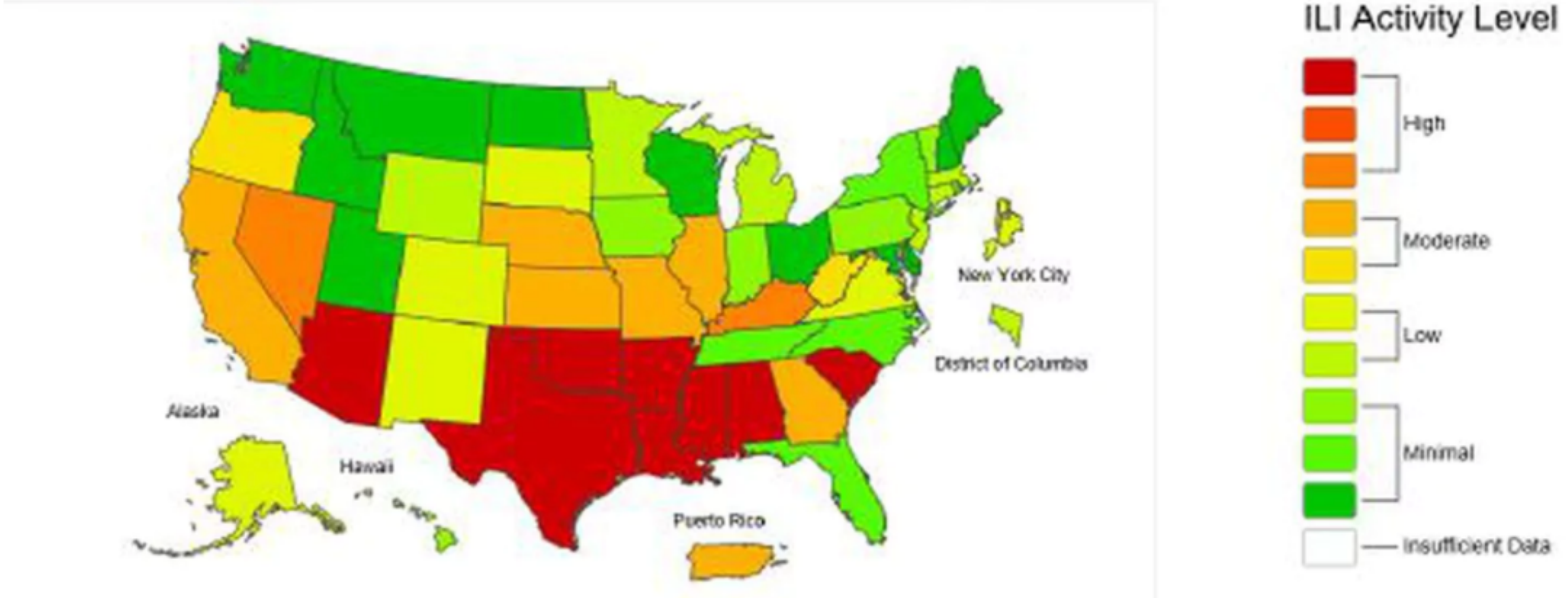
To Your Health

Why this may be a bad flu season, especially around the holidays

By Lena H. Sun December 22, 2017

CDC's flu forecasters say there's a 30 percent chance the season will peak around the end of December and a 60 percent chance that the greatest incidence will be by late January, Jernigan said. Generally, flu season peaks near the end of February.

Influenza-Like Illness (ILI) Activity Level Indicator Determined by Data Reported to ILINet
2017-18 Influenza Season Week 50 ending Dec 16, 2017



Dan Jernigan, Director of Influenza Division, CDC
September 2018

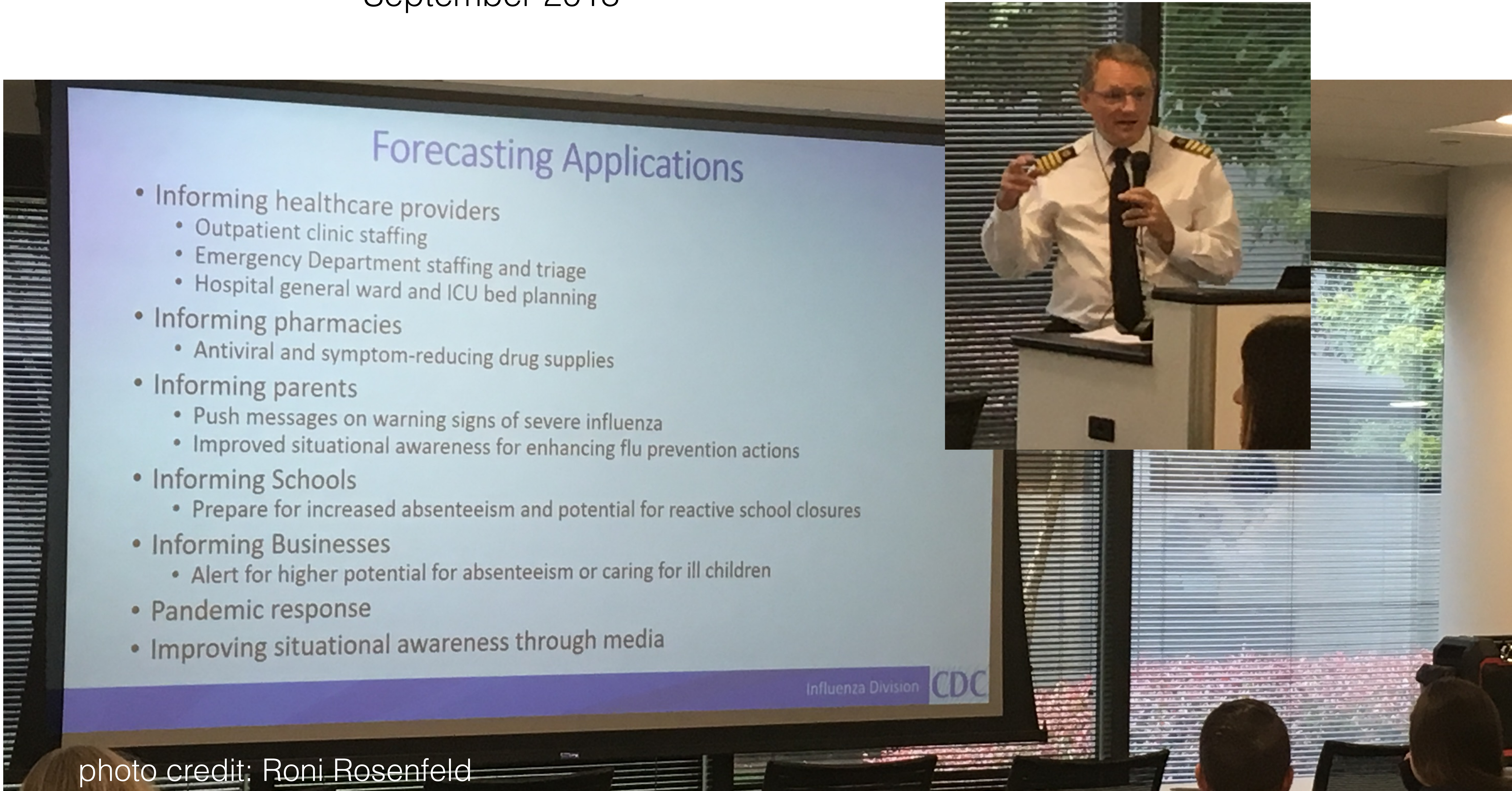


photo credit: Roni Rosenfeld

Model coordination is key

- There have been numerous government-coordinated outbreak forecasting efforts (flu, Ebola, chikungunya, Zika, dengue, etc...).
- A combination of individual forecasts is pragmatic: it reduces dependency on a single model or team.
- 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*.
COVID-19: Cramer et al. 2020, *medrxiv*.

The "Hub" idea is not new

- The idea: coordinated modeling between groups to inform policy and/or develop knowledge about a system.

- Different than a competition: involving coordination between groups and input from policy-makers, often in real-time.

Climate

ipcc

ipcc.ch

Ecology

FISHERIES & MARINE ECOSYSTEM
FISH-MIP
MODEL INTERCOMPARISON PROJECT

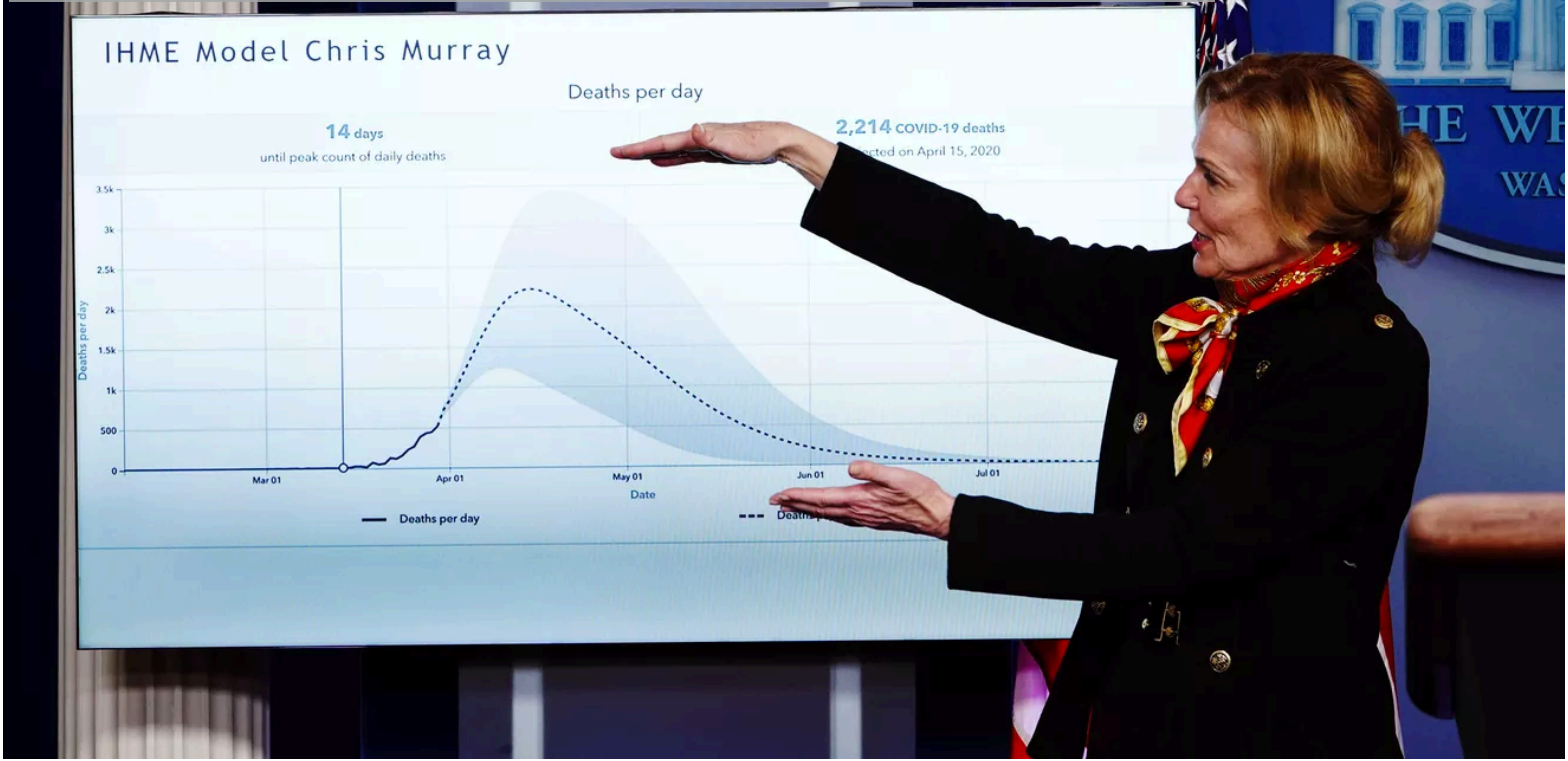
isimip.org/about/marine-ecosystems-fisheries/

Space Science

Community Coordinated Modeling Center

ccmc.gsfc.nasa.gov

Policy makers needed more than1 model



early April 2020

Chapter 3:
The US COVID-19 Forecast Hub

COVID-19 ForecastHub

Goals

covid19forecasthub.org

Launched April 6, 2020

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

2. Assess reliability of forecasts and gain insight into which modeling approaches do well.

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

Read more: https://covid19forecasthub.org/doc/research/

Clones

COVID-19 ForecastHub

European COVID-19 ForecastHub

https://covid19forecasthub.eu/

COVID-19 ScenarioModelingHub

https://covid19scenariomodelinghub.org/

COVID-19 ForecastHub

Numbers

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

- The Hub builds **an ensemble that combines quantile-based predictive distributions** from these models for 1 through 4 week ahead forecasts.

- To date, we have curated data from **10~~5~~6 models: over 4,900 submissions** and **7~~1~~2 million unique predictions**.

Data on GitHub and on Zoltar.
https://github.com/reichlab/covid19-forecast-hub/ https://zoltardata.com/project/44

COVID-19
ForecastHub

Data from the COVID-19 Forecast Hub are 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

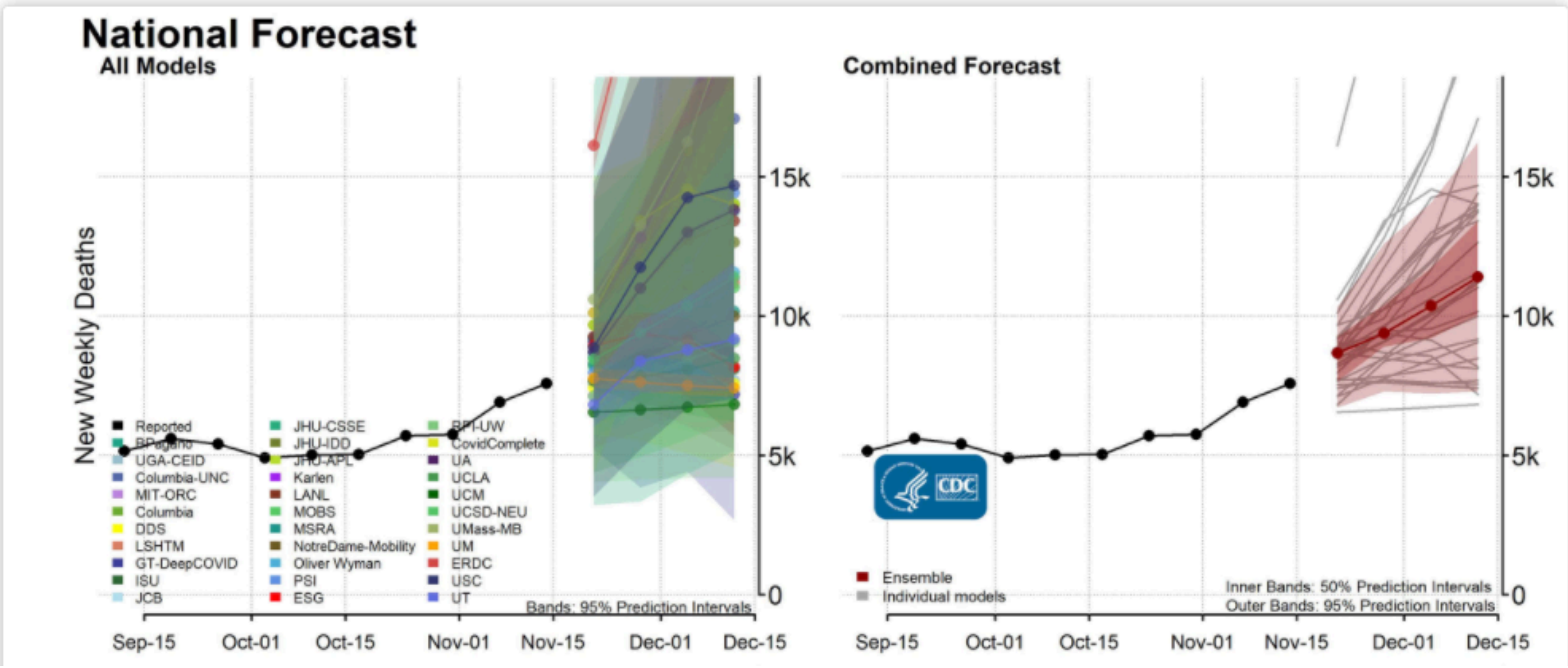
Observed and forecasted new and total reported COVID-19 deaths as of November 16, 2020.

Interpretation of Forecasts of New and Total Deaths

On This Page

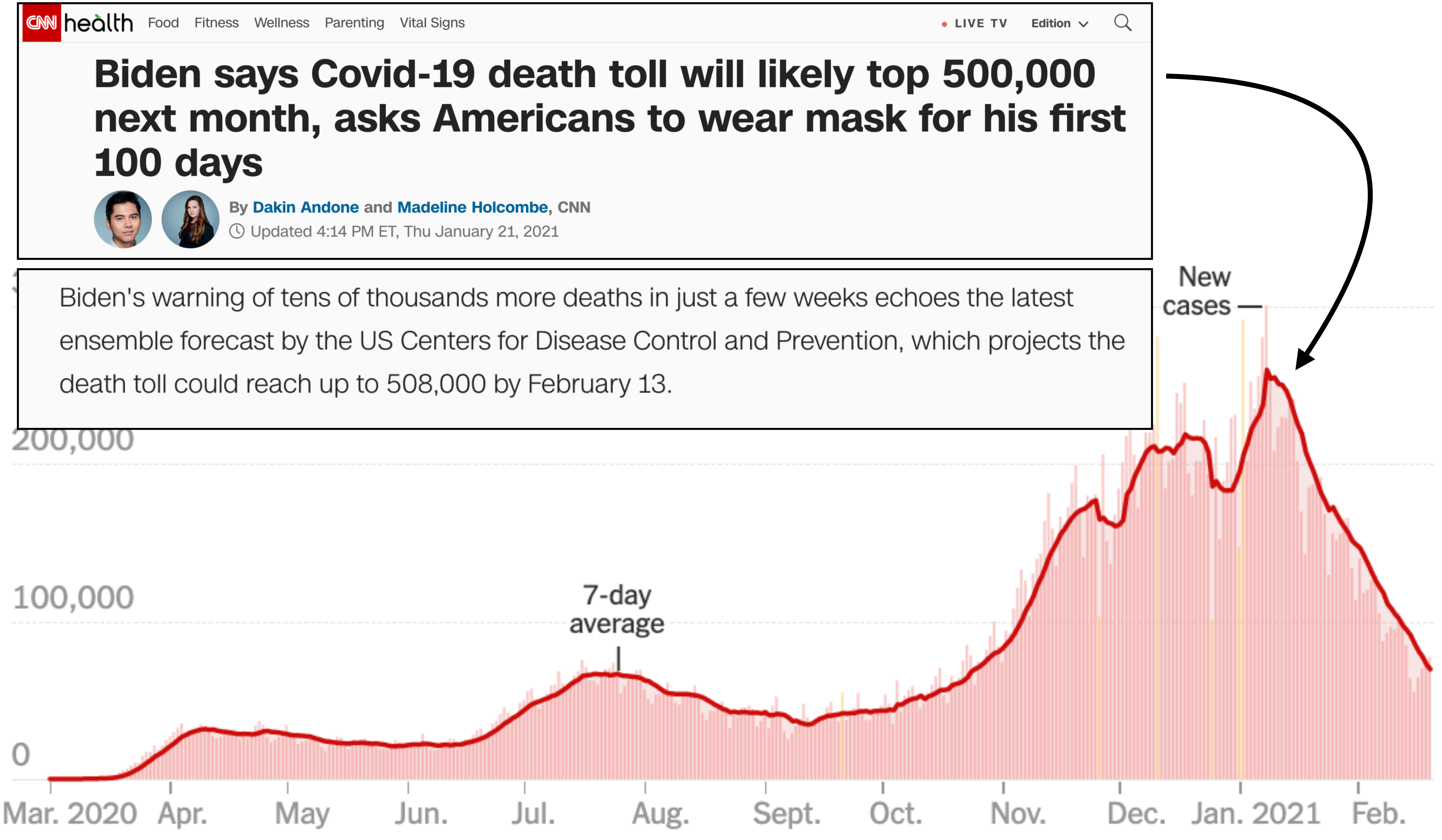
- This week CDC received forecasts of COVID-19 deaths over the next 4 weeks from 36 modeling groups that were included in the ensemble forecast. Of the 36 groups, 33 provided forecasts for both new and total deaths, two groups forecasted total deaths only, and one forecasted new death only.
- This week's 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.

National Forecast



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

Ensemble in use by government officials


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

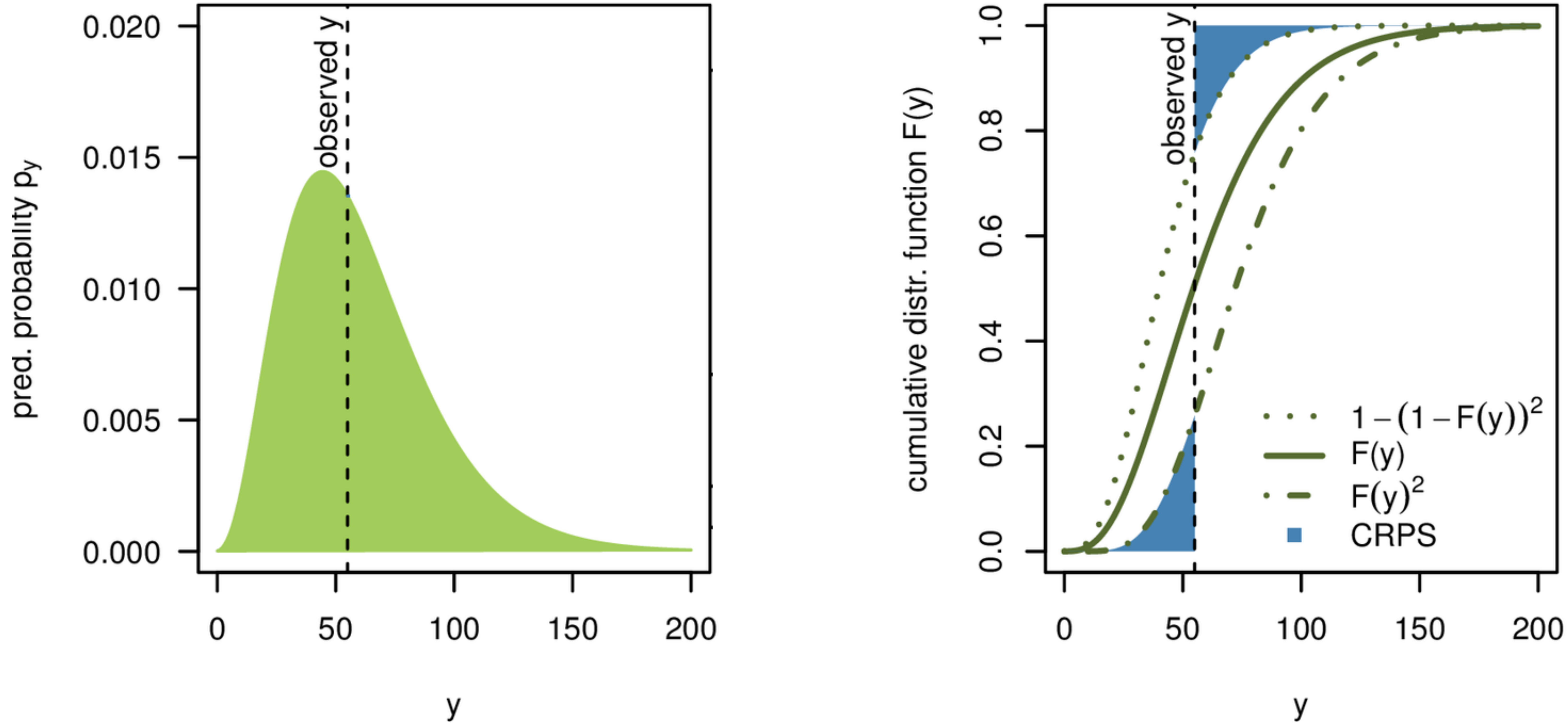
Updated 4:14 PM ET, Thu January 21, 2021

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

Evaluating forecast accuracy (1)

1. Weighted Interval Score (WIS): a score rewarding both accuracy and precision.

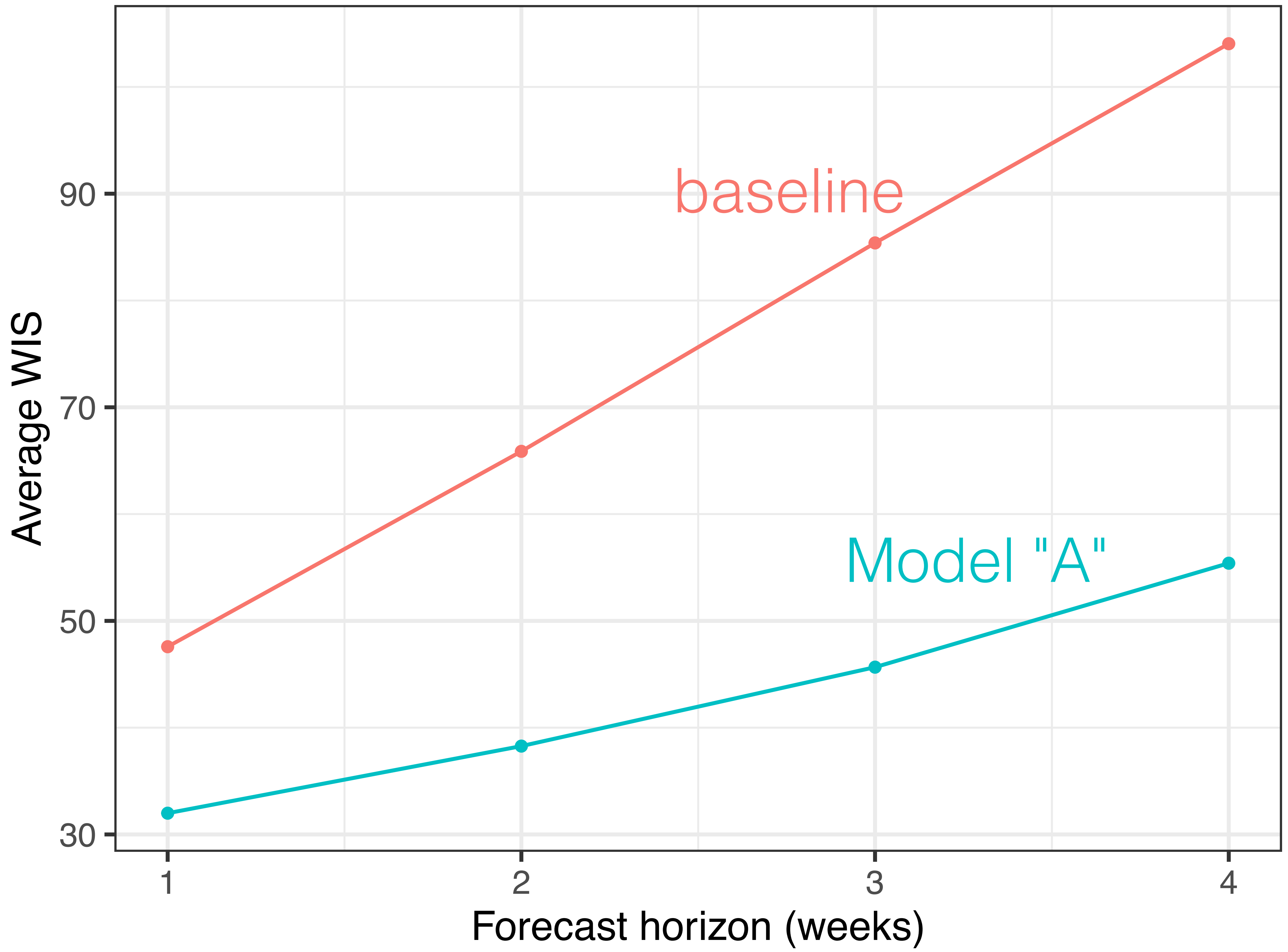


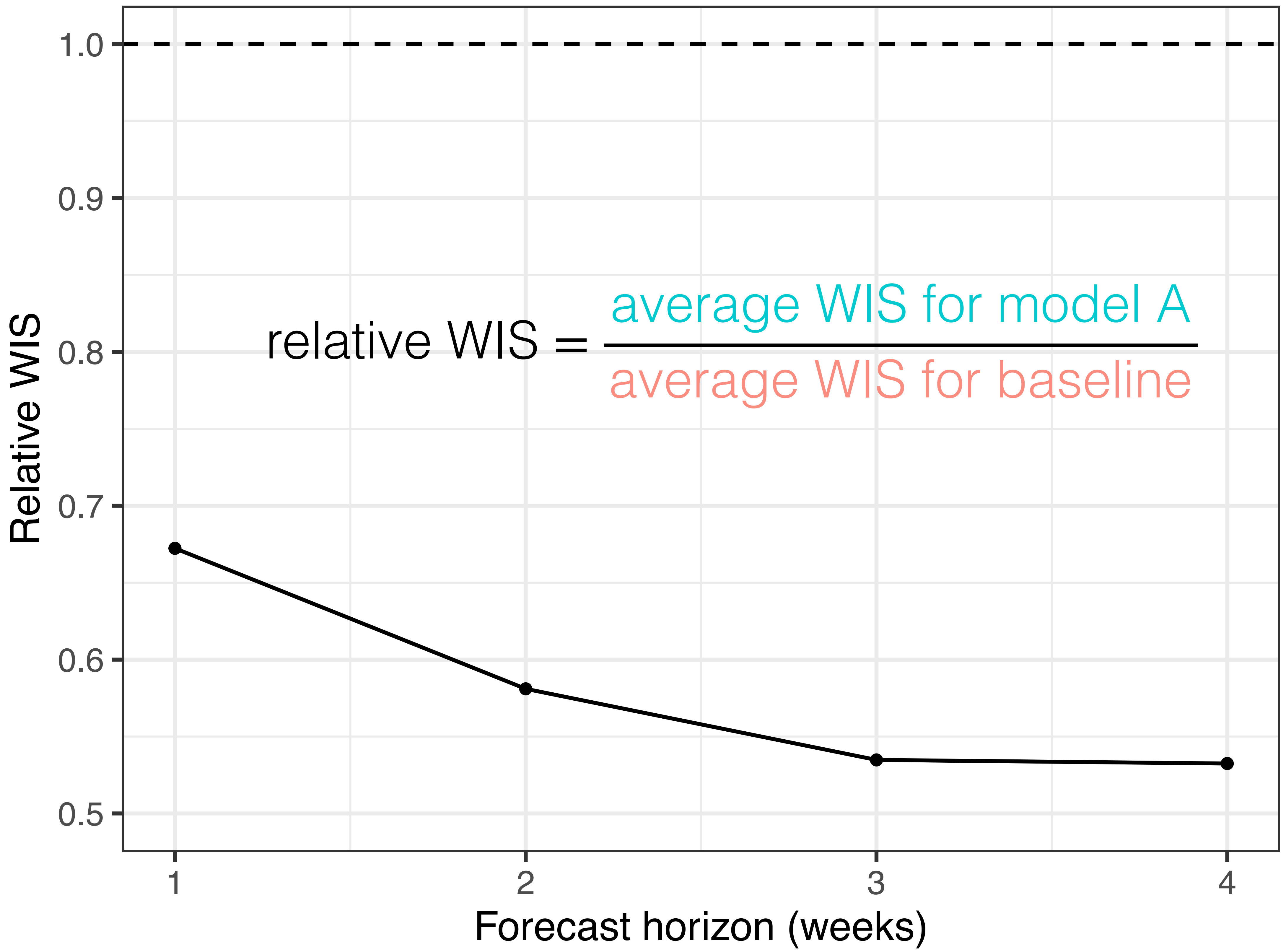
Based on the submission format of forecasts, we developed the WIS which is an approximation to the commonly used Continuous Rank Probability Score (CRPS), and equivalent to the commonly used "pinball loss" in machine learning. **Lower scores are better.**

slide adapted from Bracher et al (2021). *PLOS Comp Bio*. DOI:10.1371/journal.pcbi.1008618

Evaluating forecast accuracy (1a)

1a. Relative Weighted Interval Score (Relative WIS): how accurate is a model relative to a naïve baseline?





Relative WIS lower than 1 means the forecasts were better than the baseline.

Evaluating forecast accuracy (2)

2. Prediction interval coverage: when you say something has a 50% chance of occurring does it really happen with that frequency?

With millions (!!) of observations across different weeks and locations, we can develop a good sense of how "reliable" the models are.

Individual COVID-19 models vary

- IHME-CurveFit: "**hybrid modeling approach** to generate our forecasts, which incorporates elements of statistical and disease transmission models."

- 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**"

- CMU-TimeSeries: "A **basic AR-type time series model** fit using lagged values of case counts and deaths as features. No assumptions are made regarding reopening or governmental interventions."

Individual COVID-19 models vary

- IHME-CurveFit: "**hybrid modeling approach** to generate our forecasts, which incorporates elements of statistical and disease transmission models."

- 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**"

- CMU-TimeSeries: "A **basic AR-type time series model** fit using lagged values of case counts and deaths as features. No assumptions are made regarding reopening or governmental interventions."

Individual COVID-19 models vary

- IHME-CurveFit: "**hybrid modeling approach** to generate our forecasts, which incorporates elements of statistical and disease transmission models."

- 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**"

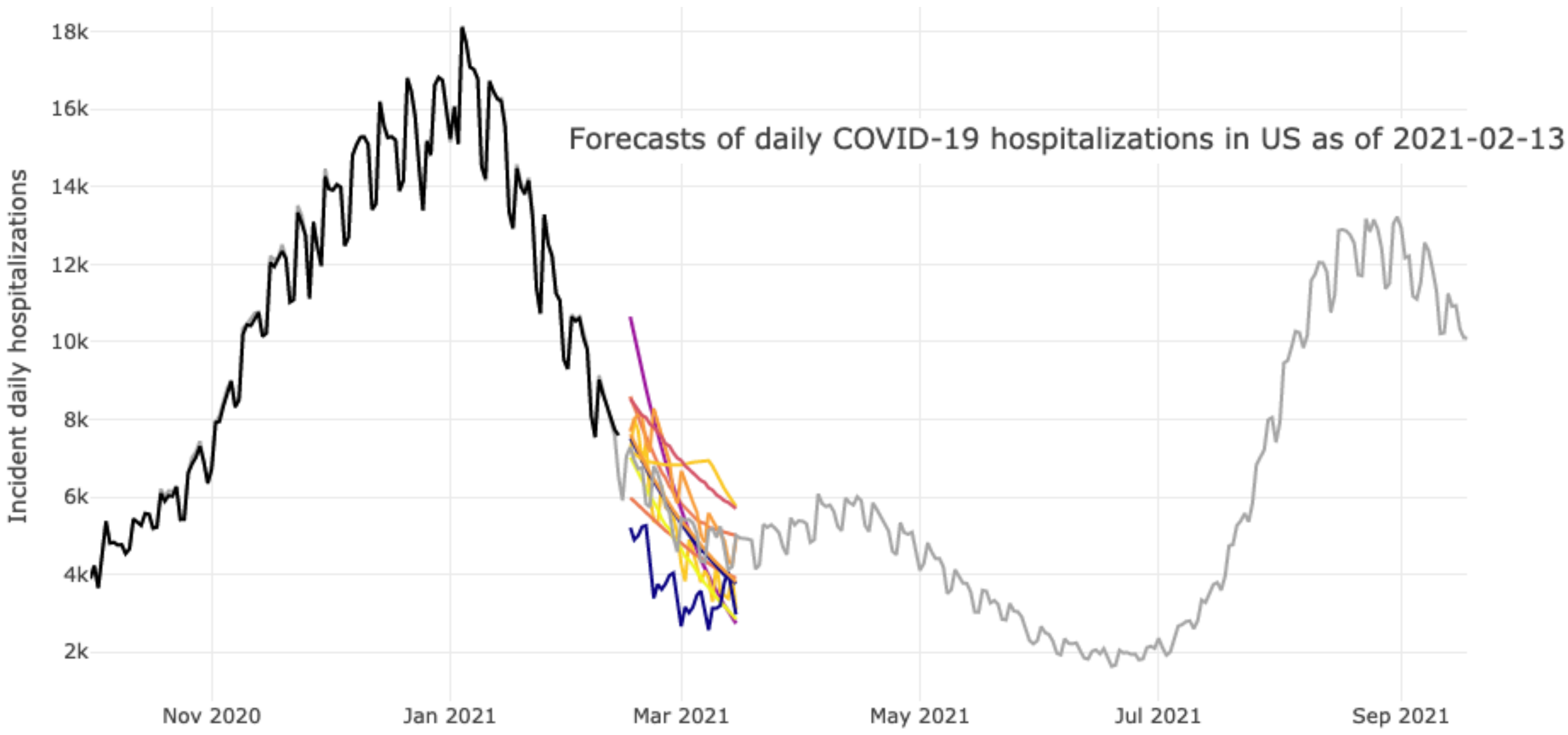
- CMU-TimeSeries: "A **basic AR-type time series model** fit using lagged values of case counts and deaths as features. No assumptions are made regarding reopening or governmental interventions."

Individual COVID-19 models vary

- IHME-CurveFit: "**hybrid modeling approach** to generate our forecasts, which incorporates elements of statistical and disease transmission models."
- 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**"
- CMU-TimeSeries: "A **basic AR-type time series model** fit using lagged values of case counts and deaths as features. No assumptions are made regarding reopening or governmental interventions."

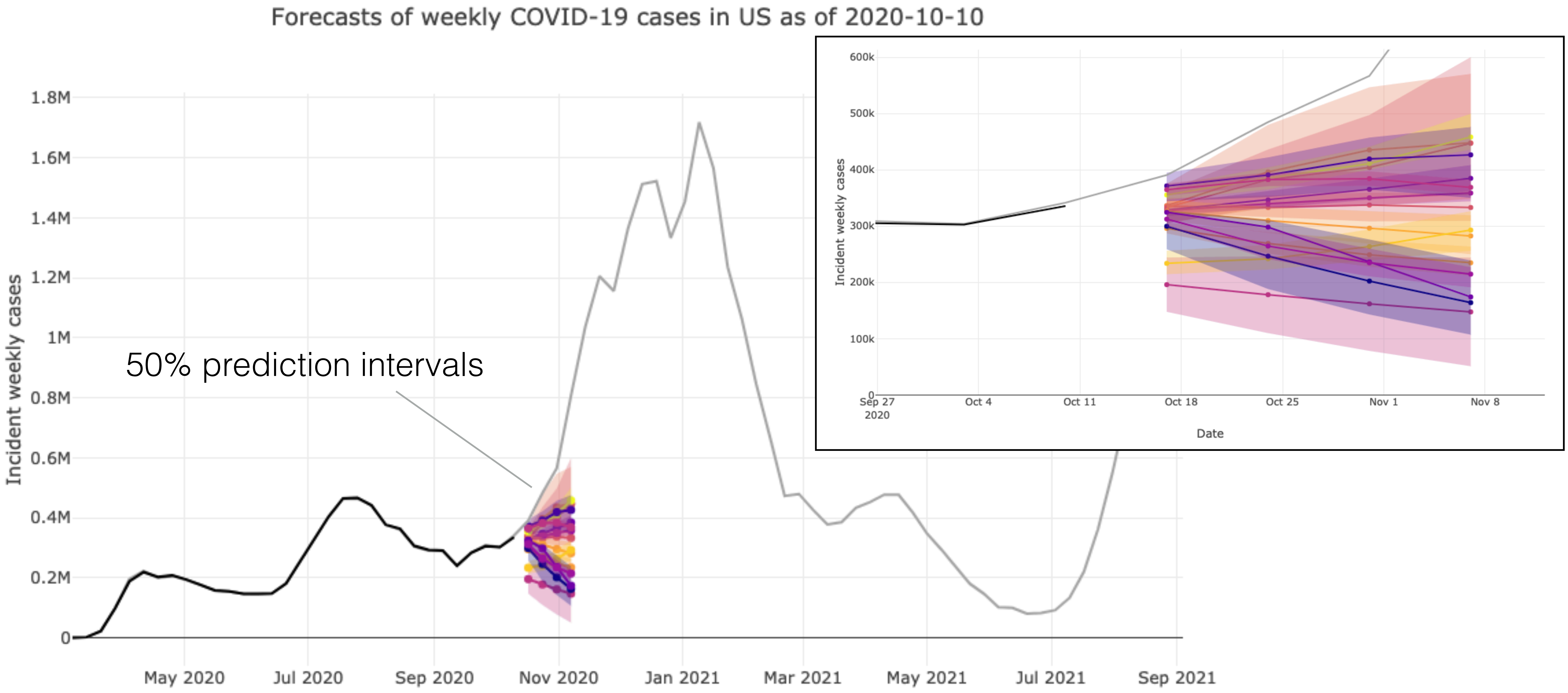
Forecasts often have made reasonable statements

especially during times where a stable trend continues



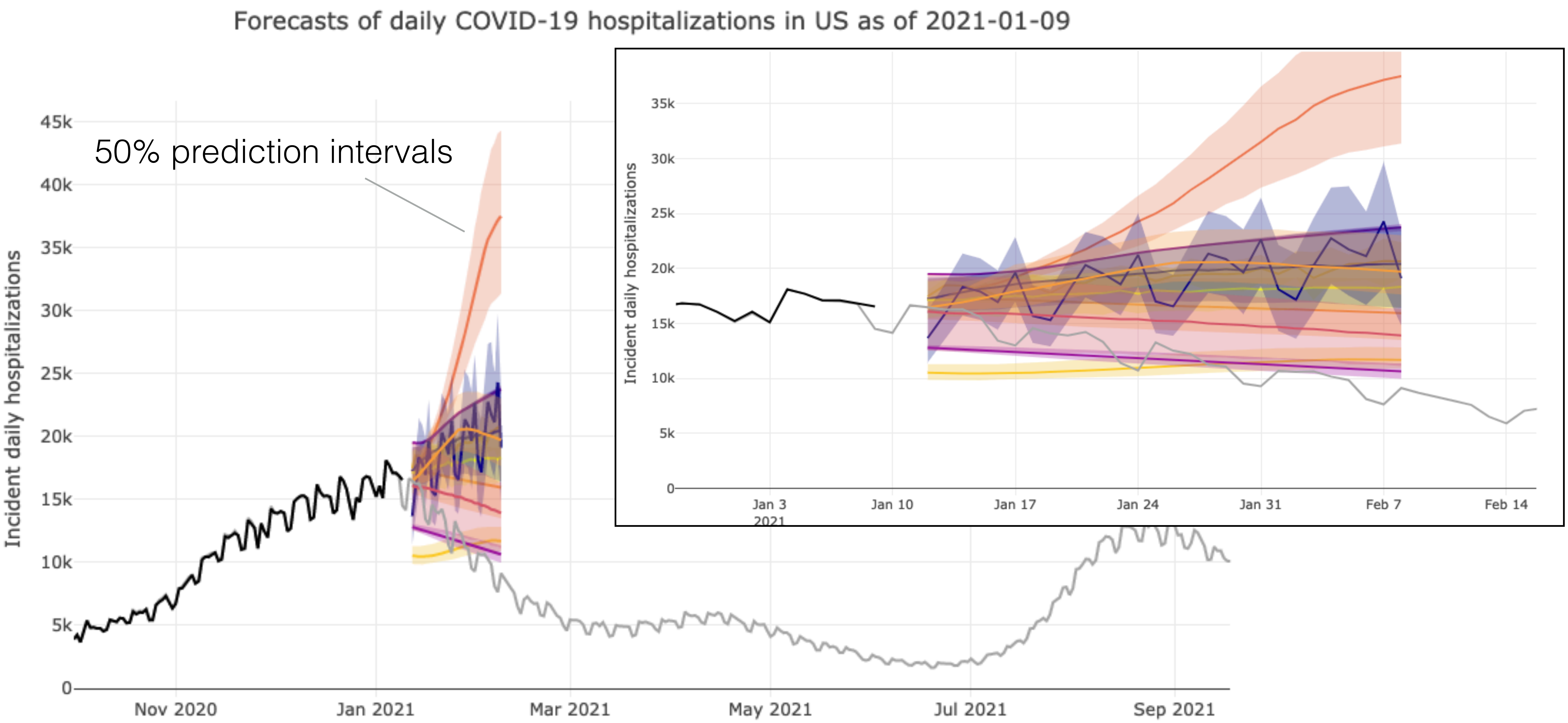
Forecasts have missed change-points

especially for case forecasts



Forecasts have missed change-points

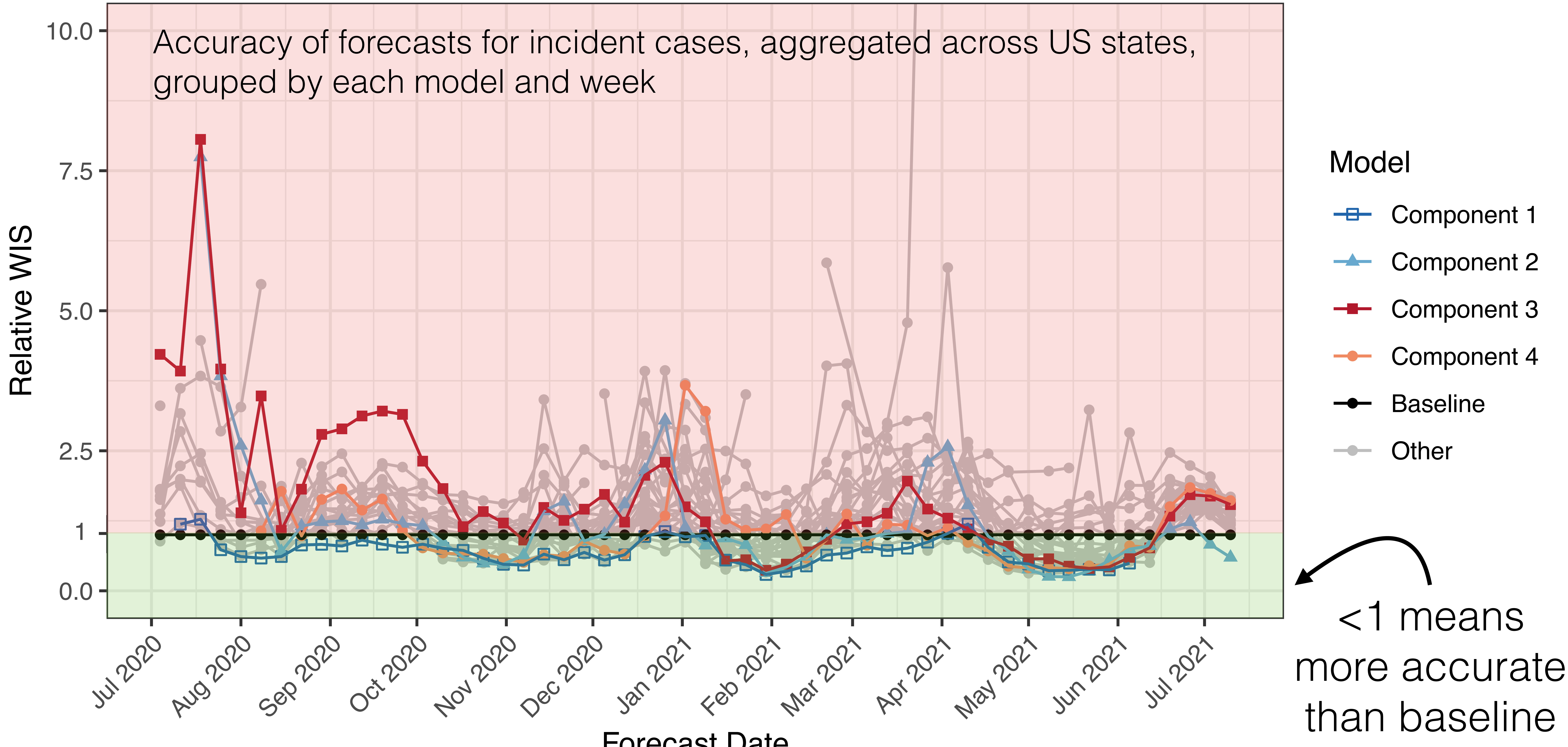
also for hospitalization forecasts



Forecasts have missed change-points

and for deaths as well

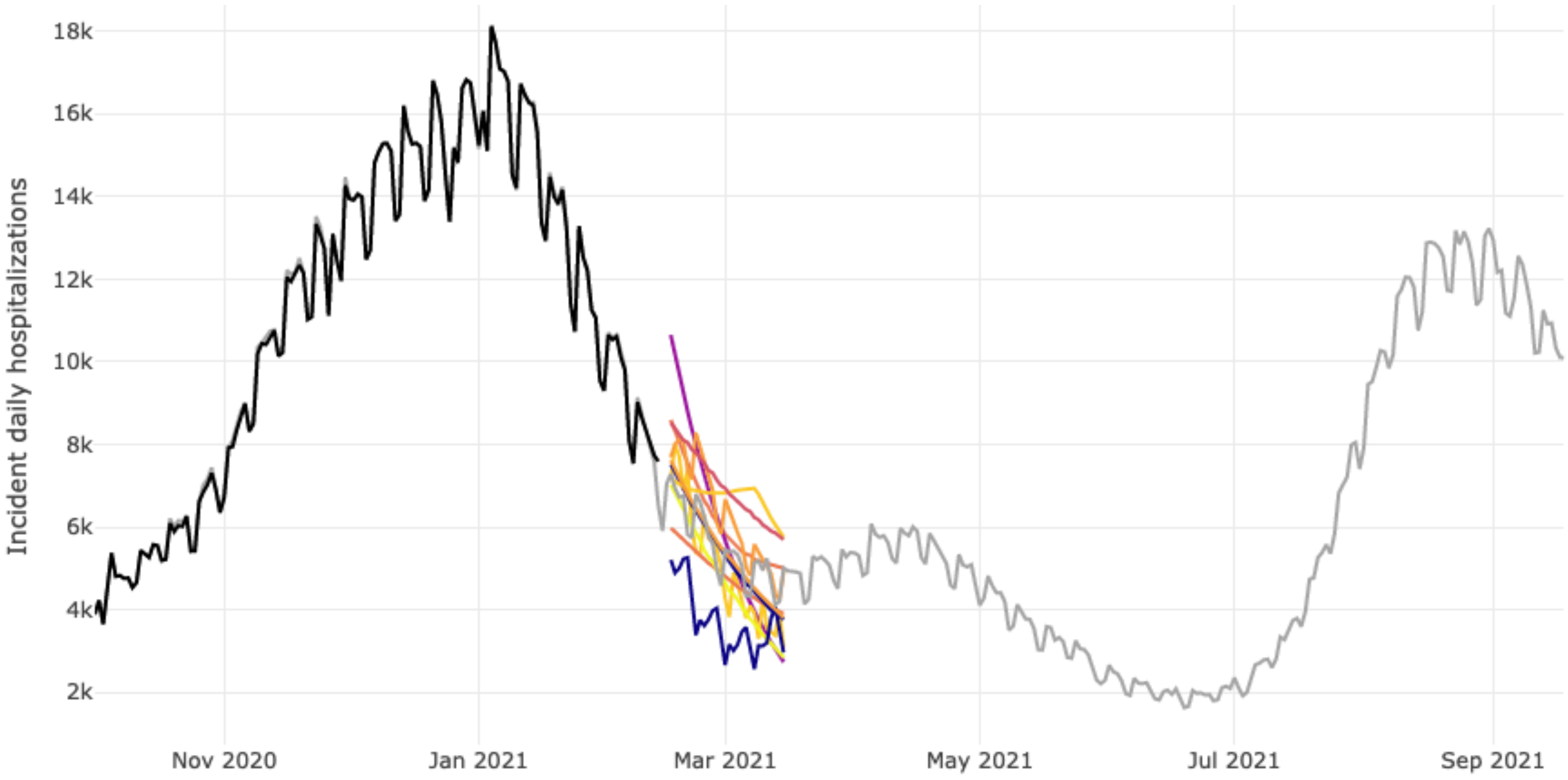
Model performance changes over time



<1 means more accurate than baseline

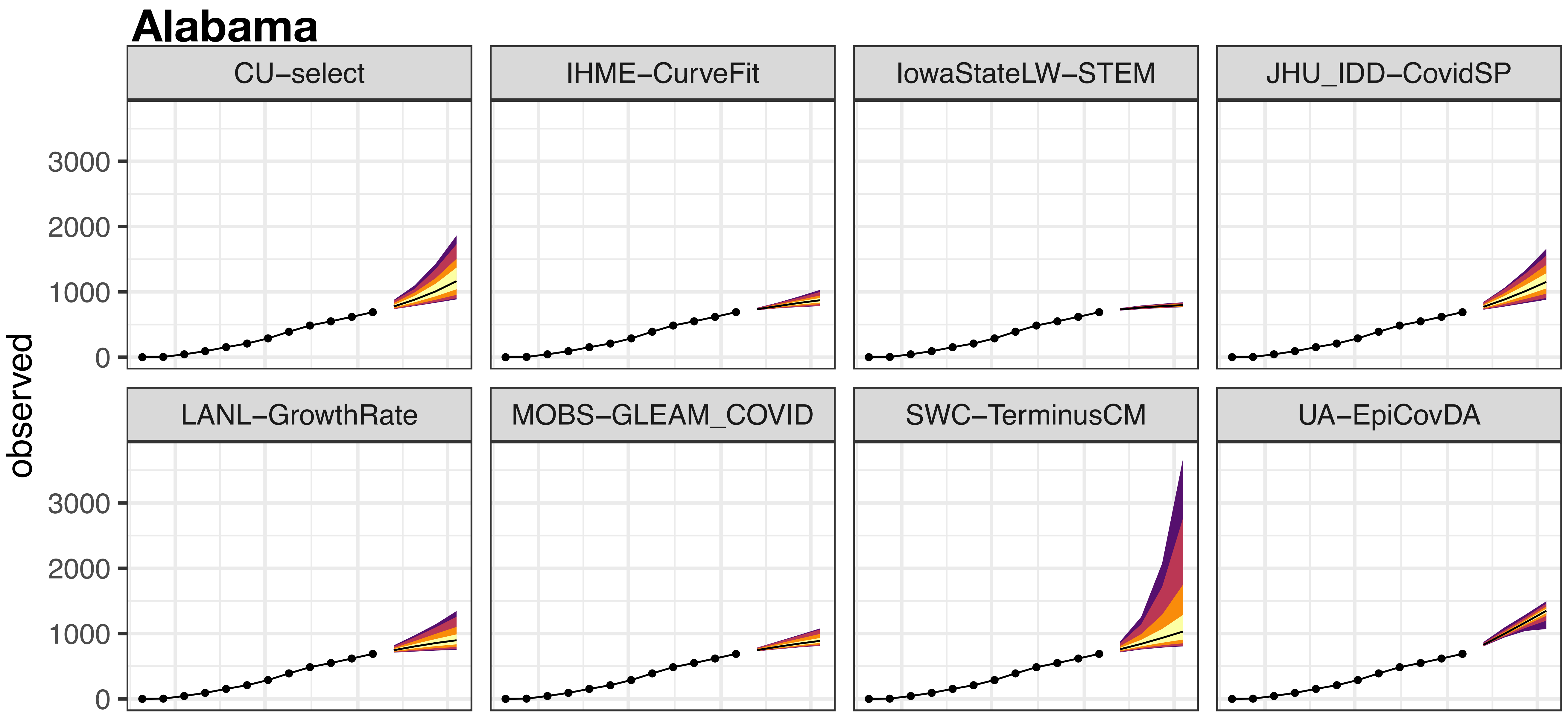
slide adapted from content by Evan Ray

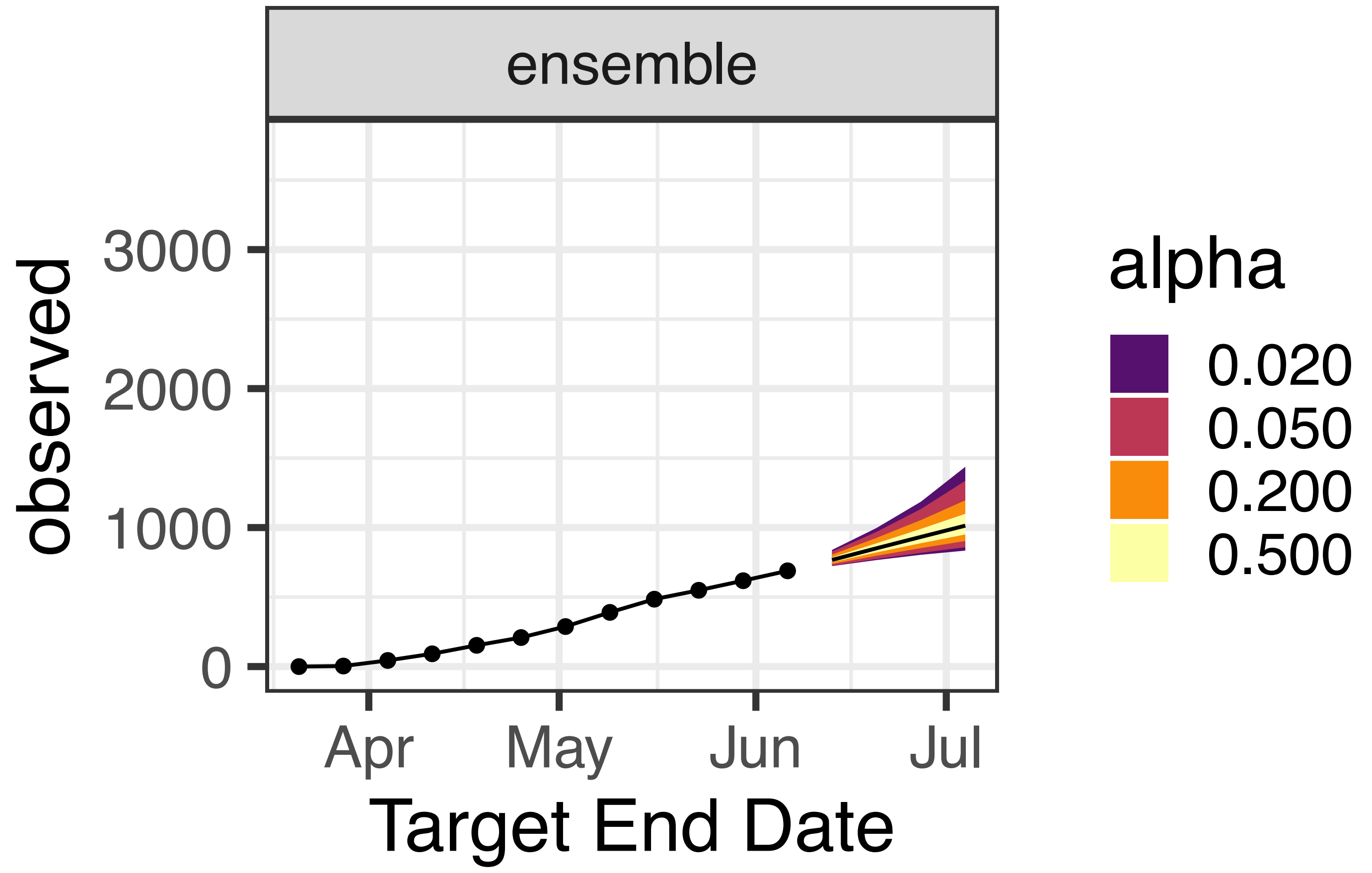
Ensemble model to the rescue?



Ensemble model to the rescue?

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 (or 7) quantiles of a predictive distribution:

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

The predictive median

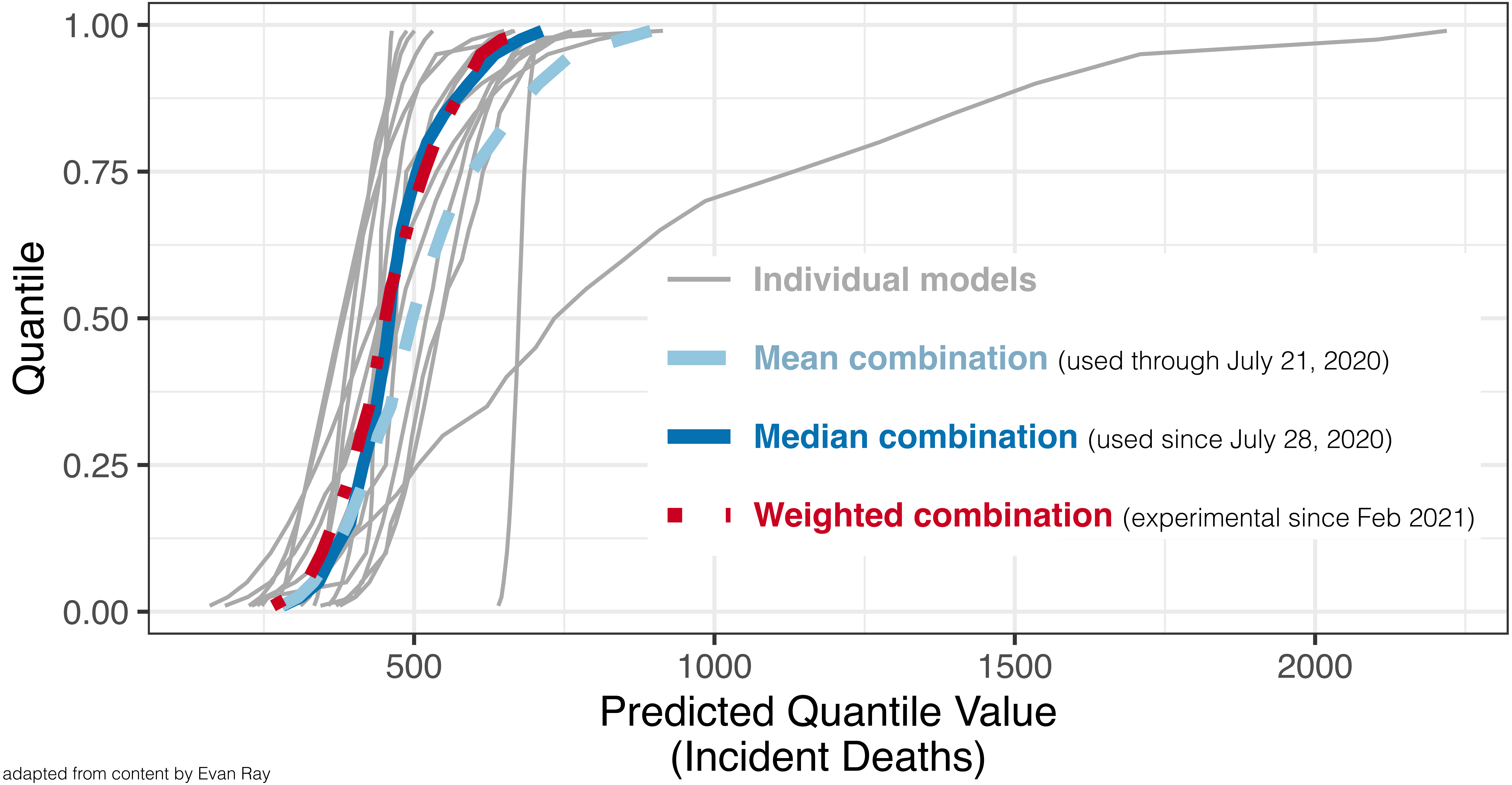
Limits of a 98% prediction interval

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

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

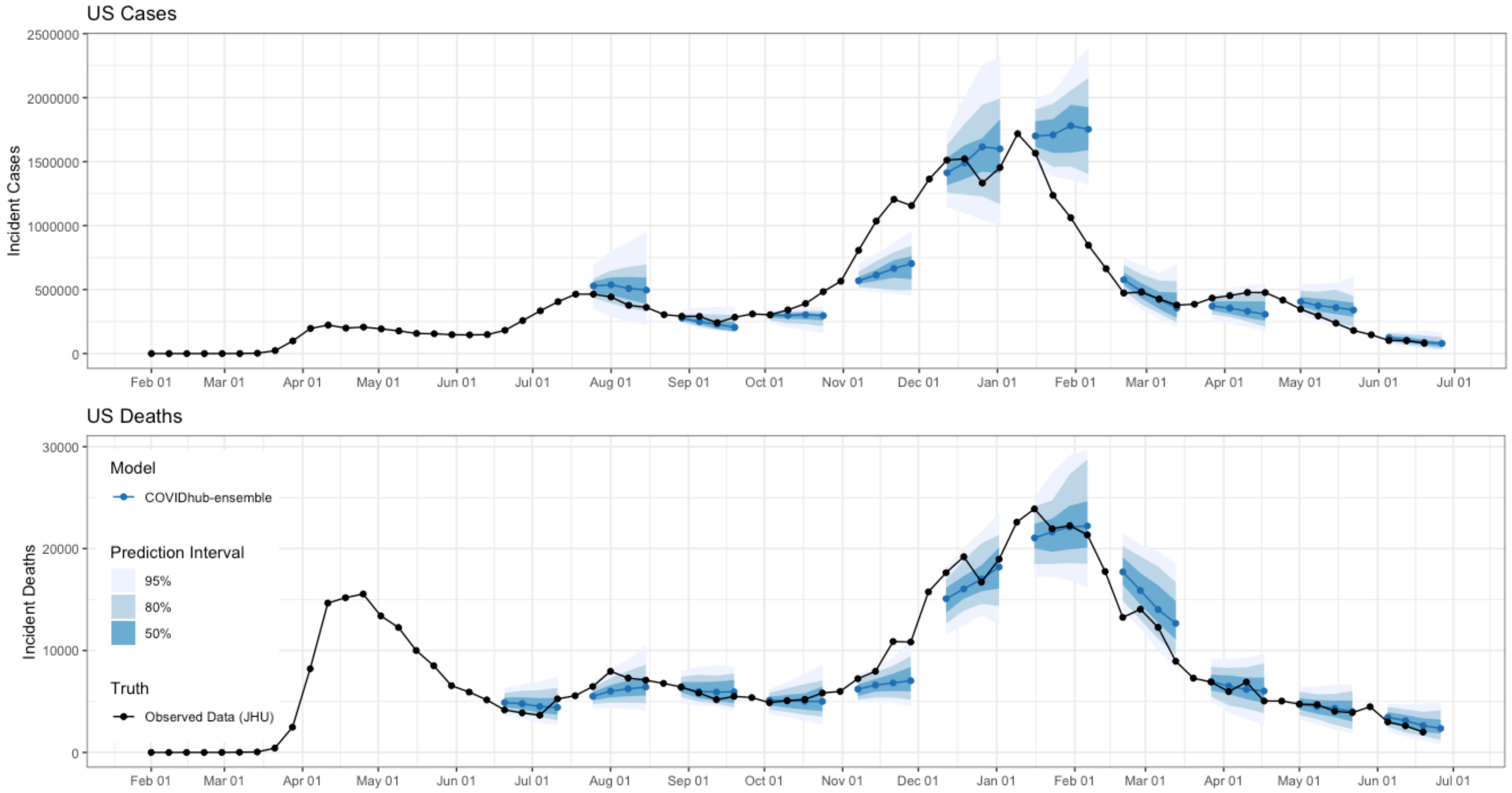
slide adapted from content by Evan Ray

Building an Ensemble: View 2

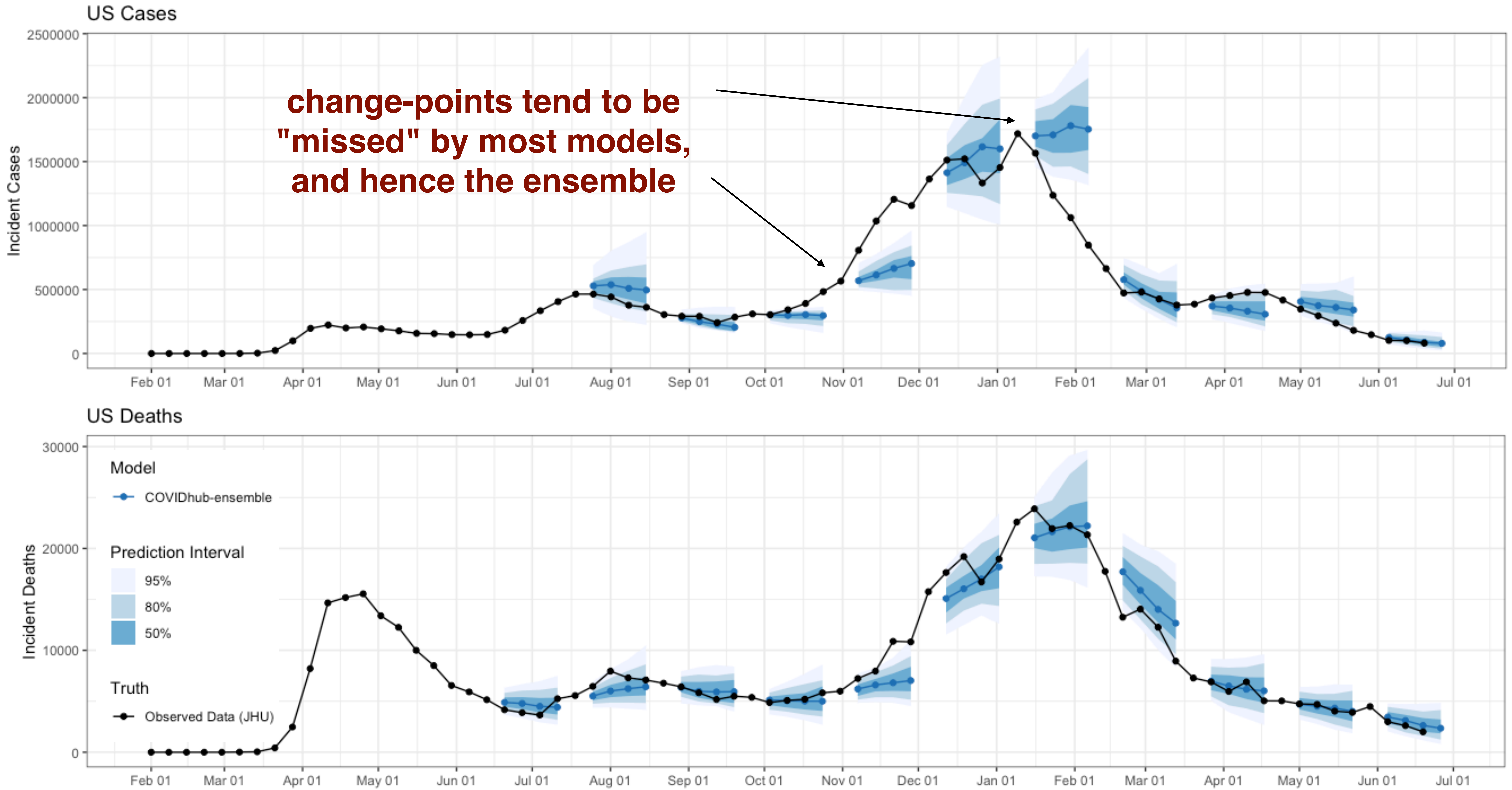


slide adapted from content by Evan Ray

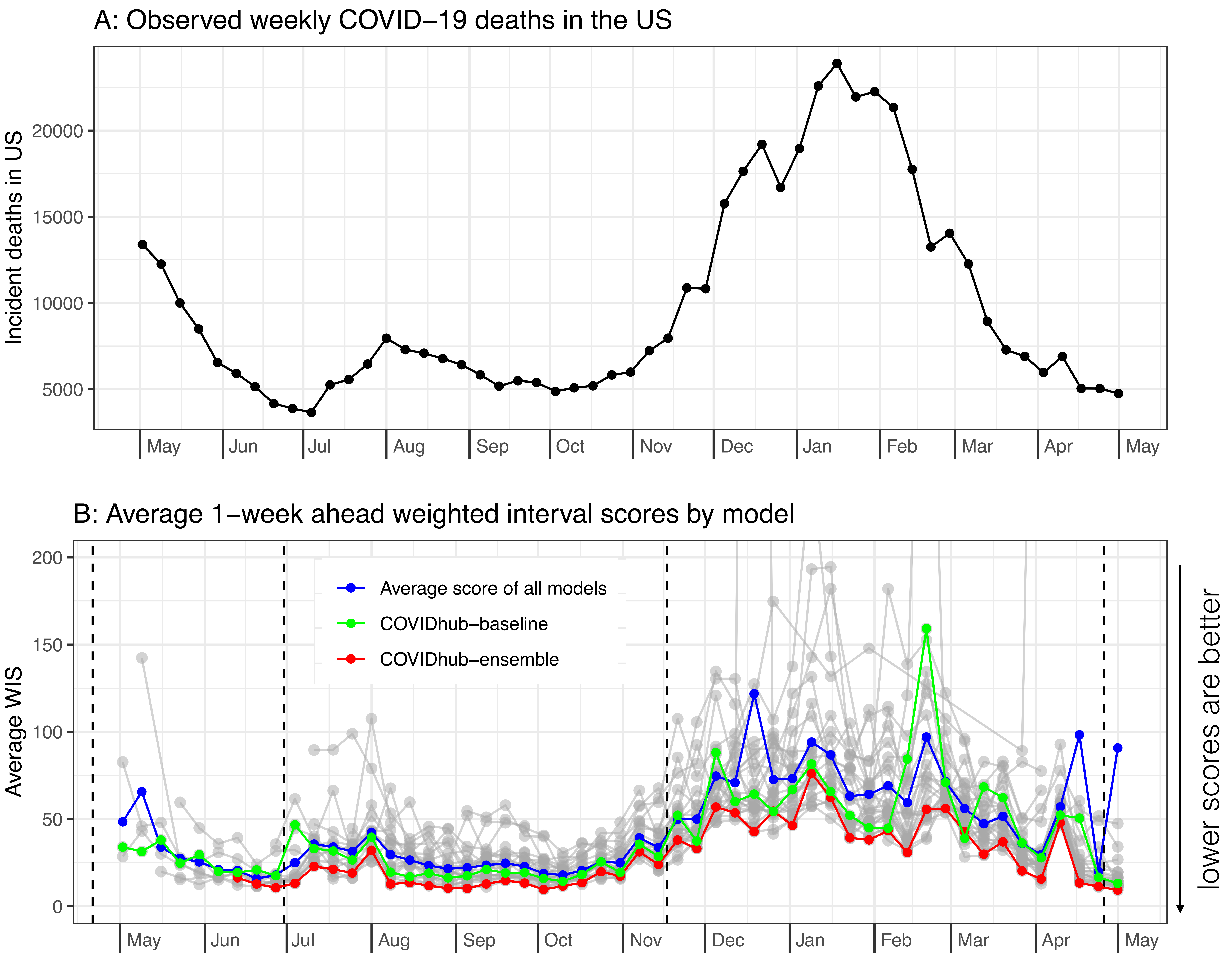
Ensemble forecasts over time



Ensemble forecasts over time



Forecast accuracy over time



Ensembles have shown top performance

Metrics computed using forecasts from the last 6 months...

Forecasts of COVID-19 cases

model	rel. WIS	95% cov.
COVIDhub-ensemble	0.90	0.72
USC-SI_kJalpha	0.95	0.44
CU-select	0.99	0.53

⋮

- 24 individual models for cases*
- 3 models (13%) better than baseline at 3 & 4 weeks ahead

Forecasts of COVID-19 deaths

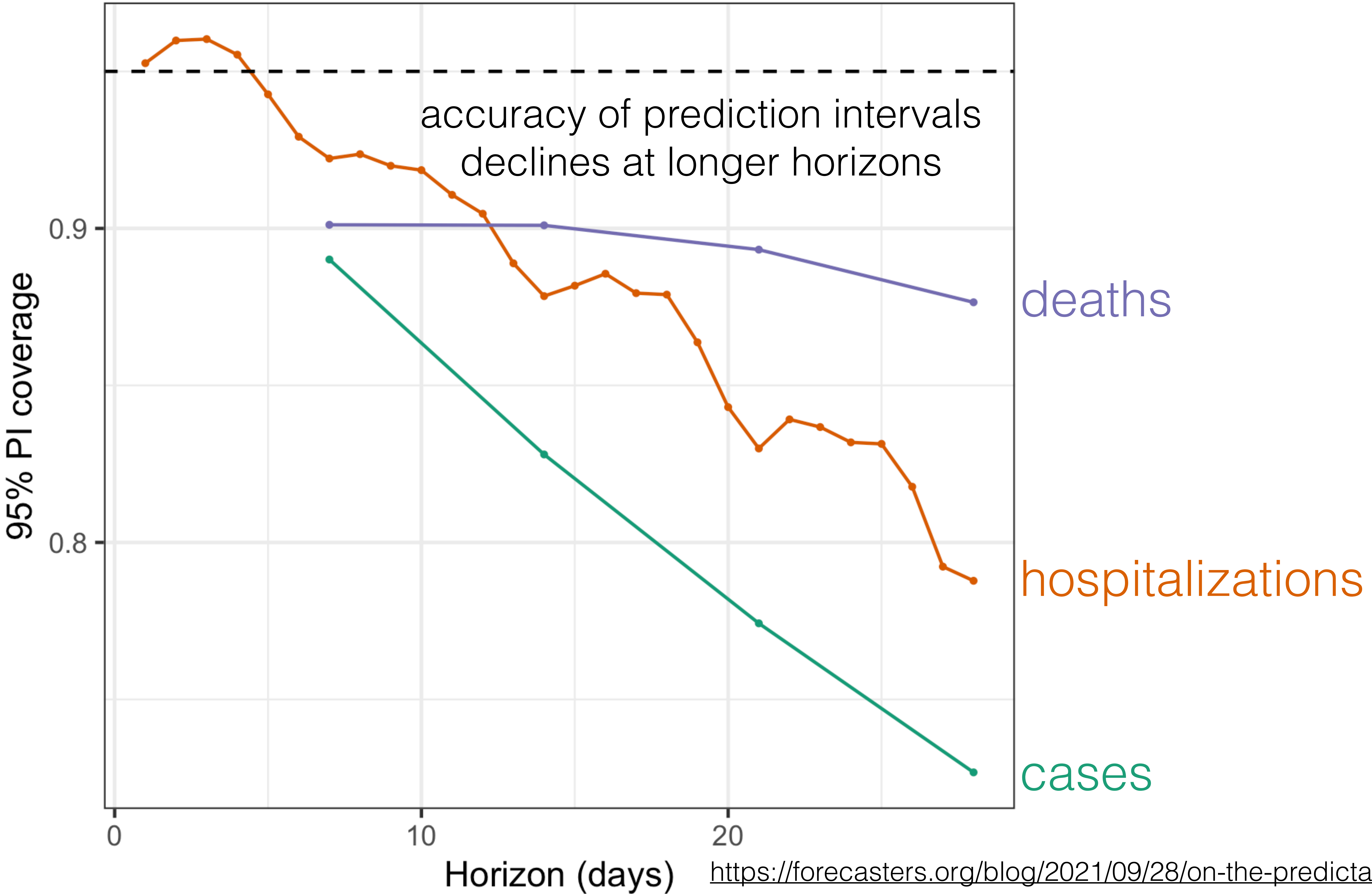
model	rel. WIS	95% cov.
COVIDhub-ensemble	0.54	0.94
CMU-TimeSeries	0.57	0.91
JHU_CSSE-DECOM	0.59	0.76
IHME-CurveFit	0.60	0.52
SteveMcConnell-CovidComplete	0.61	0.75
MOBS-GLEAM_COVID	0.61	0.78
UMass-MechBayes	0.62	0.90

⋮

- 33 individual models for deaths*
- 24 models (73%) better than baseline at 4 weeks ahead

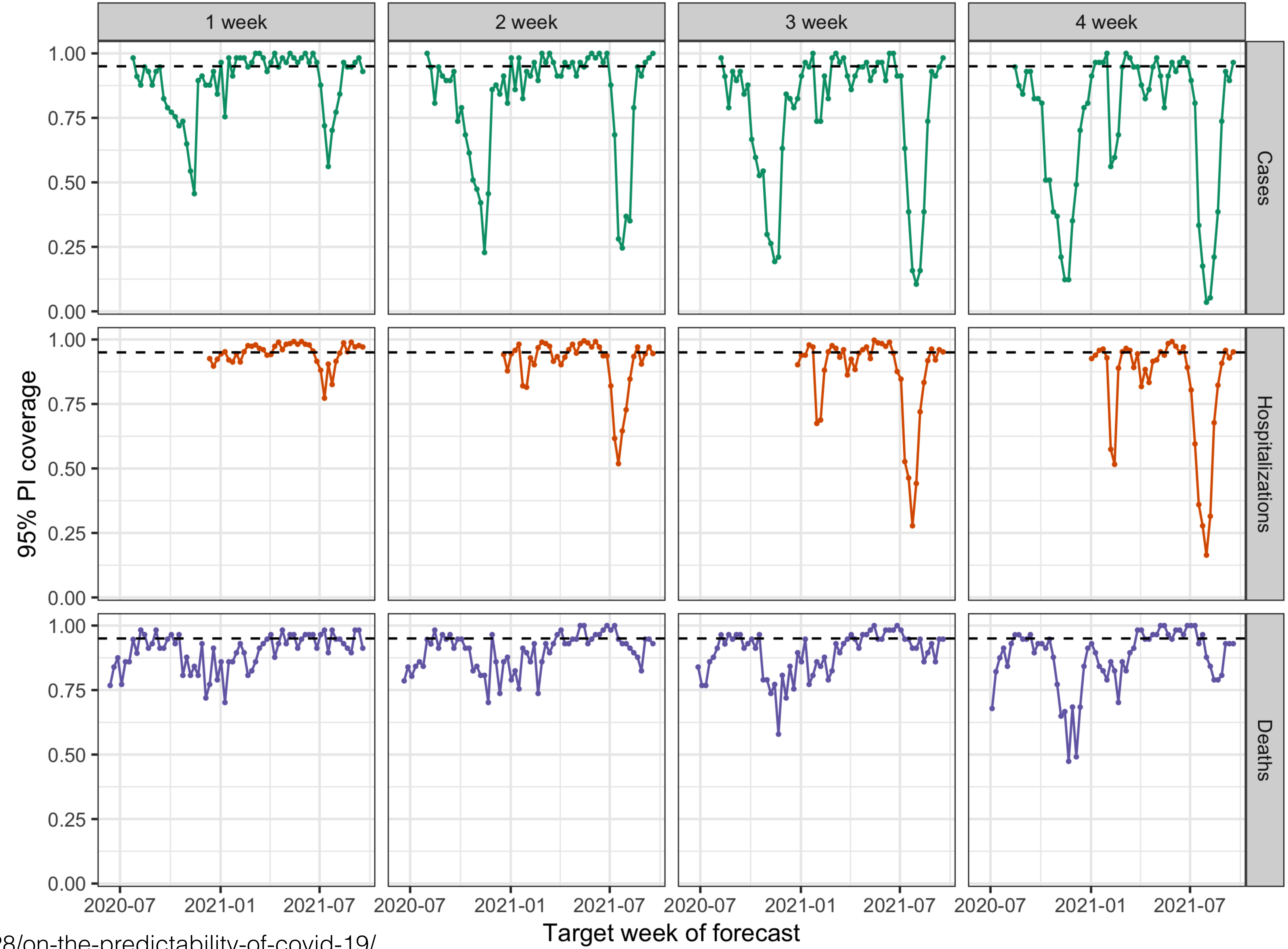
* Models had to have submitted forecasts for at least 13 of the past 26 weeks.
More scores at: https://covid19forecasthub.org/eval-reports/?state=US&week=2021-09-22

Overall, ensemble accuracy varies by target and error increases as we look further into the future



https://forecasters.org/blog/2021/09/28/on-the-predictability-of-covid-19/

Interval coverage over time reveals systemic problems



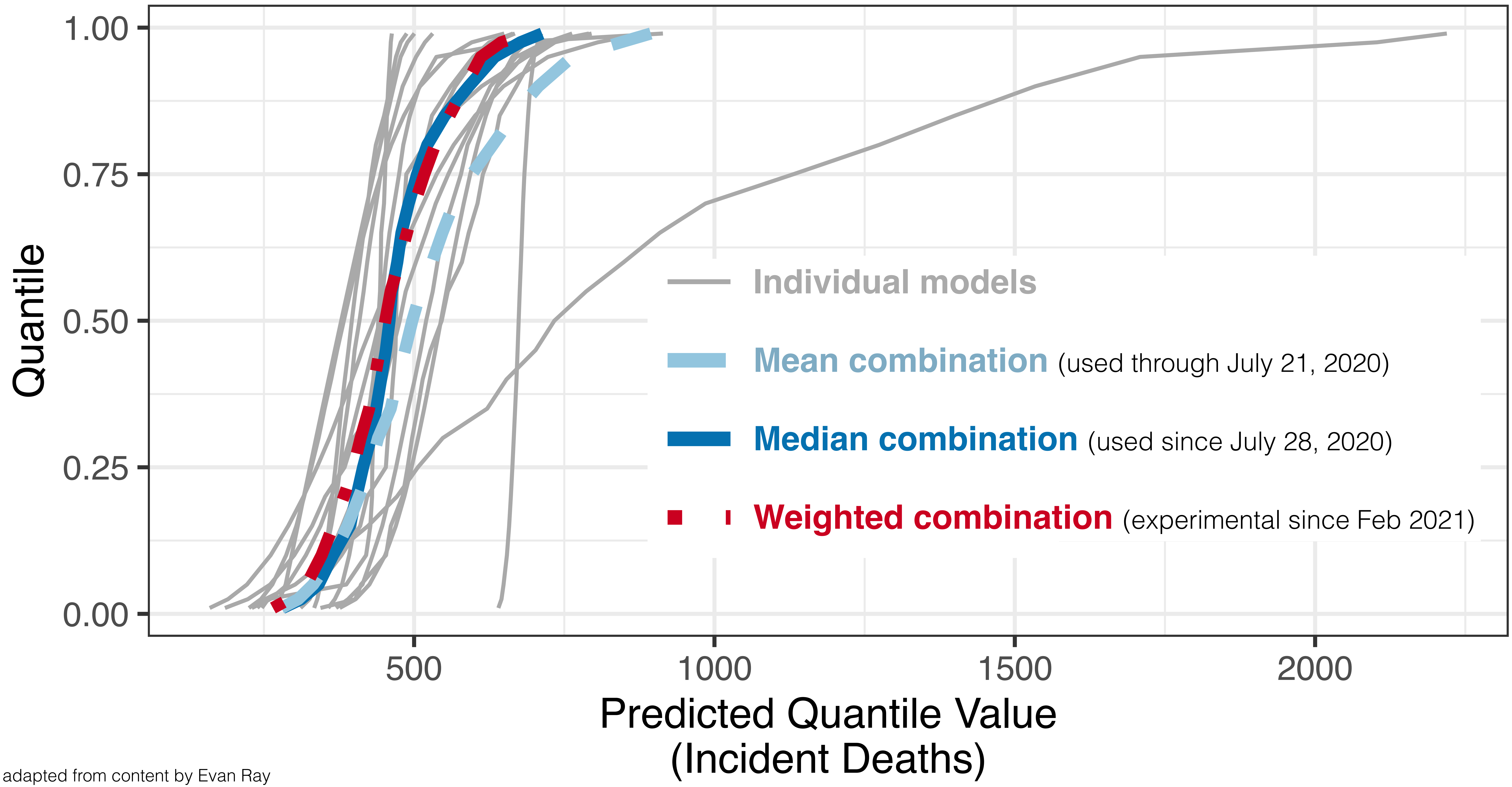
https://forecasters.org/blog/2021/09/28/on-the-predictability-of-covid-19/

Interval coverage over time reveals systemic problems

target suspended as of late Sept 2021

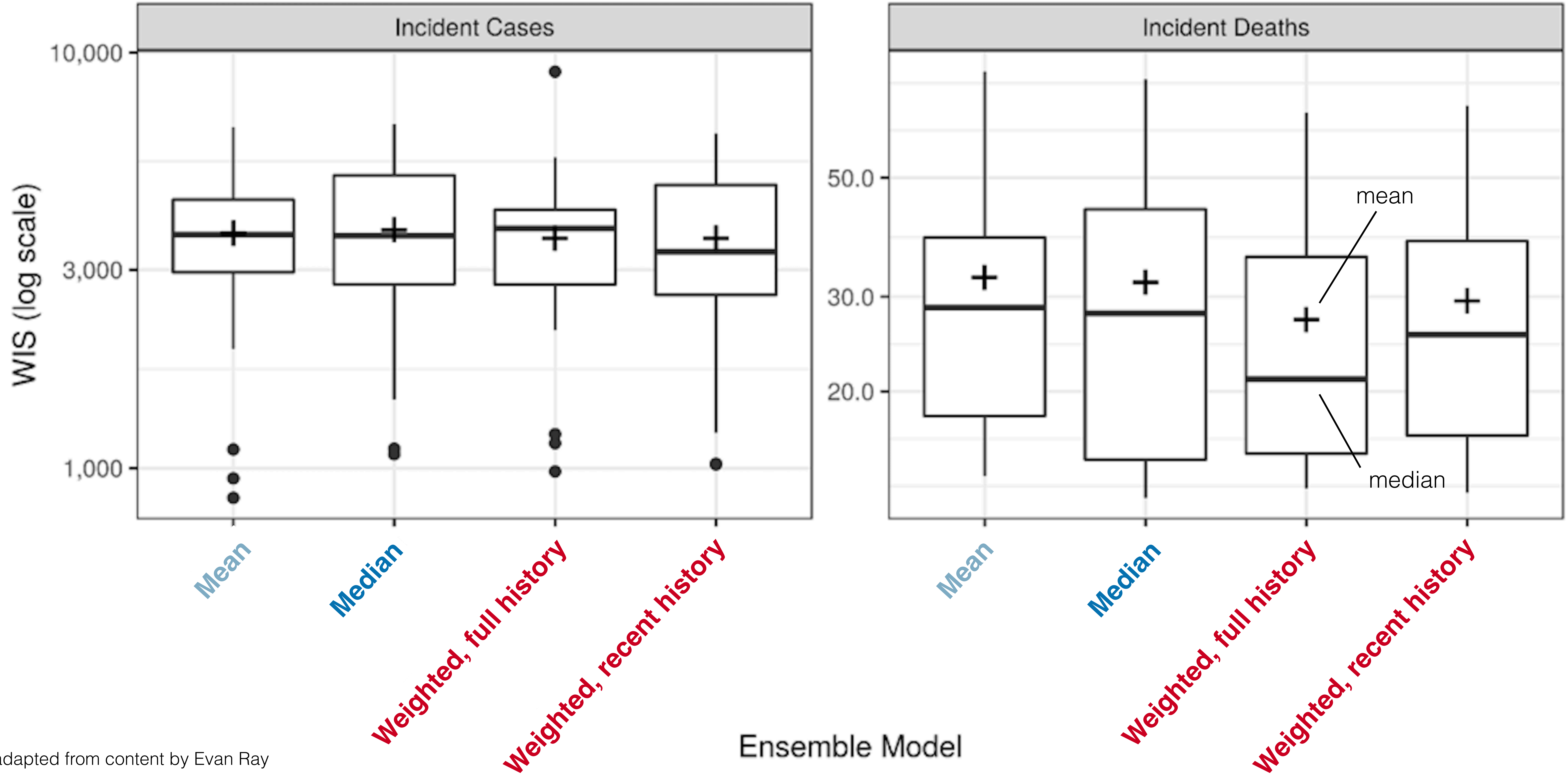
https://forecasters.org/blog/2021/09/28/on-the-predictability-of-covid-19/

Building an Ensemble: View 2



slide adapted from content by Evan Ray

Weighted ensembles yield modest accuracy gains

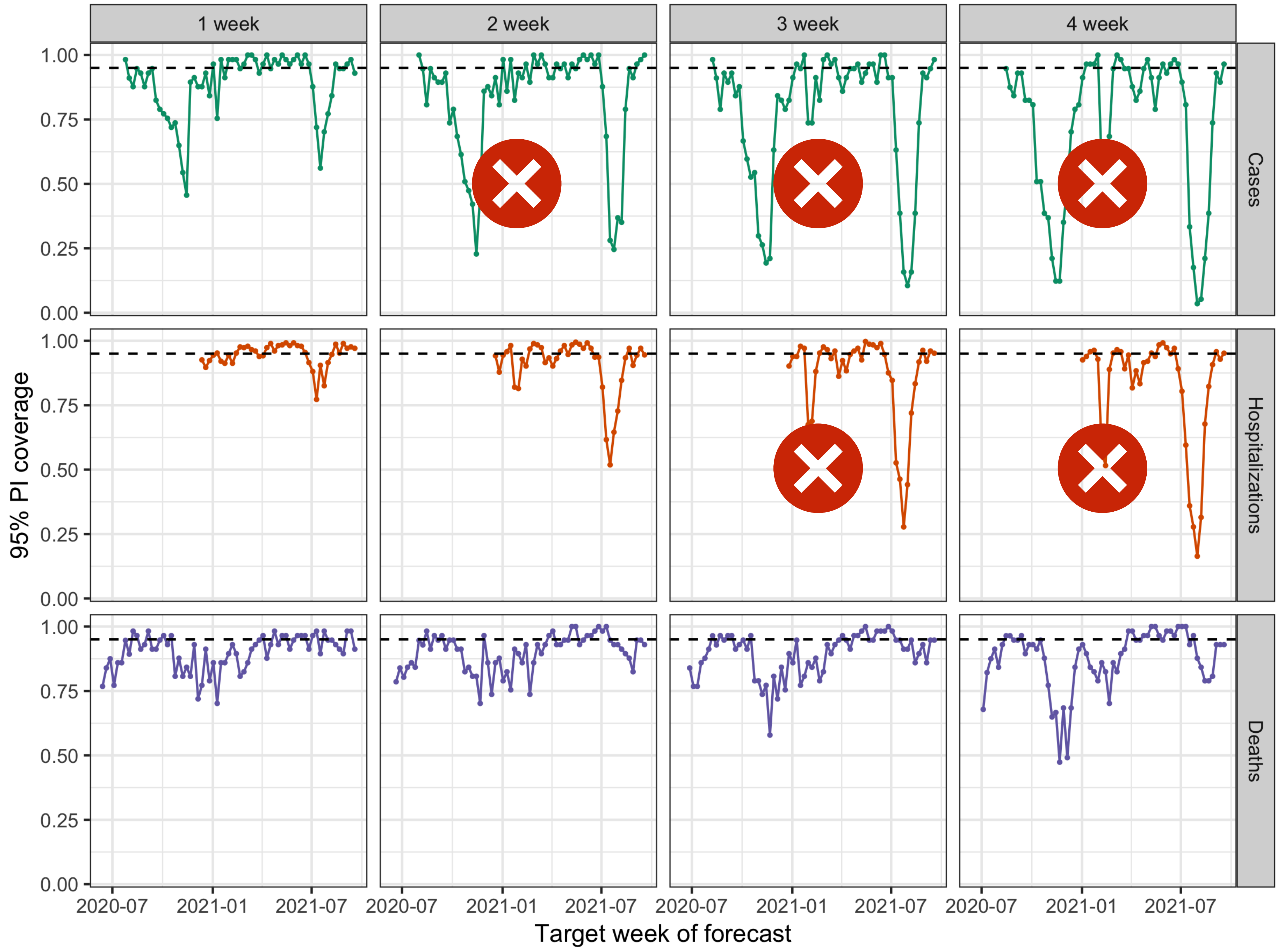


slide adapted from content by Evan Ray

Lessons from COVID-19 forecasting

1. In the very near term, forecasts for all outcomes are usually accurate.
   For deaths, forecasts up to 4 weeks are reliable.

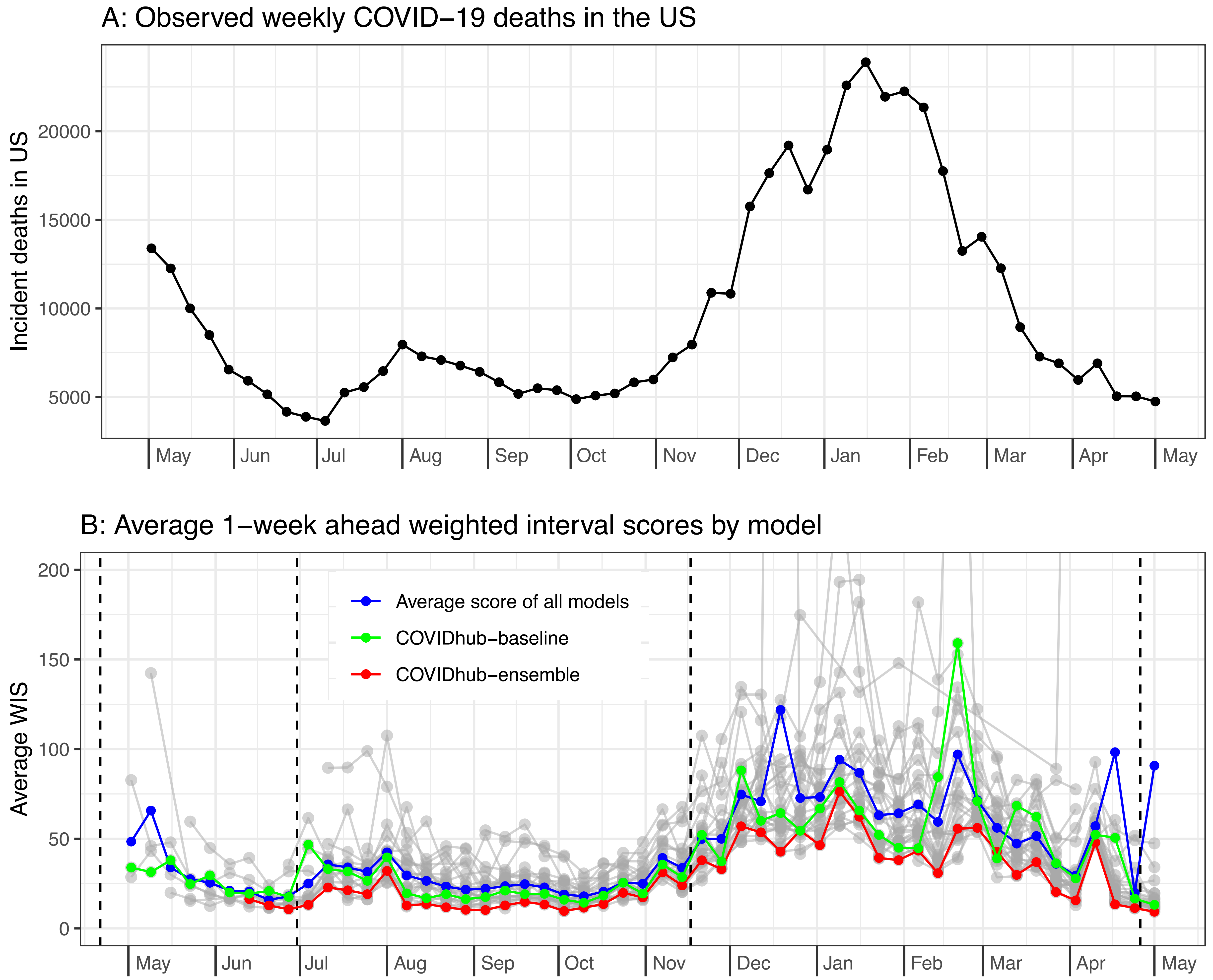
2. Treat any specific 4+ week-ahead prediction with skepticism!
   (Especially for cases.)



Lessons from COVID-19 forecasting

3. Individual model performance varies substantially.

4. Data quality issues (e.g. revisions after the fact and imperfect "ground truth" measurements) pose challenges for models.



Lessons from COVID-19 forecasting

5. Don't rely on one model, unless it is an ensemble. And even then, remain skeptical!

model	rel. WIS	95% cov.
COVIDhub-ensemble	0.54	0.94
CMU-TimeSeries	0.57	0.91
JHU_CSSE-DECOM	0.59	0.76
IHME-CurveFit	0.60	0.52
SteveMcConnell-CovidComplete	0.61	0.75
MOBS-GLEAM_COVID	0.61	0.78
UMass-MechBayes	0.62	0.90

Afterword

September 17, 2021

REUTERS/Joshua Roberts

The use of multiple perspectives to gain insight into the truth is not just a statistical idea, but one rooted in story-telling.

American Journal of Sociology > Volume 124, Number 2

Faulkner's Assembly of Memories into History: Narrative Networks in Multiple Times[1]

John F. Padgett

The core stylistic problematic, indeed the obsession, for Faulkner was how ***the diverse multiple perspectives*** of his narrators and characters do or do not blend together to ***make a collective ensemble***, which at minimum is gripping to narrators and which at maximum is true—in the sense that an elegant multiperspectival account emerges from the composition of partial perspectives, consistent with all known facts about the behaviors of characters in the story.

The use of multiple perspectives to gain insight into the truth is not just a statistical idea, but one rooted in story-telling.

American Journal of Sociology > Volume 124, Number 2

Faulkner's Assembly of Memories into History: Narrative Networks in Multiple Times[1]

John F. Padgett

The core stylistic problematic, indeed the obsession, for Faulkner was how ***the diverse multiple perspectives*** of his narrators and characters do or do not blend together to ***make a collective ensemble***, which at minimum is gripping to narrators and which at maximum is true—in the sense that an elegant multiperspectival account emerges from the composition of partial perspectives, consistent with all known facts about the behaviors of characters in the story.

I'm telling you stories. Trust me.

from "The Passion" by Jeanette Winterson

COVID-19 ForecastHub

https://covid19forecasthub.org/

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

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

CMU Delphi Group: Ryan Tibshirani, Roni Rosenfeld, Logan Brooks, and others

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

Modeling groups: Over 80 groups at various institutions have contributed forecasts

UMassAmherst
School of Public Health
& Health Sciences
Biostatistics and Epidemiology

Evan Ray

Thank you!

With acknowledgments to Dr. Evan Ray,
all members of Reich Lab and COVID-19 Forecast Hub team,
CDC collaborators, and modeling contributors.

The Reich Lab and the COVID-19 Forecast Hub have 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.

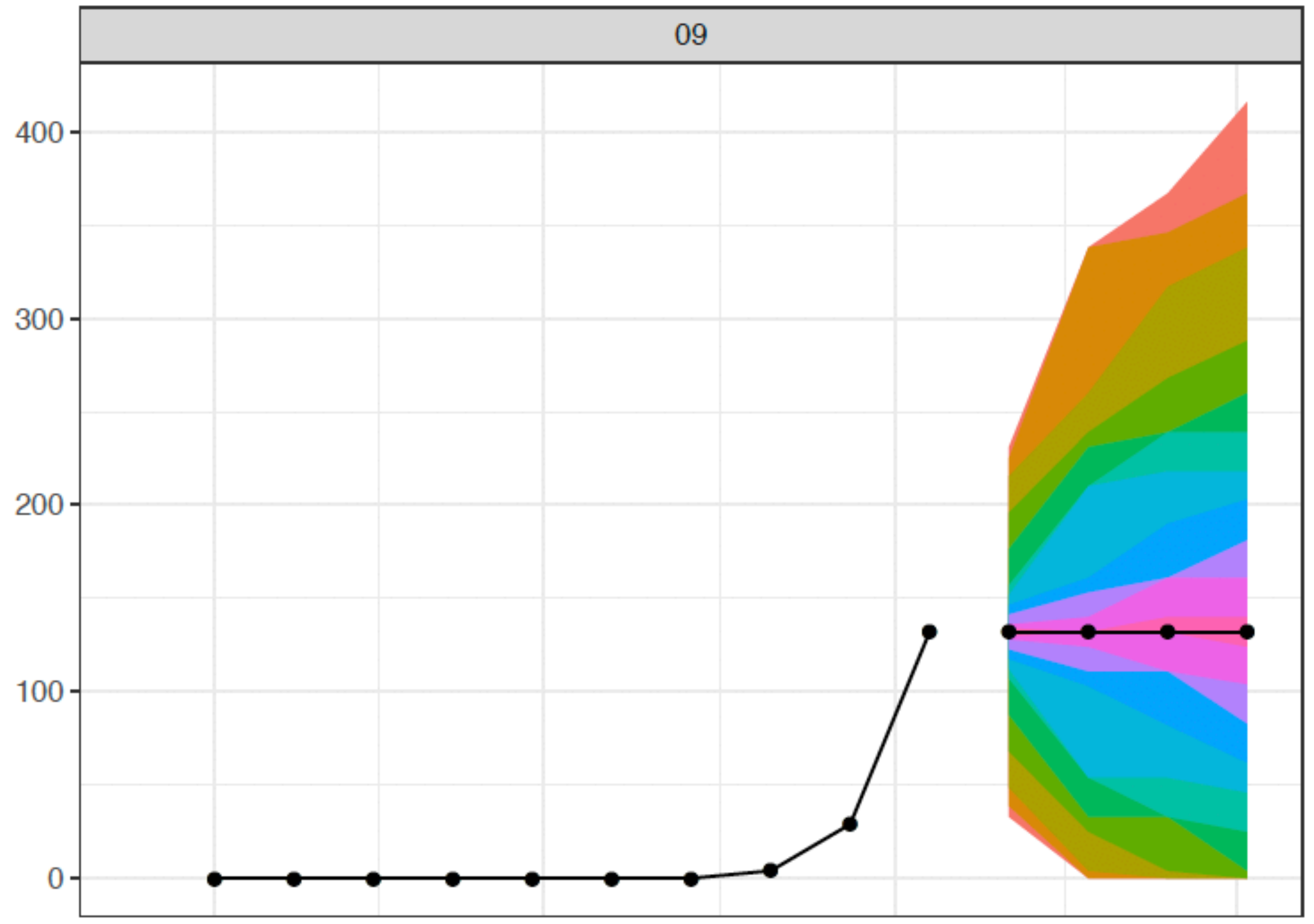
reichlab.io
reichlab

Reich Lab | AT UMASS AMHERST

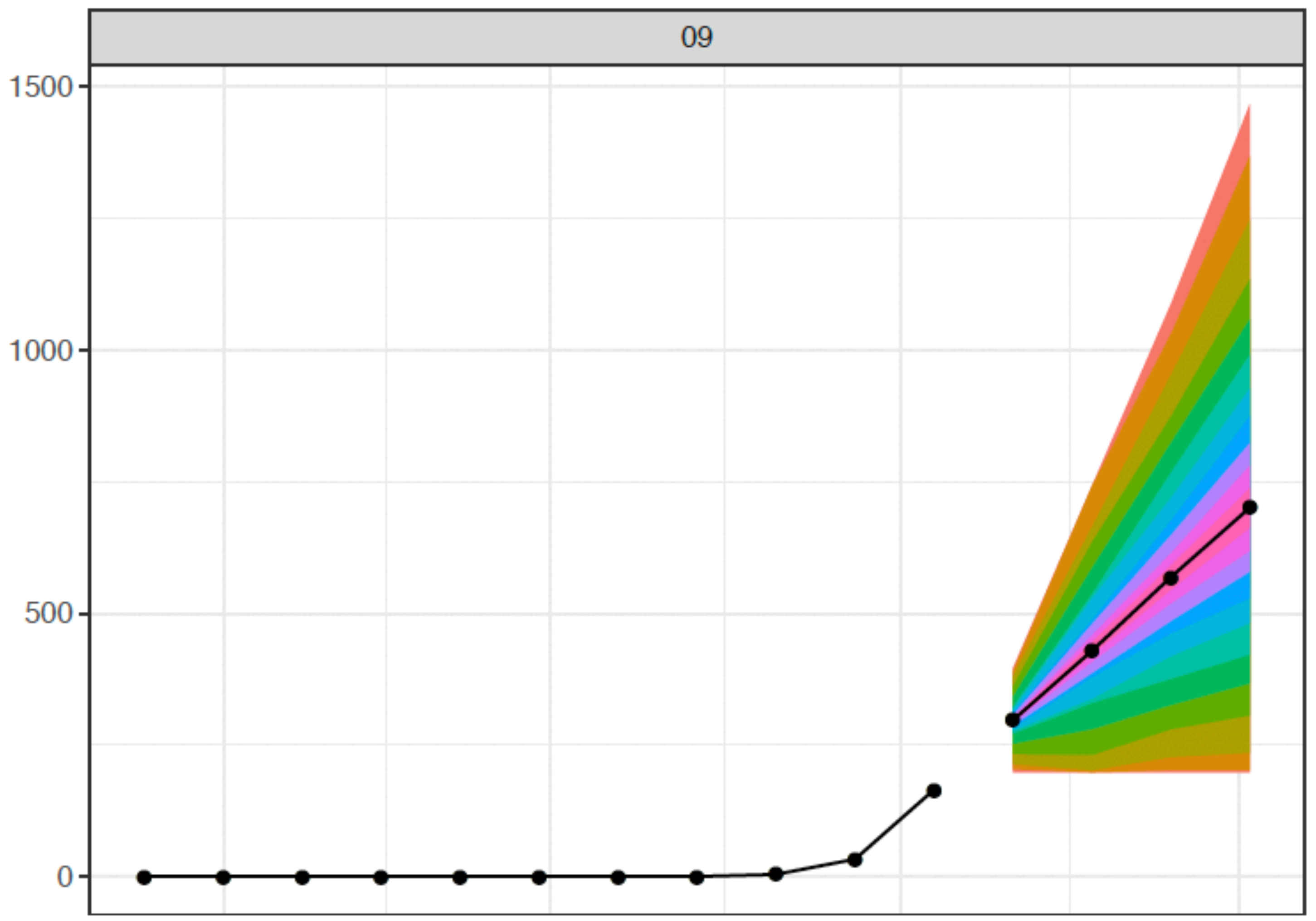
Baseline Model

- 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



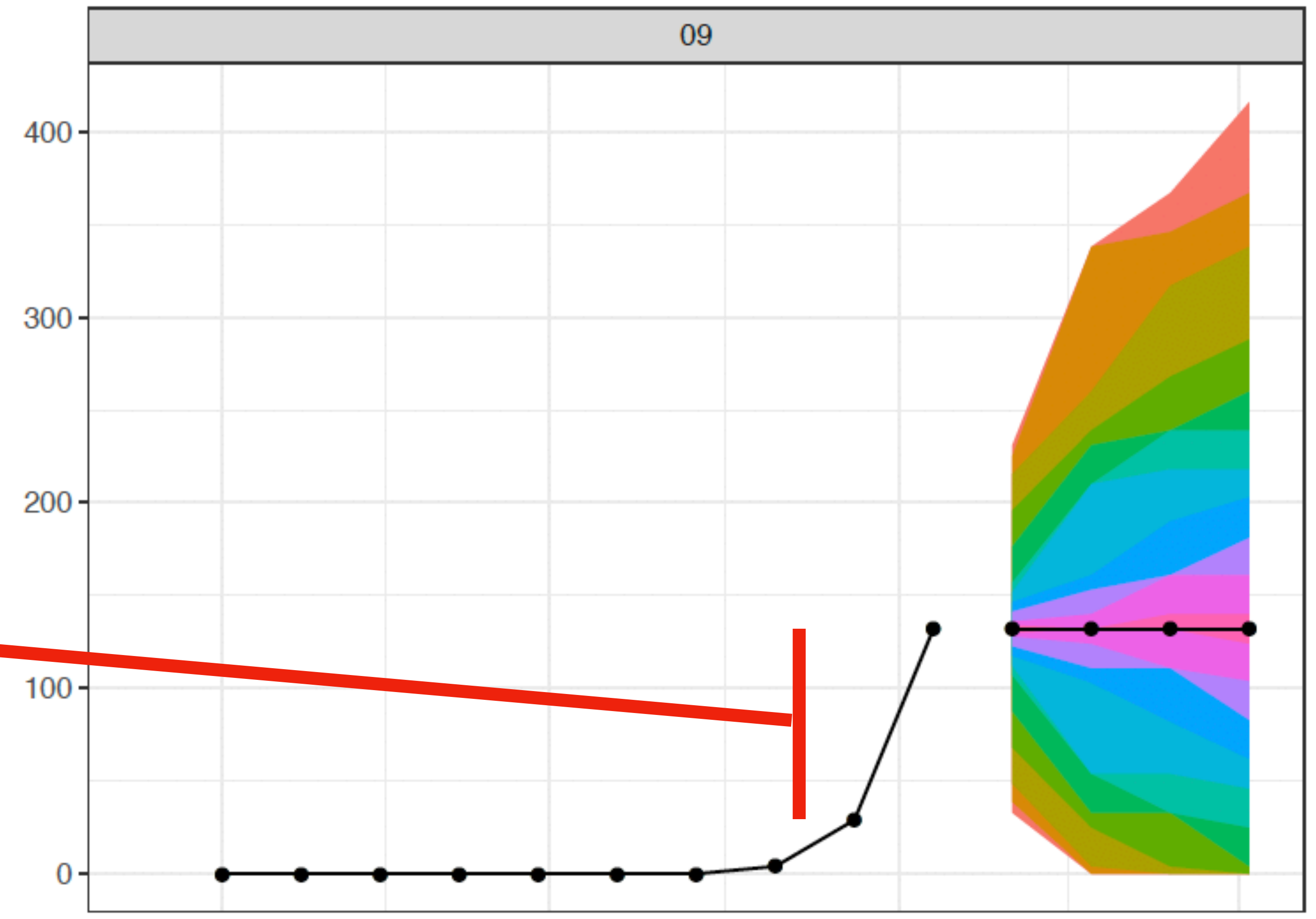
Cumulative Deaths



Baseline Model

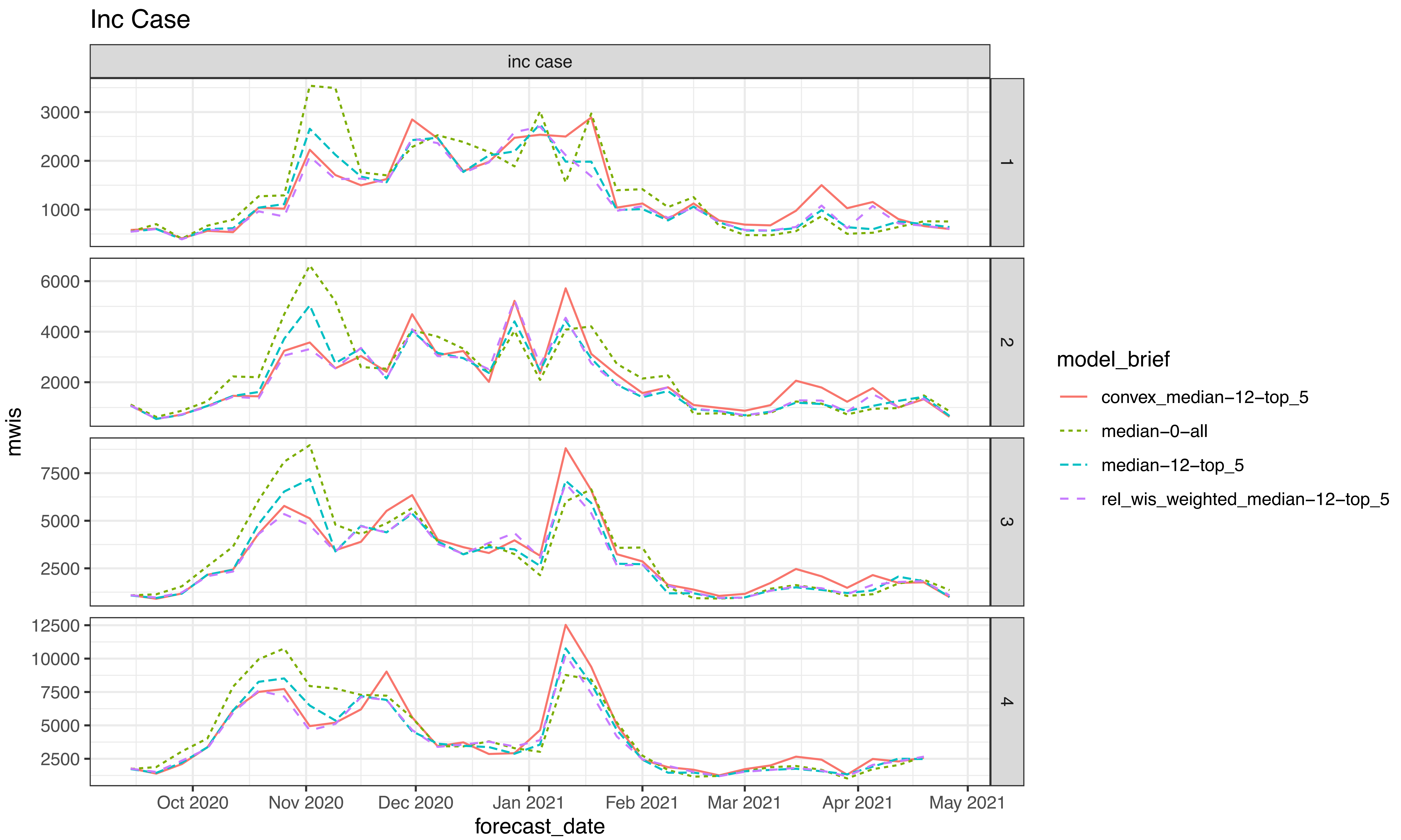
- Procedure:
  - Compute first differences of historical incidence:

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

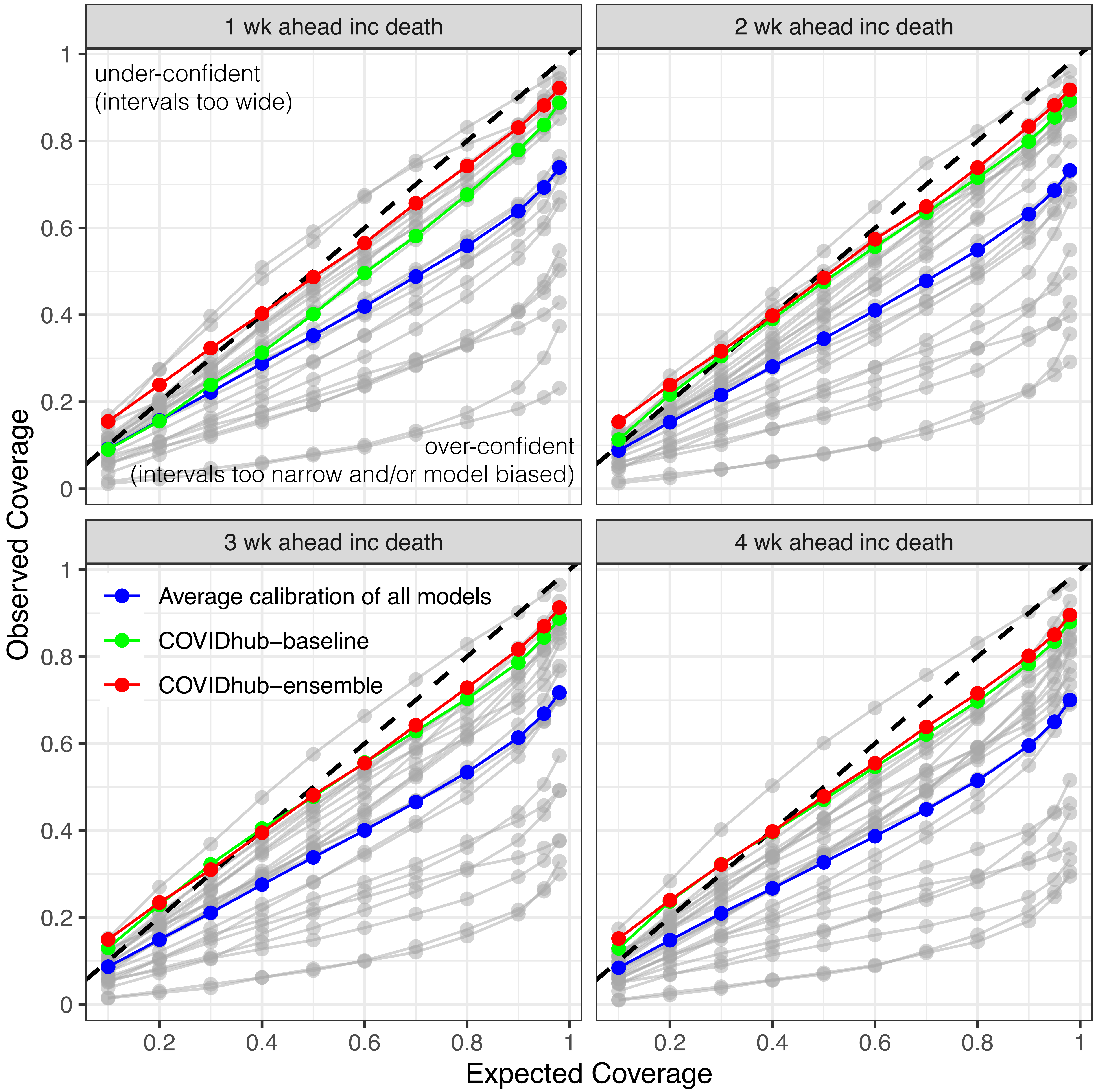


  - Collect first differences and their negatives
  - Sample first differences and add to last observed incidence; take quantiles of the resulting distribution
  - Iterate for horizons > 1
  - Adjustments for “niceness”:
    - Force median = last observed incidence
    - Truncate at 0

Evaluation Over Time — Cases

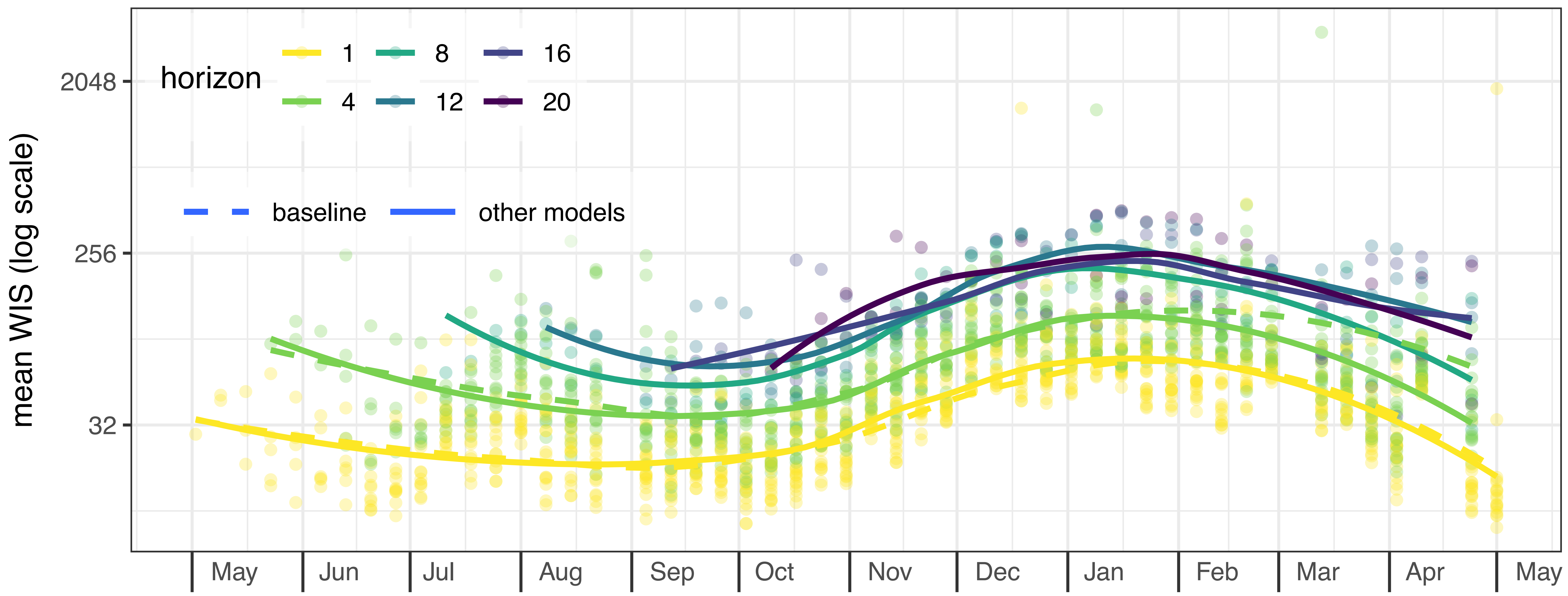


Model calibration — Deaths



Errors increase with horizon

B: mean WIS across time, stratified by forecast horizon



C: 95% prediction interval coverage across time, stratified by forecast horizon