COVID-19: What's New for May 29, 2020

Main updates on IHME COVID-19 predictions since May 26, 2020

Updated IHME COVID-19 Projections: new death model

Consistently collecting and reporting health data in a timely manner is far from easy, even in settings where monitoring is well resourced. Amid a global pandemic, where understanding of the disease can change each day and many agencies have to multi-task epidemic response with reporting, health information system functions can be quickly strained.

Over the last few months, our team has developed various approaches to better process highly noisy COVID-19 death data inputs (e.g., three-day smoothing algorithms prior to modeling, as described in the May 4 estimation update); incorporated leading variables such as COVID-19 cases and hospitalizations; and refined model parameters to allow for more variable functional forms in predicting epidemic peaks and shapes. These improvements have occurred in tandem with the world's expanding knowledge base of COVID-19 and how data systems have sought to keep pace with each location's epidemic trajectory.

Today's release involves a new approach to modeling deaths, an improvement spurred by the need to develop methods that are more robust in handling variable COVID-19 input data and perform well for smaller epidemics. This method has been implemented for the US, and we are now running it for all currently included locations. In parallel, we are expanding included locations to a more global scale and incorporating additional covariates to inform the transmission dynamics component of our modeling strategy.

A new modeling approach for COVID-19 death predictions

Our improved COVID-19 death modeling strategy can be summarized as a five-step process, which we detail below. In sum, we are now using a more parsimonious model than CurveFit, combining deaths and cases into a time series of estimated COVID-19 deaths for the past and predicted eight days into the future. This update builds off of our previous model which used COVID-19 cases as a leading indicator (added in our April 17 release). We adjust reported cases by taking into account overall testing volumes and populations tested (i.e., symptomatic individuals and active case detection efforts among high-risk populations in factories, prisons, nursing homes, and homeless shelters).

This new death model involves five steps:

- **Step 1: Fit a spline to log cumulative deaths *versus* log cumulative cases.** This step generates estimates of the case-fatality rate for COVID-19 (as captured by the slope of this curve). When case-fatality rates (CFR) are decreasing, this reflects increases in testing and shifting toward screening non-symptomatic or mildly symptomatic individuals.
- **Step 2:** Using CFR estimates in Step 1, COVID-19 deaths are predicted based on cases over time.
- **Step 3:** Reported daily deaths (unsmoothed, so "noisy") and predicted daily deaths based on cases (again unsmoothed, so "noisy") are combined using a spline fit in log daily death space.
- **Step 4:** Drawing from the residuals estimated in Step 3, we estimate the standard deviation and sample residuals to produce 1,000 draws of COVID-19 deaths per location-day.

- **Step 5:** We then use the resampled death draws as inputs for our transmission dynamics model.

So what does all of this mean? While the new death model does not substantially affect estimates for most locations, especially those with larger epidemics like the US, its performance is more robust in handling daily fluctuations in reported deaths and cases, and also for locations with smaller epidemics to date. Incorporating this new model into our broader COVID-19 estimation platform sets the foundation for substantially expanding included locations.

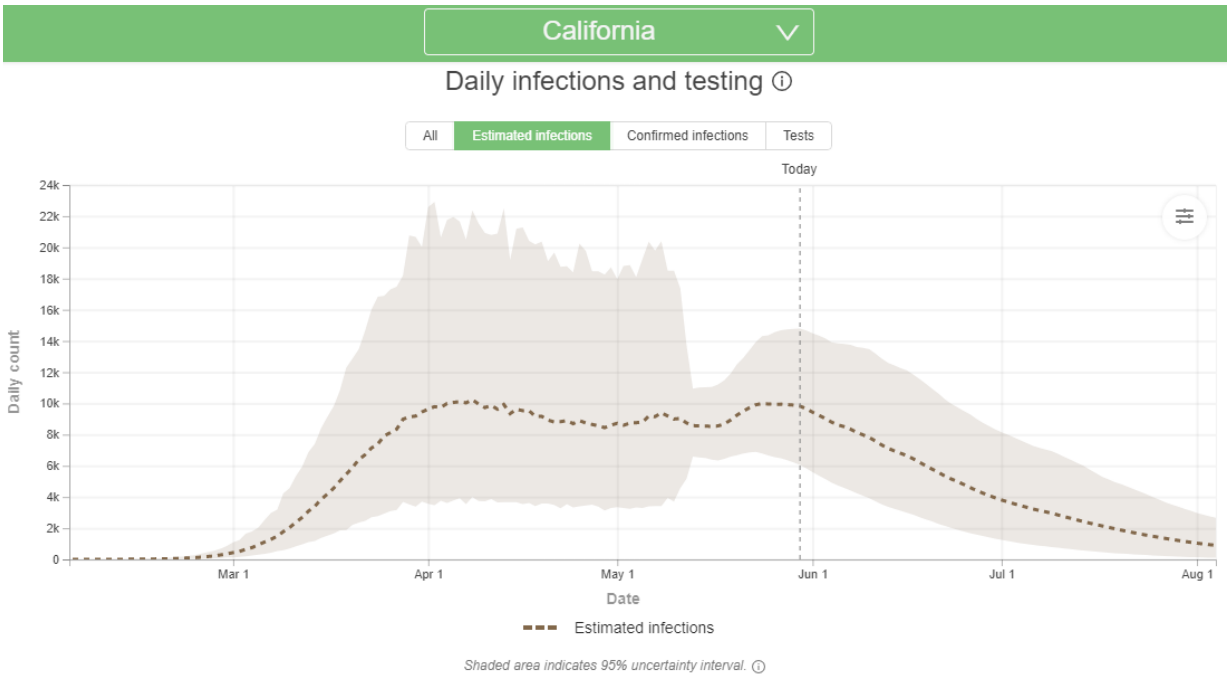
Key US findings from today's release (May 29, 2020)

- Nationally, cumulative COVID-19 deaths could reach 135,109 (estimate range of 123,344 to 157,715) by August. This is higher than projections from the May 26 release, where the cumulative COVID-19 death toll was projected to be 131,967 (estimate range of 115,839 to 173,528); however, the uncertainty intervals overlap considerably.

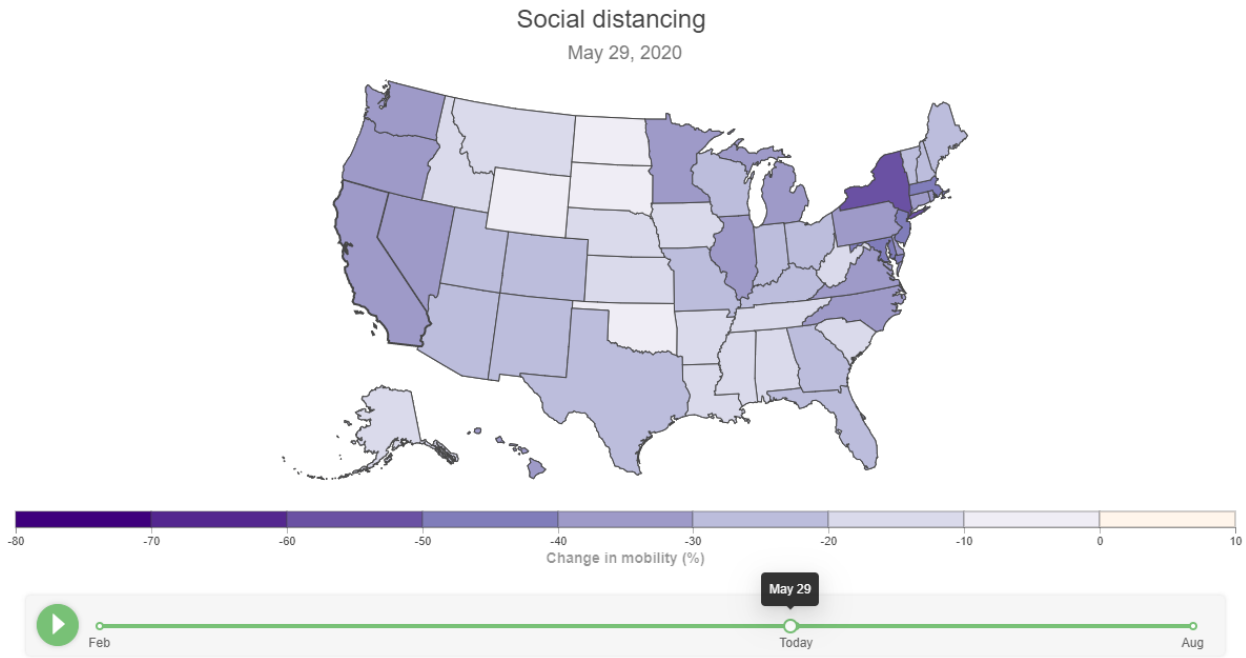
- Based on the latest available data and model updates, California's cumulative COVID-19 deaths could be 12,951 (5,896 to 34,137) by August. These projections are higher than the state's estimates as of the May 26 release (average prediction of 7,558 deaths, with an estimate range of 5,612 to 10,312). Mobility patterns have been drifting upward in California, a factor in the state's higher projected cumulative COVID-19 deaths. The trend in cases is also increasing, even when taking into account testing rates, implying that it is likely that we will see an increase in reported deaths in the coming weeks.

Location	**Total COVID-19 deaths to date (estimates to May 29)**	**Predictions for cumulative COVID-19 deaths through August**
United States (national)	102,159 (102,085 to 102,239)	135,109 (123,344 to 157,715)
New York	29,756 (29,716 to 29,801)	31,295 (30,682 to 32,388)
California	4,058 (4,038 to 4,085)	12,951 (5,896 to 34,137)
New Jersey	11,451 (11,433 to 11,472)	12,577 (12,067 to 13,500)
Massachusetts	6,915 (6,898 to 6,936)	9,003 (7,644 to 12,203)
Pennsylvania	5,311 (5,294 to 5,332)	6,418 (5,827 to 7,554)
Illinois	5,205 (5,188 to 5,224)	6,371 (5,822 to 7,409)
Michigan	5,331 (5,323 to 5,342)	5,986 (5,602 to 6,886)
Results as of 05/29/2020 		

- While national estimates of COVID-19 infections are declining, several states may be experiencing stagnated progress or increased infections – concerning trends given upward trends in mobility and eased social distancing measures throughout the US. These states include Arizona, California (see screenshot below), Utah, and Wyoming, among others.



- As shown in the screen shot below, the majority of states are increasingly closer to baseline levels of human mobility (lightest shades of purple). New York state remains the only location in the US where mobility remains more than 50% lower than baseline levels; however, even New York has seen mobility rates increasing in recent days.



Data updates since our last release on May 26, 2020

Data and locations

- For all currently included locations, we have added reported data points on COVID-19 deaths, cases, hospitalization, testing, and mobility, as well as available information on social distancing policies, through May 28.
- Currently included locations are the US (nationally) and 50 states plus the District of Columbia, Puerto Rico, four provinces in Canada, European Economic Area (EEA) countries plus Switzerland and Argentina, 19 states in Brazil, Bolivia, Chile, Colombia, Cuba, the Dominican Republic, Ecuador, Egypt, Honduras, Israel, Japan, Malaysia, seven states in Mexico, Moldova, Panama, Peru, the Philippines, Russia, Serbia, South Korea, Turkey, and Ukraine. Three EEA countries – Germany, Italy, and Spain – also have subnational estimates at the first administrative level.

What's in the development pipeline for IHME COVID-19 predictions

Before we introduce new model components or improvements to our current analytical platform for predictions, IHME's COVID-19 development team members test these additions or changes.

Based on currently available data and model testing progress, our immediate- and medium-term priorities are as follows:

- **COVID-19 projections for the world.** Data collation and processing for all locations and countries worldwide continue, and we are conducting development model runs for a global location set, which we hope to release next week.

- **Additional key epidemic drivers**. Pending data availability across currently included locations, we are exploring how to incorporate additional model covariates such mask or facial covering use by the broader public, human contact rates, household size, and use of public transit, as well as exploring whether trends in diseases such as pneumonia can be used to predict trends in COVID-19.

A note of thanks

We would like to extend a special thanks to the Pan American Health Organization (PAHO) for key data sources; our partners and collaborators in Argentina, Brazil, Bolivia, Chile, Colombia, Cuba, the Dominican Republic, Ecuador, Egypt, Honduras, Israel, Japan, Malaysia, Mexico, Moldova, Panama, Peru, the Philippines, Russia, Serbia, South Korea, Turkey, and Ukraine for their support and expert advice; and to the tireless data collection and collation efforts of individuals and institutions throughout the world.

In addition, we wish to express our gratitude for efforts to collect social distancing policy information in Latin America to University of Miami Institute for Advanced Study of the Americas (Felicia Knaul, Michael Touchton); Fundación Mexicana para la Salud with support from the GDS Services International: Tómatelo a Pecho A.C.; and Centro de Investigaciones en Ciencias de la Salud, Universidad Anáhuac

(Héctor Arreola-Ornelas); Lab on Research, Ethics, Aging and Community-Health at Tufts University (REACH Lab) and the University of Miami Institute for Advanced Study of the Americas (Thalia Porteny).

Further, IHME is grateful to the Microsoft AI for Health program for their support in hosting our COVID-19 data visualizations on the Azure Cloud.

For all COVID-19 resources at IHME, visit http://www.healthdata.org/covid.
Questions? Requests? Feedback? Please contact us here.