Stanford team uses data to help California track and prevent COVID-19

July 24 2020

As the COVID-19 pandemic begins to spike again in California, a team of Stanford modeling experts is working around the clock to pump data into a new assessment tool that is helping California hospitals and public health officials determine their next moves.

The California COVID Assessment Tool, or CalCAT, contains assessments of the spread of COVID-19 in short-term forecasts of disease trends, and presents scenarios of the course of the disease across the 21 counties that represent 95% of the cumulative cases in the Golden State.

Instead of relying on one or two projection models—as some countries and U.S. states did when the pandemic first hit their shores—the CalCAT tool incorporates COVID-19 estimates from a number of respected organizations, including Stanford, UCLA, MIT, Johns Hopkins University and the Imperial College of London. The RAND Corporation focuses on long-term scenarios if shelter-in-place orders are lifted and non-essential business and schools reopen.

"It's like using the wisdom of the crowd," said Jeremy Goldhaber-Fiebert, an associate professor of medicine at Stanford Health Policy and one of the principal investigators of the Stanford-CIDE Coronavirus Simulation Model, or SC-COSMO. "Instead of hanging your hat on one model, you're looking at a range of predictions to help you do planning and forecasting—and leverage the whole community of researchers and analysts who are working on this problem."

The SC-COSMO project has pulled in Stanford faculty, researchers, graduate and medical students to work on the disease estimates. They are also working with the California prison system and public health officials in India and Mexico to help prevent the spread of COVID-19.

The SC-COSMO model also incorporates non-pharmaceutical interventions, such as recommendations on social distancing, and the timing and effects on reductions in contacts which may differ by demography.

When Governor Gavin Newsom unveiled CalCAT at a news conference in late June, he instructed all state agencies and departments to make COVID-19 data publicly accessible, provided it does not include information that would violate privacy.

"California is home to some of the world's most accomplished researchers, technologists, scientists, acclaimed universities, and leading technology companies," Newsom. "While these models and forecasts make different assumptions, all of them show that individual actions can dramatically change the trajectory of the virus."

The CalCAT tool includes:

- "Nowcasts," the rate at which COVID-19 is estimated to be spreading;
- Short-term forecasts, which show what various models predict will happen over the next few weeks in California;
- And scenarios showing what could happen over the next few months under various conditions.

Some 20 Stanford faculty and graduate, law and medical students are involved in the project.

"I love modeling infectious diseases because I get to focus on impactful research and work with great, interdisciplinary teams of researchers like the SC-COSMO team," said Anneke Claypool, a Ph.D. student in management science and engineering. "COVID-19 has affected everyone's daily life—and I'm glad to be helping the state of California fight this deadly virus."

The project provides the state with county-level COVID-19 estimates, including the number of infections and detected cases and projections of future needs for hospitals. They are also developing intuitive tools for those who are not themselves modeling experts.

Tess Ryckman, a Ph.D. student at Stanford Health Policy focused on decision science, said working on the team has taught her new skills and burnished her expertise. "I've been on an accelerated learning trajectory these past few months and have picked up a lot of coding and modeling skills that will be valuable for me not only in my current research but in future career," she said. "I also feel that I'm putting a lot of what I've learned during my Ph.D. training to good use by applying it to such a pressing and important issue."

How to Read CalCAT—Ingredients of a Model

For those who aren't modeling experts, reading the CalCAT tool can be challenging.

Goldhaber-Fiebert explains that the main measurement of the tool is the effective reproduction number—known as the R-effective—which is the average number of people onto which each infected person is likely to pass the virus. It also represents the rate at which COVID-19 is spreading.

He explains that if the R-effective is above 1, that means the infection is growing and would be one signal for concern. That might help the state focus on the counties that are of particular concern and hospitals prepare their ICUs and build up their PPE supply.

As of Tuesday, July 13, for example, the R-effective number was 1.09 for Los Angeles County, as compared to 1.05 in San Francisco County. The overall R-effective for California was 1.12.

The team gets its data from a bundle of data sources.

"The first is paying attention to the clinical and epidemiological literature that's being published and pre-published, and that's where Jason comes in." Jason Andrews, another principal investigator for the SC-COSMO modeling project, is an infectious disease physician and associate professor of medicine at Stanford Medicine.

They also get secure data feeds from the state. Then the team data analysts, SHP's Kim Babiarz and Lea Prince, do a granular analysis case, testing, hospitalization, and death series to create targets for model calibration and refine model inputs.

Finally, they track numerous public data sites, such as the U.S. Census Bureau for demographics, and Google Mobility and Foursquare, which track the movement of people after public health orders are put into effect. After six counties in the Bay Area issued its shelter-in-place order on March 16, they could see a large reduction in mobility.

"All of that data ultimately results in the ability of the model to make a range of future projections that are consistent with what's happened in the past—and then the uncertainty in these projections widens out as you move forward in time," Goldhaber-Fiebert said. "The tricky part, of course, is that there is always a delay in seeing effects in the data, especially when interventions and policies change," he added. "It's like when you try to turn a big ship—you turn the wheel, but it takes a while for the ship to turn."

Provided by Stanford University