Research team creates a computer model that can predict how COVID-19 spreads in cities
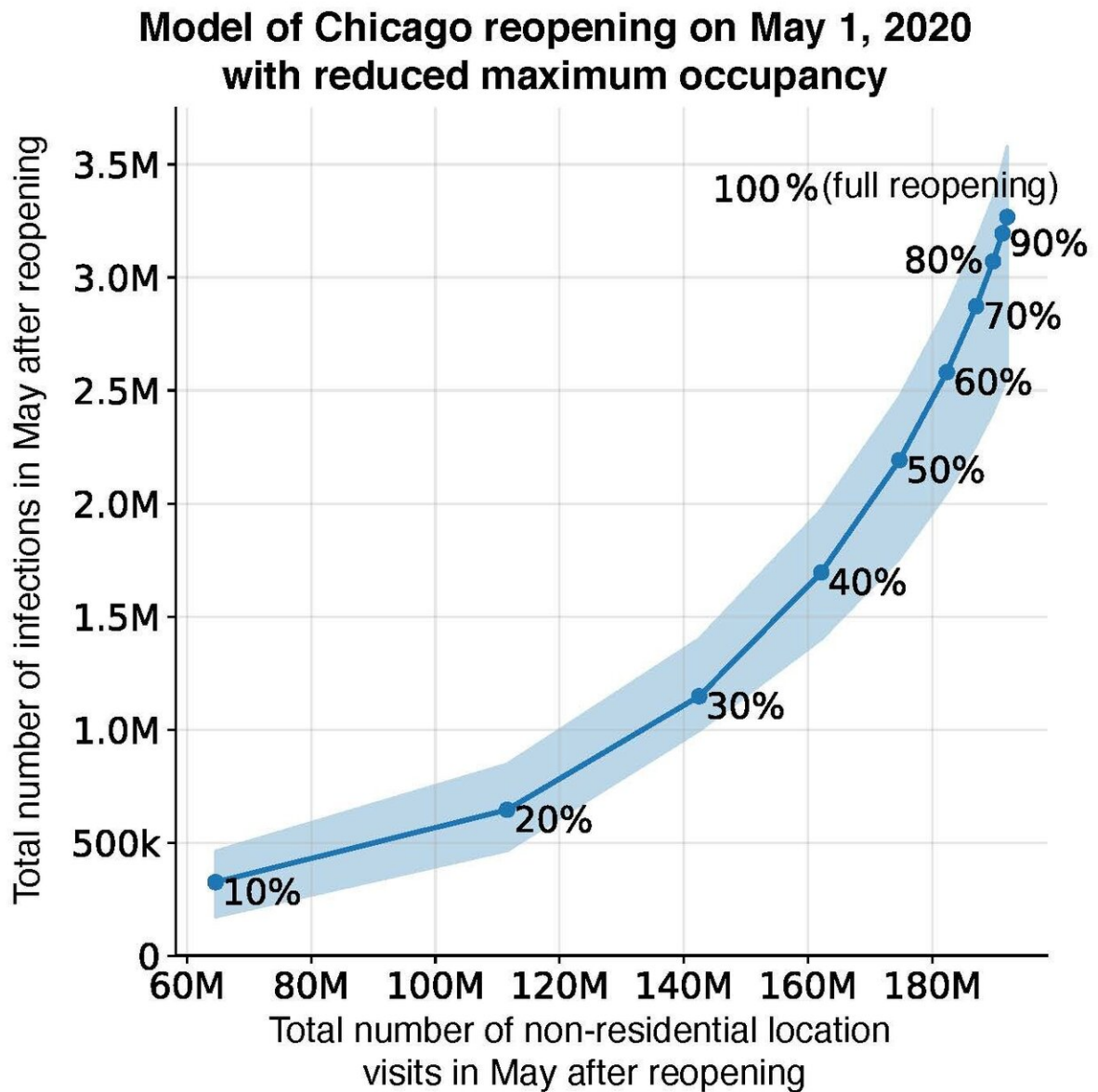
November 10 2020, by Tom Abate

Model of Chicago reopening on May 1, 2020 with reduced maximum occupancy

A new computer model predicts the COVID-19 infection-versus-activity trade-off for Chicago. According to the figure, COVID-19 infections will rise as the number of visits to businesses and public places approach pre-pandemic levels. However, restricting maximum occupancy can strike an effective balance: for example, a 20 percent occupancy cap would still permit 60 percent of pre-pandemic visits while risking only 18 percent of the infections that would occur if public places were to fully reopen. Credit: Serina Yongchen Chang

A team of researchers has created a computer model that accurately predicted the spread of COVID-19 in 10 major cities this spring by analyzing three factors that drive infection risk: where people go in the course of a day, how long they linger and how many other people are visiting the same place at the same time.

"We built a computer model to analyze how people of different demographic backgrounds, and from different neighborhoods, visit different types of places that are more or less crowded. Based on all of this, we could predict the likelihood of new infections occurring at any given place or time," said Jure Leskovec, the Stanford computer scientist who led the effort, which involved researchers from Northwestern University.

The study, published today in the journal *Nature*, merges demographic data, epidemiological estimates and anonymous cellphone location information, and appears to confirm that most COVID-19 transmissions occur at "superspreader" sites, like full-service restaurants, fitness centers and cafes, where people remain in close quarters for extended periods. The researchers say their model's specificity could serve as a tool for officials to help minimize the spread of COVID-19 as they reopen businesses by revealing the tradeoffs between new infections and lost sales if establishments open, say, at 20 percent or 50 percent of capacity.

Study co-author David Grusky, a professor of sociology at Stanford's School of Humanities and Sciences, said this predictive capability is particularly valuable because it provides useful new insights into the factors behind the disproportionate infection rates of minority and low-income people. "In the past, these disparities have been assumed to be driven by preexisting conditions and unequal access to health care, whereas our model suggests that mobility patterns also help drive these disproportionate risks," he said.

Grusky, who also directs the Stanford Center on Poverty and Inequality, said the model shows how reopening businesses with lower occupancy caps tend to benefit disadvantaged groups the most. "Because the places that employ minority and low-income people are often smaller and more crowded, occupancy caps on reopened stores can lower the risks they face," Grusky said. "We have a responsibility to build reopening plans that eliminate—or at least reduce—the disparities that current practices are creating."

Leskovec said the model "offers the strongest evidence yet" that stay-at-home policies enacted this spring reduced the number of trips outside the home and slowed the rate of new infections.

Following footsteps

The study traced the movements of 98 million Americans in 10 of the nation's largest metropolitan areas through half a million different establishments, from restaurants and fitness centers to pet stores and new car dealerships.

The team included Stanford Ph.D. students Serina Chang, Pang Wei Koh and Emma Pierson, who graduated this summer, and Northwestern University researchers Jaline Gerardin and Beth Redbird, who assembled study data for the 10 metropolitan areas. In population order, these cities

include: New York, Los Angeles, Chicago, Dallas, Washington, D.C., Houston, Atlanta, Miami, Philadelphia and San Francisco.

SafeGraph, a company that aggregates anonymized location data from mobile applications, provided the researchers data showing which of 553,000 public locations such as hardware stores and religious establishments people visited each day; for how long; and, crucially, what the square footage of each establishment was so that researchers could determine the hourly occupancy density.

The researchers analyzed data from March 8 to May 9 in two distinct phases. In phase one, they fed their model mobility data and designed their system to calculate a crucial epidemiological variable: the transmission rate of the virus under a variety of different circumstances in the 10 metropolitan areas. In real life, it is impossible to know in advance when and where an infectious and susceptible person come in contact to create a potential new infection. But in their model, the researchers developed and refined a series of equations to compute the probability of infectious events at different places and times. The equations were able to solve for the unknown variables because the researchers fed the computer one, important known fact: how many COVID-19 infections were reported to health officials in each city each day.

The researchers refined the model until it was able to determine the transmission rate of the virus in each city. The rate varied from city to city depending on factors ranging from how often people ventured out of the house to which types of locations they visited.

Once the researchers obtained transmission rates for the 10 metropolitan areas, they tested the model during phase two by asking it to multiply the rate for each city against their database of mobility patterns to predict new COVID-19 infections. The predictions tracked closely with the

actual reports from health officials, giving the researchers confidence in the model's reliability.

Predicting infections

By combining their model with demographic data available from a database of 57,000 census block groups—600 to 3,000-person neighborhoods—the researchers show how minority and low-income people leave home more often because their jobs require it, and shop at smaller, more crowded establishments than people with higher incomes, who can work-from-home, use home-delivery to avoid shopping and patronize roomier businesses when they do go out. For instance, the study revealed that it's roughly twice as risky for non-white populations to buy groceries compared to whites. "By merging mobility, demographic and epidemiological datasets, we were able to use our model to analyze the effectiveness and equity of different reopening policies," Chang said.

The team has made its tools and data publicly available so other researchers can replicate and build on the findings.

"In principle, anyone can use this model to understand the consequences of different stay-at-home and business closure policy decisions," said Leskovec, whose team is now working to develop the model into a user-friendly tool for policymakers and public health officials.

Jure Leskovec is an associate professor of computer science at Stanford Engineering, a member of Stanford Bio-X and the Wu Tsai Neurosciences Institute. David Grusky is Edward Ames Edmonds Professor in the School of Humanities and Sciences, and a senior fellow at the Stanford Institute for Economic Policy Research (SIEPR).

 More information: Mobility network models of COVID-19 explain

inequities and inform reopening, *Nature* (2020). DOI:
10.1038/s41586-020-2923-3 ,
www.nature.com/articles/s41586-020-2923-3

Provided by Stanford University