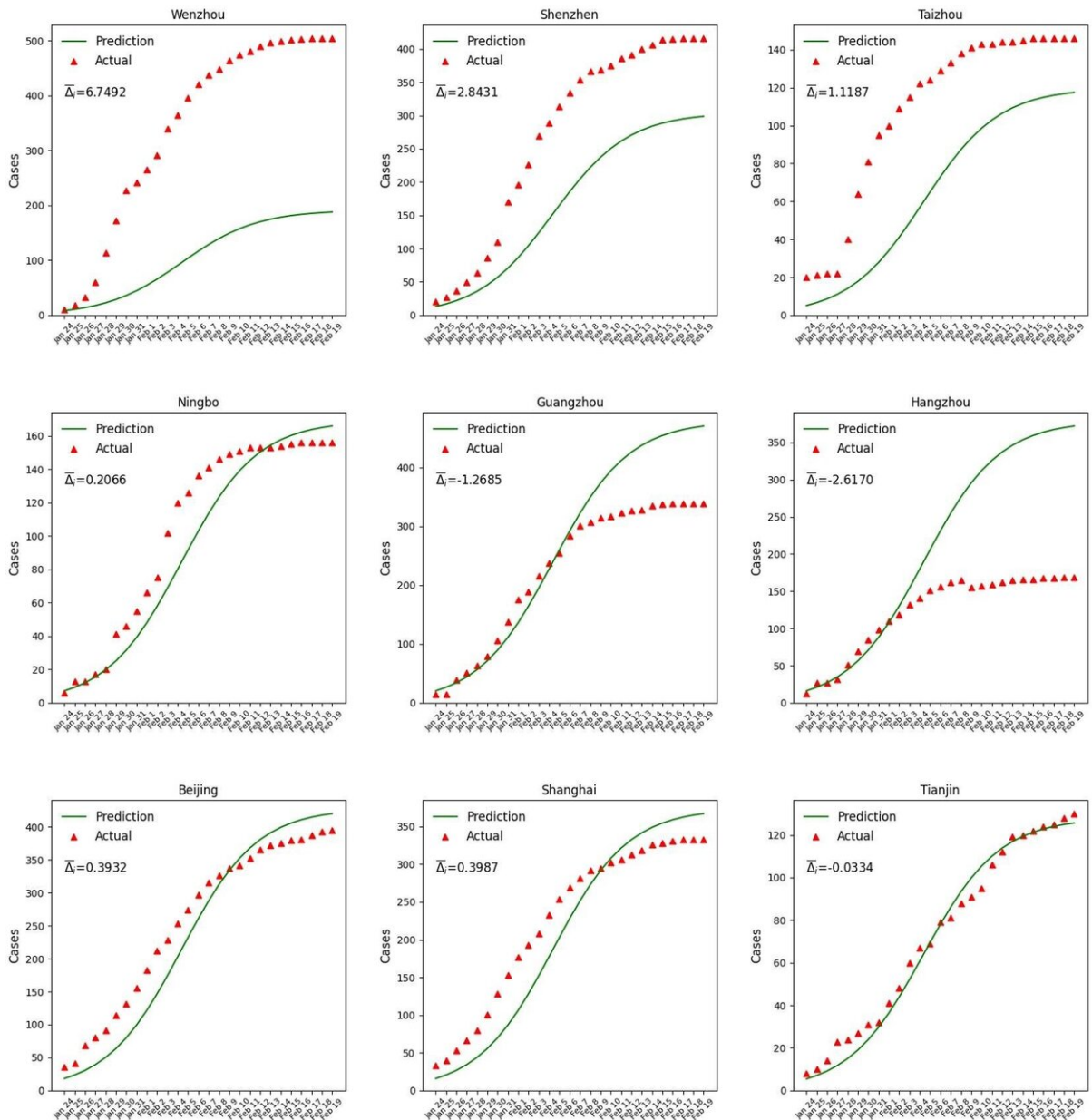# Researchers devise new model to track COVID-19's spread

April 29 2020, by Mike Cummings

Predicted versus observed epidemic curves over time. (copyright: Nature) Our model aggregates population outflow from Wuhan from January 1 to 24, 2020 to provide a reference growth pattern (i.e. epidemic curves) for COVID-19's spread. Differences in the predicted and confirmed growth in confirmed cases can signal higher levels of COVID-19 community transmission. Credit: University of Hong Kong

Yale University researchers and colleagues in Hong Kong and China have developed an approach for rapidly tracking population flows that could help policymakers worldwide more effectively assess risk of disease spread and allocate limited resources as they combat the COVID-19 pandemic.

The approach, described in a study published early online on April 29 in the journal *Nature*, differs from existing epidemiological models by exploiting real-time data about population flows, such as phone use data and other "big data" sources that can accurately quantify the movement of people.

"This work shows that it is possible to very accurately forecast the timing, intensity, and geographic distribution of the COVID-19 outbreak based on population movement alone," said Yale's Nicholas A. Christakis, Sterling Professor of Social and Natural Science and a co-author of the study. "Moreover, by tracking population flows in real time, our model can provide policymakers and epidemiologists a powerful tool to limit an epidemic's impact and save lives."

In developing the model, the researchers used nationwide mobile-phone geo-location data to track about 11.5 million occasions of people transiting through Wuhan, a prefecture city in China's Hubei Province,

between Jan. 1 and Jan. 24, 2020—a period covering the run-up to the Chinese Lunar New Year and the annual chunyun mass migration in China. People moved through Wuhan to 296 prefectures in 31 provinces and regions throughout the country. The researchers linked the population-flow data, which was provided by a major national wireless telecommunications carrier, to COVID-19 infection counts, provided by the Chinese Center for Disease Control and Prevention (Chinese CDC), by location and time at the prefecture level.

Their analysis demonstrates the effectiveness of the quarantine imposed on Wuhan on Jan. 23. By the end of the day on Jan. 24, movement out of the city had almost completely ceased, according to their findings.
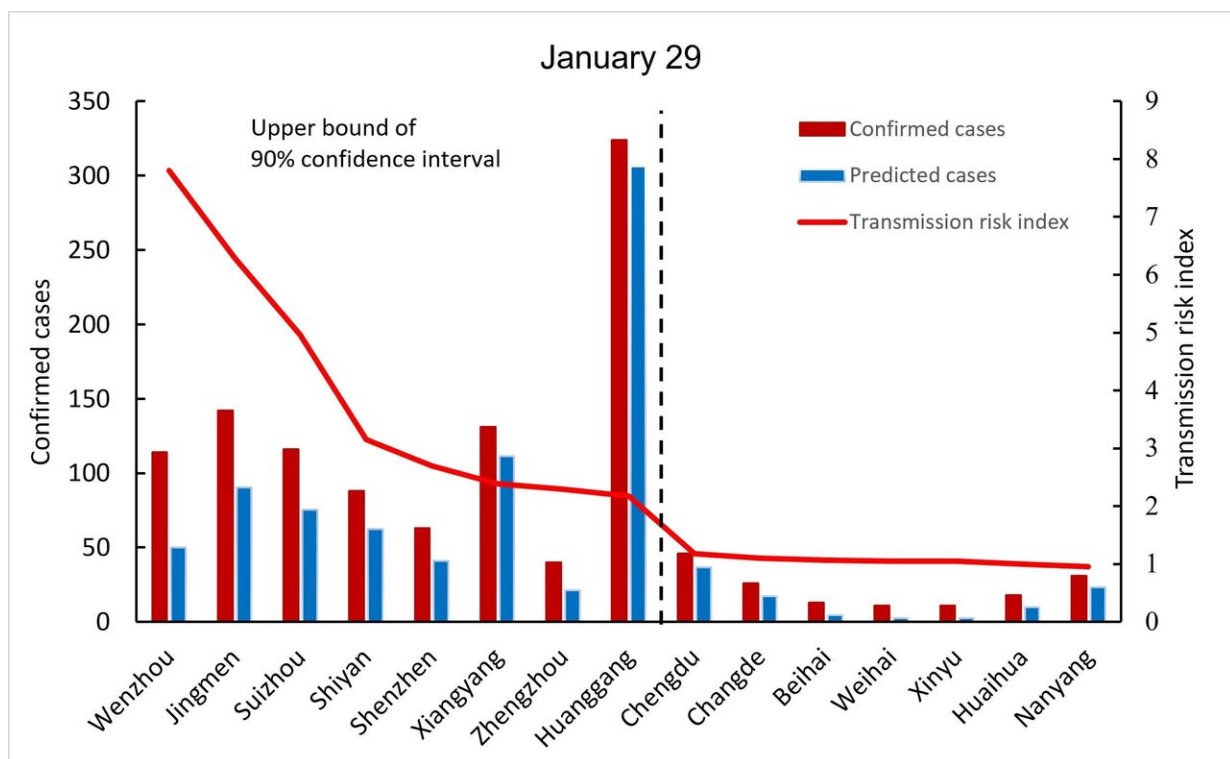
Illustrative example of using model to track COVID-19 community spread risk. (copyright: Nature) Our model uses population movement to predict expected cases. The predicted spread of the SARS-CoV-2 virus can be used as a

benchmark to identify which locales are 'outliers', which have significantly more or less cases than expected (given the movement data). The graph is an illustration of what our model showed on January 29. Prefectures to the left of the dashed line are outliers that have significantly more than expected cases, i.e., a higher level of unexplained or community transmission. Our model identified Wenzhou as having the most severe community transmission risk on January 29, 2020. The government announced a full quarantine of the prefecture on February 2, 2020. Credit: University of Hong Kong

The researchers found that the distribution of people leaving Wuhan accurately predicted the relative frequency of subsequent COVID-19 infections across China through Feb. 19, 2020. Researchers also developed a "risk source" model that leveraged population flow data to accurately forecast confirmed cases and identify places at risk of high transmission rates during the outbreak's early stages.

Their analysis also corroborates the data released by the Chinese CDC through Feb. 19 (for the prefectures outside of Wuhan itself) because it shows that a totally independent source of information—the telecom carrier—is very well correlated with official COVID-19 case counts.

The new model can be applied using any dataset that accurately captures people's movements, such as train ticketing or car tolling data, researchers noted, meaning that policymakers worldwide could use it to inform efforts to contain the virus' spread if data regarding population movements is available.

"People spread contagious diseases when they move," said Christakis, director of the Yale Institute for Network Science. "By accurately capturing population movements over time, we can predict how a contagion will spread geographically and use data-analytic techniques to help control it before a devastating epidemic erupts or re-erupts."

Provided by Yale University