MedicalXpress

Ups and downs in COVID-19 data may be caused by data reporting practices

July 14 2020

As data accumulates on COVID-19 cases and deaths, researchers have observed patterns of peaks and valleys that repeat on a near-weekly basis. But understanding what's driving those patterns has remained an open question.

A study published this week in *mSystems* reports that those oscillations arise from variations in testing practices and data reporting, rather than from societal practices around how people are infected or treated. The findings suggest that epidemiological models of infectious disease should take problems with diagnosis and reporting into account.

"The practice of acquiring data is as important at times as the data itself," said computational biologist Aviv Bergman, Ph.D., at the Albert Einstein College of Medicine in New York City, and microbiologist Arturo Casadevall, M.D., Ph.D., at the Johns Hopkins Bloomberg School of Public Health in Baltimore, Maryland. Bergman and Casadevall worked on the study with Yehonatan Sella, Ph.D., at Albert Einstein, and physician-scientist Peter Agre, Ph.D., at Johns Hopkins.

The study began when Agre, who co-won the 2003 Nobel Prize in Chemistry, noticed that precise weekly fluctuations in the data were clearly linked to the day of the week. "We became very suspicious," said Bergman.

The researchers collected the total number of daily tests, positive tests, and deaths in U.S. national data over 161 days, from January through the

end of June. They also collected New York City-specific data and Los Angeles-specific data from early March through late June. To better understand the oscillating patterns, they conducted a power spectrum analysis, which is a methodology for identifying different frequencies within a signal. (It's often used in signal and image processing, but the authors believe the new work represents the first application to epidemiological data.)

The analysis pointed to a 7-day cycle in the rise and fall of national new cases, and 6.8-day and 6.9-day cycles in New York City and Los Angeles, respectively. Those oscillations are reflected in analyses that have found, for example, that the mortality rate is higher at the end of the week or on the weekend.

Alarmed by the consistency of the signal, the researchers looked for an explanation. They reported that an increase in social gatherings on the weekends was likely not a factor, since the time from exposure to the coronavirus to showing symptoms can range from 4-14 days. Previous analyses have also suggested that patients receive lower-quality care later in the week, but the new analysis didn't support that hypothesis.

The researchers then examined reporting practices. Some areas, like New York City and Los Angeles, report deaths according to when the individual died. But national data publishes deaths according to when the death was reported—not when it occurred. In large datasets that report the date of death, rather than the date of the report, the apparent oscillations vanish. Similar discrepancies in case reporting explained the oscillations found in new case data.

The authors of the new study note that weekend interactions or health care quality may influence outcomes, but these societal factors do not significantly contribute to the repeated patterns.

"These oscillations are a harbinger of problems in the public health response," said Casadevall.

The researchers emphasized that no connection exists between the number of tests and the number of cases, and that unless data reporting practices change, the oscillations will remain. "And as long as there are infected people, these oscillations, due to fluctuations in the number of tests administered and reporting, will always be observed," said Bergman, "even if the number of cases drops."

Provided by American Society for Microbiology