

COVID-19: Tracking, data privacy and getting the numbers right

May 18 2020



Credit: CC0 Public Domain

As plans for re-opening businesses, communities and schools emerge, it becomes increasingly important to better understand how many people are being infected and dying from COVID-19, and where and how the

new coronavirus is transmitted.

In Wednesday's Berkeley Conversations: COVID-19 event led by Nobel Laureate Saul Perlmutter, director of the Berkeley Institute for Data Science and professor of physics, three Berkeley faculty opened different windows onto what they're discovering about how to track and limit the spread of COVID-19. They also discussed what data they need to learn more and how we can use techniques like [data encryption](#) to advance our understanding while protecting private information.

Presenters included Jacob Steinhardt, assistant professor of statistics, Uros Seljak, professor of physics and senior fellow with the Berkeley Institute for Data Science (BIDS) and Shafi Goldwasser, professor of electrical engineering and computer sciences and director of the Simons Institute for the Theory of Computing.

Steinhardt has assessed the relative benefits of data sources available to track infection rates, from smart thermometers to sampling wastewater for genetic traces of the [coronavirus](#).

Ideally, he said, "we want short time lag, low error" data sources that provide the most accurate picture in the shortest time after exposure. Some of his most recent research examines what measures are necessary to increase mobility levels more safely. In the live, online talk, Steinhardt also discussed how the sources of transmission have shifted under shelter-in-place measures.

Seljak, meanwhile, has examined death rates connected with the SARS-COV-2 virus. He recently published a paper based on data from Italy that suggested that death rates are far higher than many initial estimates, particularly for older people. He and his co-authors compared [death rates](#) by age in the hard-hit region of Lombardy in 2020 with those over the previous five years.

While the earlier five years showed little variation, 2020 saw a substantial increase, beginning with the outbreak of COVID-19. The researchers hypothesized that many of the additional deaths were among older people who had died from SARS-COV-2 virus infections outside of hospital settings but had not been tested. They found similar results in other communities.

"What is the risk of dying if you get infected? The answer to that is that it depends very strongly on age," he said. "If you are, for example, in the age range of 30 to 39, then your risk of dying (if infected) is one in 10,000, so that's a very small number, and it goes up from there ... By the time we get to 70 to 79, it's one in 40 or 2.5%; and 80 to 89, it's one in 15, or 6.6%; and then finally, if you are 90 or above, it's one in six, or roughly 17%."

Steinhardt and Seljak noted that while the currently available data has provided critical insights that are helping to inform decisions about hospital capacity and re-opening business and communities, more robust and accurate data are still urgently needed.

Details about not just who is infected, but where they work and who they see can build understanding about how the disease spreads and whom it affects. Some of that data has, in fact, already been collected and is held in private databases that belong to hospitals and medical centers that are prevented by law from sharing it.

Shafi Goldwasser is among the researchers pioneering ways to use this powerful arsenal of data without actually "seeing" it. In other words, she and her team are developing tools and approaches to aggregate and compute on huge volumes of encrypted data and enable insights without violating privacy.

In the conversation, Goldwasser explained how this process, called

homomorphic encryption, works.

Such an approach could be used if people agreed to share encrypted data from their phones to enable contact tracing, she said. With that data, it would be possible to see trends related to where and how people are becoming infected.

"And you won't know which household, who was infected, who was close to (someone) who was infected," she said. "The kind of computations we're talking about are not complicated and can be done efficiently under encryption."

Perlmutter concluded: "These are really great examples of what a public research university does best. And their work is just what we need at this time."

Provided by University of California - Berkeley

Citation: COVID-19: Tracking, data privacy and getting the numbers right (2020, May 18)
retrieved 19 April 2024 from
<https://medicalxpress.com/news/2020-05-covid-tracking-privacy.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--