

National Academies publishes guide to help public officials make sense of COVID-19 data

July 30 2020, by James Urton



Credit: CC0 Public Domain

As the COVID-19 pandemic continues, officials across the country have had to make decisions about opening and closing schools, businesses and community facilities. They have relied in large part on information about the pandemic—from hospitalization statistics to test results—to inform these decisions. But different facts and figures about COVID-19 can paint different pictures of the pandemic, according to Adrian Raftery, a

professor of statistics and sociology at the University of Washington.

"The COVID-19 pandemic is generating many different types of data about this disease in communities—things like the number of confirmed cases or the number of deaths in a particular area," said Raftery. "None of these data sources on their own are perfect in terms of capturing a complete and accurate summary of the prevalence of COVID-19 and the risks of doing certain things like opening businesses or schools. All have their own strengths and weaknesses."

Raftery is lead author of a new [guide](#) published June 11 by the National Academies of Sciences, Engineering and Medicine that is intended to help officials nationwide make sense of these different COVID-19 data sources when making public health decisions.

Officials looking for COVID-19 statistics have plenty to choose from: confirmed cases, deaths, hospitalizations, intensive care unit occupancy, emergency room visits, antibody tests, nasal-swab tests and the ratio of positive test results—to name a few of the more common data points collected and distributed by hospitals and public health agencies. But officials don't necessarily have all of these statistics on hand when making decisions, or have enough information to interpret them.

"We intend for this guide to help these decision-makers and their advisors interpret the data on COVID-19 and understand the upsides and downsides of each data source," said Raftery.

For example, the number of positive test results for the novel coronavirus is likely an underestimate of its true prevalence in a community. Many people who have the virus are asymptomatic and aren't likely to seek out a test, and even people with symptoms may not have access to tests and medical care, according to Raftery. As another example, the number of COVID-19 deaths in a region does not reflect

the disease's current prevalence because the number of deaths lag behind the number of cases by several weeks. In addition, some deaths may be misattributed to COVID-19, Raftery said.

The guide highlights some criteria for officials to take into account when assessing the usefulness of particular COVID-19 data points, including:

- Assessing how representative the data are for a community or region
- Whether there may be systemic biases in some data sources
- Thinking about the types of uncertainties in data sources, due to factors like sample size, how data were collected and the population surveyed
- Whether there's a time lag due to delays in reporting data, the course of the disease and other factors

"There are no perfect data sources, but all of these data sources are still useful for making decisions that directly impact public health," said Raftery.

Raftery has worked extensively on statistical methods to measure and estimate the prevalence of other viruses, including HIV in Africa. Though HIV and the novel coronavirus cause different types of diseases, there are similarities in how the two viruses spread among susceptible populations, as well as how types of social distancing—condom use for HIV and physical distancing and mask usage for the novel [coronavirus](#)—can decrease transmission. COVID-19 is also generating the same types of data sources, with the same limitations, as HIV/AIDS, such as test results, hospitalization rates and deaths.

Over time, it may be possible to collect more revealing data about COVID-19 from what are known as "representative random samples" within a population. In representative sampling, people are surveyed at

random for a disease, and certain populations can be more heavily sampled than others based on what scientists and officials have learned about a disease's prevalence and susceptibility. Representative sampling avoids biases and can more accurately estimate the disease's prevalence in a region, according to Raftery.

"As we learn more about COVID-19, how it spreads, how different populations are more or less susceptible, we may be able to move more in the direction of representative sampling," said Raftery. "The State of Indiana has already done a survey of this kind, and others should follow suit. But there is also a lot that officials can do with the statistics and data sources that hospitals and agencies are providing right now—provided that officials can be made aware of the strengths and weaknesses of each piece of data."

The guide is the first completed by the National Academies' Societal Experts Action Network—or SEAN—an eight-member committee tasked by the National Academies to provide rapid expert assistance on issues related to the social and behavioral sciences during the pandemic. Raftery is a member of the SEAN and spearheaded this inaugural project.

More information: Evaluating Data Types, (2020). [DOI: 10.17226/25826](https://doi.org/10.17226/25826)

Provided by University of Washington

Citation: National Academies publishes guide to help public officials make sense of COVID-19 data (2020, July 30) retrieved 16 June 2024 from <https://medicalxpress.com/news/2020-07-national-academies-publishes-covid-.html>

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.