World view

By Natalie Dean

Tracking COVID-19 infections: time for change

To manage the pandemic effectively, channel the power of random sampling.

One of the best ways the world has to get a clear view of COVID-19 is going underused. It's time to exploit the power of random sampling.

Last September, the US Centers for Disease Control and Prevention estimated that only one in four SARS-CoV-2 infections in the United States had been reported. Across Africa, the average is closer to one in seven. Why? Many people who are quite ill, or worried about their symptoms, can't get tested. Those with mild or no symptoms often don't seek testing.

And undercounts are getting worse. Reinfections and breakthrough infections are rising, but they are often mild, so people go untested. The onslaught of Omicron cases has far outstripped many countries' testing capacities. Last December, a testing site near me in Atlanta, Georgia, had a wait of three to four hours. In the United States, at-home lateral flow tests are finally becoming more readily available, so fewer people will seek PCR confirmation.

All this undercounting renders many important questions unanswerable. For example, if a surge in cases slows, is transmission down, or is testing maxed out? Waiting to find out means that hospitals can't prepare and policymakers are two to four weeks behind. Who can drive looking only in their rear-view mirror?

Wastewater surveillance is an innovative part of the solution. It shows whether virus levels are increasing or decreasing across a community, and does not depend on people seeking or reporting test results. In my home state of Massachusetts, waste water was one of the earliest reliable indicators that infections were declining last month.

But waste water can't pinpoint who in a community is getting infected and who is getting sick. With Omicron, hospitalizations in children have reached record highs. Yet infections in this age group are frequently missed. It's clear there are more infections, but are those infections more severe? Knowing that is important for risk–benefit calculations around schooling, vaccinations and much more.

Random sampling can answer those sorts of question. As long as participants are selected randomly, they will on average mimic characteristics of the wider population. Roughly speaking, testing fewer than 1,000 people can yield crucial information about 10 million, or even more.

Shining examples of random sampling are the Coronavirus (COVID-19) Infection Survey run throughout the United Kingdom by the Office of National Statistics (ONS), and Imperial College London's REACT-1 study. The ONS initiative aims to obtain swab test results at least fortnightly from around 180,000 people across the United Kingdom, and blood tests monthly from around 150,000 people. In late January, one in 20 people tested positive for current infection. But age really mattered: one in 10 of the youngest children tested positive, as did one in 15 of the older children. The results signalled an enormous pool of infections, and were quickly made available to guide policy and family decisions.

> "Without random sampling, there's a vicious cycle of guesswork."

Forecasting the course of the pandemic demands reliable estimates of current infection levels. Without accurate knowledge of these levels, epidemiologists must make many assumptions (on the likelihood that, for example, infected people will develop symptoms, or be tested). That guesswork informs mathematical models and, consequently, public discussions about the trajectory of the pandemic. Models that overestimate how many infections have been missed overestimate population immunity, and can underestimate the risk of resurgence. Those estimates are used for decisions about everything from opening schools to planning policies and targeting vaccination campaigns. Without random sampling, there's a vicious cycle of guesswork.

The UK data are informative elsewhere, but generalizing too much from one country's data is perilous. In the United States, a few random-sampling surveys have been conducted by health departments and academic partners, for example, in Indiana, Georgia and California. These have bolstered local understanding of disparities across racial and ethnic groups. At a national level, researchers at Emory University in Atlanta (where I also work), carried out a representative household survey (P. S. Sullivan *et al. Clin. Infect. Dis.* https://doi.org/hfvm; 2021). A new round of antibody and nasal-swab testing is conducted every four to nine months. But a situation that's evolving quickly requires more frequent samples.

Why isn't random-sampling for infection happening more widely? These studies require sustained resources and coordinated effort. The patchwork US public-health system makes collaboration across states challenging. The studies also require a public that's willing and able to participate. Low participation rates in surveys are a major challenge. As an incentive to take part in the ONS survey, the UK government has offered more than £200 million (US$270 million) of shopping vouchers.

More than two years into the COVID-19 pandemic, it is clear that the virus SARS-CoV-2 will be circulating for a long time to come. Millions of people are being infected daily, and the threat of new variants looms. Investing in random sampling can better prepare governments for the future. A single sampling framework can be used for multiple pathogens, such as influenza and other respiratory viruses. For infectious diseases, failing to see the whole picture will mean poor decisions. Yes, random sampling will cost, but bad information is expensive, too.

Natalie Dean is a biostatistician at Emory University in Atlanta, Georgia.
e-mail: nataliedean@emory.edu

UNIV. FLORIDA