

# News in focus



ROBERT NICKELSBURG/GETTY

A health-care worker checks data on a computer before treating a person with COVID-19 in New York City.

## WHY THE UNITED STATES IS HAVING A COVID-19 DATA CRISIS

Political meddling and years of neglect for public-health data management are to blame.

By Amy Maxmen

**S**outh Korea's grip on the coronavirus faltered this month when a large church in Seoul had an outbreak – involving 915 cases as of 25 August. The government has reinstated restrictions in the city to prevent a surge, but it's also reporting details of the outbreak publicly. For instance, it has shared that 120 people infected at the church have spread the coronavirus to people at 22 venues, including 4 call centres and 3 hospitals in Seoul.

Almost every day for the past seven months, the Korea Centers for Disease Control and

Prevention has updated its website with near-real-time information on local outbreaks. The site also reports several COVID-19 statistics for every region of the country.

Data dashboards in Singapore and New Zealand offer similar windows into how the coronavirus is spreading within their borders. This helps policymakers and citizens determine how to go about daily life, while reducing risks – and provides researchers with a wealth of data. By contrast, the United States offers vanishingly few details on how the disease is spreading, even as people increasingly socialize and travel, and authorities reopen schools and businesses. This state of affairs is

frustrating data researchers, who want to help authorities make decisions that can save lives.

“We shouldn't be flying blind at this point,” says Natalie Dean, a biostatistician at the University of Florida in Gainesville. “We shouldn't have to speculate.”

Experts told *Nature* that political meddling, privacy concerns and years of neglect of public-health surveillance systems are among the reasons for the dearth of information in the United States.

Although information isn't the only tool that can be used against a pandemic, South Korea's attention to data correlates with its overall success at controlling the outbreak:

## News in focus

the country has had about 3.5 cases per 10,000 people overall, and there have been around 2 COVID-19 deaths per week over the past month. By contrast, the United States has had 175 cases per 10,000 people overall, and about 7,000 people have died of the disease every week for the past month.

South Korea owes its detailed intel to a coordinated network of public-health centres in 250 districts that send information rapidly to the Korea Centers for Disease Control and Prevention. Sung-il Cho, an epidemiologist at Seoul National University, attributes the system's success to this centralization of power – along with rapid hiring of 'temporary epidemiologists' to meet demand during the COVID-19 pandemic. These scientists have helped to lead contact-tracing investigations that produce succinct, anonymized details such as those listed about the outbreak at SarangJeil Church in Seoul.

The United States isn't pursuing contact tracing for COVID-19 to nearly the same extent as South Korea, but its disease surveillance also flows from local health departments to the federal level. For years, the US Centers for Disease Control and Prevention (CDC) has used this system to follow the spread of outbreaks such as a current surge in infections with the bacterium *Salmonella*, and to trace them back to their sources. However, the surveillance system has run into problems at several levels during the pandemic. The result is that a lot of data, such as information on where people were exposed to the coronavirus, are missing. And the data that do exist are made public only slowly.

The CDC and four US health departments declined to speak to *Nature* about how they're managing COVID-19 data. However, former scientific employees and researchers who collaborate with them offered suggestions about why data are lagging and lacking in the United States.

### Increased scrutiny

Some speculate that because the pandemic is politically charged, data describing the situation are guarded closely by officials in the administration of President Donald Trump. Researchers say that investigations published in the CDC's journal *Morbidity and Mortality Weekly Reports* have been thorough, but are posted online long after they might influence outcomes. For example, on 31 July, the CDC reported that 260 staff members and children at an overnight camp in Georgia had been infected more than a month earlier (C. M. Szablewski *et al.* *Morb. Mortal. Weekly Rep.* **69**, 1023–1025; 2020). Samuel Groseclose, a public-health specialist who retired from the CDC in 2018, suggests that the reports are undergoing an unusual amount of review within the agency, and perhaps within its parent agency, the US Department for Health and Human Services (HHS).

The CDC was further sidelined in July, scientists say, when the Trump administration announced that data on COVID-19 cases and hospitalizations would be diverted away from the agency, and handled instead by a new system launched at the HHS, the head of which reports directly to the president. So far, the HHS's dashboard has been up to a week behind in reporting data, and it includes information only on case counts and hospital capacity, rather than details such as the location of disease clusters. An HHS spokesperson says that the new system streamlines reporting from 6,000 hospitals in the country.

But Georges Benjamin, the executive director of the American Public Health Association (APHA) in Washington DC, says that rather than streamlining data collection, the change has confounded information gathering further. Some hospital administrators are now confused about which agencies to report to, he says. He's also frustrated because the US\$10 million spent on the HHS's system might have been put to better use improving

**“We've begged for money over the years to build a solid information highway.”**

public-health data management at the CDC and at the health departments it collaborates with across the country. This outdated system has buckled under the pressure of 5.7 million COVID-19 cases in the United States.

Benjamin notes that many health departments still share data by fax, which is more time-consuming than digital methods. A lack of funds also means that overburdened staff don't have enough time to analyse the data they have. The APHA and other scientific organizations have long called for resources to improve data surveillance in the US public-health system.

“We've begged for money over the years to build a solid information highway so that we can collect data rapidly and share it with the people that need it in a timely way,” says Benjamin. “But we've never gotten what we needed.”

This long-standing neglect has been exacerbated by the lack of national leadership during the pandemic, say the researchers who spoke to *Nature*. There are no nationwide requirements for the information that hospitals and testing laboratories must report to health departments. Ranu Dhillon, an epidemic-response specialist at Harvard Medical School in Boston, Massachusetts, who is currently treating people with COVID-19 in Vallejo, California, says that neither the local health department nor the CDC tells him to report where patients might have been exposed to the coronavirus. He records these data in their

health records voluntarily, but he's not sure whether the state or local health department use the information, which troubles him. He worked on the Ebola response in Guinea during the outbreak in 2015, and says authorities there collected data on how people were infected with the Ebola virus, to curb its spread. “This crucial information would help us drill down transmission further” for COVID-19, he says. “It's crazy we aren't collecting it.”

### Access denied

Epidemiologists at universities would like to help overburdened state and local health departments to analyse whatever data they have so they can help officials to direct an efficient coronavirus response. Typically, health departments share disease-surveillance data with researchers on request. But during the COVID-19 pandemic, epidemiologists have been denied access. For instance, Steffanie Strathdee, an HIV epidemiologist at the University of California, San Diego, often requests case data broken down by parameters such as locality, race and probable exposure route, such as intravenous drug use. “Surveillance is done and it's made transparent, and this is what we have been accustomed to,” she says. “Only for this epidemic, things have changed.”

This year, Strathdee and other epidemiologists asked for anonymized COVID-19 data from the California department of health. E-mails seen by *Nature* show that the researchers' requests considered the need for individual privacy with measures such as asking the state for age ranges instead of exact ages. But the request was denied – something that had never happened to Strathdee before. The department's director, Mark Ghaly, explained in a 3 July e-mail that releasing information from people's records would “necessitate a careful and time-consuming analysis of each record to determine what data fields may be made public from each individual report”.

Without up-to-date, reliable information on who is infected, why and where, US scientists, policymakers and the public must instead rely on media reports and independent efforts to consolidate data, such as the COVID Tracking Project from magazine *The Atlantic* and the COVID-19 Dashboard compiled by researchers at Johns Hopkins University in Baltimore, Maryland.

But data from news outlets aren't necessarily comprehensive or universally trusted, and the dashboard doesn't detail where transmission is happening.

There is an urgent need for such information, says Caitlin Rivers, an epidemiologist at Johns Hopkins, because people are returning to work, to socializing and to schools. That means that precisely tailored interventions are more important than ever. “It's not enough to just tell people to be cautious.”