It shouldn’t be so hard to get the data needed to protect public health

By Dr. Christine Cassel, Susan L. Graham and William H. Press

Large companies and organizations exist in an era of evidence-based decisionmaking, fueled by digital data and analytics. Yet the U.S. public health system lacks the data needed to manage the current pandemic.

Modern data science, were it put to use, could both serve public health needs and also make our healthcare delivery system more efficient. Real-time information about who is harboring disease, who has been exposed to infection, and where clusters of cases occur would enable effective contact tracing and isolation strategies. In this pandemic, we could have avoided closing down all businesses and all schools by targeting interventions to where the risk of illness was high, not keeping every restaurant and every school shuttered and throwing the country into a recession.

The public-health data system we should have had in place was described 10 years ago in reports by the President’s Council of Advisors on Science and Technology, or PCAST, during the Obama administration and by independent advisers such as the Jason Study Group. That system would have used a modern, cloud-based approach with the kind of secure, private data flows already used for financial records and consumer transactions.

The backbone of such a public-health data system is already in place. The vast majority of U.S. healthcare activity is already recorded electronically in electronic health records. Yet although billions of dollars have been spent on EHRs for the healthcare delivery system—hospitals, clinics and emergency departments—almost nothing has been invested so that public health can unlock that same data.

It would not be a big additional step for the Centers for Disease Control and Prevention and other public health authorities to collect the information needed. It is a scandal that the best reporting now comes not from the government but from reports by universities and news organizations, produced by agglomerating incomplete reports from state and local entities.

**Why isn’t public health** information managed in 2020 at least as well as other large data assets? We see two reasons: Public health technology infrastructure has been tragically underfunded. And intentional design decisions by private-sector EHR vendors inhibit using the data for tracking infectious diseases like COVID-19. Both these issues can be addressed by Congress and the administration through a few key steps. A group of independent scientists, former PCAST members including ourselves, have spelled these out in a series of reports available at opcast.org. The group’s three most important recommendations for unlocking existing data for public health are:

- Interoperability requirements for EHRs must be accelerated to share all patient information with every provider caring for the same patient, with patients themselves, and also to share with public health organizations.
- Some COVID-19 recovery money should be used to build the digital expertise and infrastructure at CDC and at state public health offices to allow them seamless communication and coordination; $500 million allocated in the CARES Act could be used for this purpose.
- Effective shared governance between states and the CDC could support the states’ and territories’ responsibility in their jurisdictions, while also strengthening the CDC’s national leadership and coordination of tracking, contact tracing, isolation policies and public communication.

The rest of the U.S. economy benefits from modern digital infrastructures that are missing in our healthcare system. The COVID-19 crisis is a wake-up call for the nation to fix this shortcoming. Eventually, this pandemic will be over. On that bright day, we need to wake up to a better and more seamlessly integrated healthcare and public health system so we’re ready for the next health crisis. If we start now, we can make that happen.