World view

By Adam Kucharski

Fragmented outbreak data will lead to new pandemics

To break the vicious cycle of patchy understanding and poor virus control, we need to talk about privacy.

When the next pandemic emerges, what will the first few weeks look like? Will there be systems to identify contacts of infected individuals and to combine information to get crucial insights about the incubation period, disease severity and effective control measures? Or will fragmented data and improvised policies again lead to large epidemics and lengthy restrictions?

Every outbreak has a cycle of feedback between understanding and control. To design the most effective, least disruptive controls, governments must understand where and how disease transmission happens. Measures for controlling transmission — such as testing and tracing of contacts — can in turn provide more insights. Better data mean better control measures, but countries must have tough discussions about what getting those insights will take.

In early 2020, health officials weren't confident about where SARS-CoV-2 infections were, or what measures would suppress transmission. Many cities and countries went into lockdown as a result.

At the pandemic's start, my colleagues and I patched together biased and incomplete global data sets. We combined cases reported in different countries with infections on repatriation flights to estimate the impact of the lockdown in Wuhan, China, where the virus emerged. We studied the outbreak on the *Diamond Princess* cruise ship to extrapolate the severity of the infection. When we struggled with uncertainty, it was often because the data sets had not been collected to answer our questions.

In March 2020, my colleagues at the London School of Hygiene & Tropical Medicine set up CoMix, which surveyed social contacts in the United Kingdom. The first solid evidence that social-distancing behaviour had changed enough to suppress COVID-19 didn't come from cases or hospitalizations; it came from estimates in CoMix. When the Alpha variant emerged later in 2020, social-contact studies helped to show that surging infections were probably a result of the new variant, not more relaxed local behaviour (N. G. Davies *et al. Science* **372**, eabg3055; 2021).

In some places, behaviour has been analysed at finer scales. In South Korea, mobile-phone and credit-card data linked individuals to COVID-19 hotspots: 57,000 people who'd been near a nightclub outbreak received text messages telling them to get tested. In Taiwan, mobile-phone tracking ensured contacts of infected people stayed in quarantine. During test events, Singapore recorded millions of social interactions using Bluetooth sensors.

> **"Midway through a pandemic is not the time to debate how to balance data and privacy."**

Adam Kucharski is co-director of the Centre for Epidemic Preparedness and Response at the London School of Hygiene & Tropical Medicine.
e-mail: adam.kucharski@lshtm.ac.uk

The author declares competing interests; see go.nature.com/3dgko for details.

Collecting such data is important for evaluating control measures. Fresh initiatives will hopefully ensure that new diagnostics, treatments and vaccines arrive even sooner in the next pandemic. But countries will nonetheless first have to decide on measures such as isolation, quarantine, mask mandates and limits on social contact.

To understand the effectiveness of these and similar measures, researchers have aligned the timing of interventions with epidemic curves, but there have been very few studies designed specifically for this purpose. Most that have occurred, such as the UK Events Research Programme, which looked at COVID-19 risk at in-person events in 2021, have been underpowered and inconclusive.

Study design could become particularly important in the next pandemic. In 2020, COVID-19 vaccine results arrived quickly because resurgent second waves occurred in active trial sites, so evidence accrued quickly. Such evidence will arrive more slowly in partially suppressed epidemics.

Since 2016, several researchers (including my team of collaborators) contributed to the World Health Organization's R&D Blueprint, in a project to ready vaccine trial designs for health emergencies. In 2015, a 'ring trial' of Ebola vaccines successfully tracked infections among infected people's contacts and contacts-of-contacts. However, a respiratory infection that spreads faster among less clearly defined contacts poses a greater challenge, unless these contacts can somehow be rapidly identified and enrolled.

Quantifying the characteristics of the virus, identifying people at risk of being infected and evaluating the effectiveness of early measures, as well as vaccines, all require good data. And that requires planning how future data systems can fulfil multiple roles. (It will not be possible to build new infrastructure for every question.) There are examples of proactive data streams to learn from. After the 2003 severe acute respiratory syndrome outbreak, Taiwan established the National Health Command Center, which rapidly mobilized against COVID-19 in early 2020 using data-driven measures ranging from digital quarantine to triangulation of patient travel and contact history. After a 2015 Middle Eastern respiratory syndrome outbreak, South Korea amended legislation to allow health agencies to access and analyse data, including mobile-phone location and card transactions, during a serious outbreak.

I've noticed a conflict in the West when it comes to implementing data-intensive approaches. I've lost track of how many times someone has said we should copy East Asia's responses — but once they hear the details, they conclude these measures are an unacceptable invasion of privacy.

Midway through a pandemic is not the time to debate how to balance data and privacy, or which control measures and trial designs are appropriate. These are decisions that countries need to plan for now, before the next pandemic.

BRET HARTMAN/TED