

ZEYNEP TUFEKCI

Show Me the Data!

Aug. 27, 2021, 5:00 a.m. ET



By Zeynep Tufekci
Opinion Columnist

Who should get vaccine booster shots and when? Can vaccinated people with a breakthrough infection transmit the virus as easily as unvaccinated people? How many people with breakthrough infections die or get seriously ill, broken down by age and underlying health conditions?

Confused? It's not you. It's the fog of pandemic, in which inadequate data hinders a clear understanding of how to fight a stealthy enemy.

To overcome the fog of war, the Prussian general and military theorist Carl von Clausewitz called for "a sensitive and discriminating judgment" as well as "skilled intelligence to scent out the truth." He knew that since decisions will have to be made with whatever information is available in the face of an immediate threat, it's crucial to acquire as much systematic evidence as possible, as soon as possible.

In the current crisis, that has often been difficult.

These days, some experts grapple for answers on Twitter. They might be trying to figure out the effect of a vaccine booster shot by reverse engineering a bar chart in a screenshot from Israel's Ministry of Health, or arguing with one another about confounding factors or statistical paradoxes.

Why this stumbling in the fog? It may seem like we're drowning in data: Dashboards and charts are everywhere. However, not all data is equal in its power to illuminate, and worse, sometimes it can even be misleading.

Few things have been as lacking in clarity as the risks for children. Testing in schools is haphazard, follow-up reporting is poor and data on hospitalization of children appears to be unreliable, even if those cases are rare. The Food and Drug Administration has asked that vaccine trials for children aged 5 to 11 be expanded, which is wise, but why weren't they bigger to begin with?

While the pandemic has produced many fine examples of research and meticulous data collection, we are still lacking in detailed and systematic data on cases, contact-tracing, breakthrough infections and vaccine efficacy over time, as well as randomized trials of interventions like boosters. This has left us playing catch-up with emerging threats like the Delta variant and has left policymakers struggling to make timely decisions in a manner that inspires confidence.

To see the dangers of insufficient data and the powers of appropriate data, consider the case of dexamethasone, an inexpensive generic corticosteroid drug.

In the early days of the pandemic, doctors were warned against using it to treat Covid patients. The limited literature from SARS and MERS — illnesses related to Covid — suggested that steroids, which suppress the immune system, would harm rather than help Covid patients.

That assessment changed on June 16, 2020, when the results of a large-scale randomized clinical trial from Britain, one of all too few such efforts during the pandemic, demonstrated that dexamethasone was able to reduce deaths by one-fifth among patients needing supplemental oxygen and an astonishing one-third among those on ventilators.

The study also explained the earlier findings: Given too early, before patients needed supplemental oxygen, steroids could harm patients. But comprehensive data from the randomized trial showed that when given later, as the disease progressed in severity, dexamethasone was immensely helpful.

Dexamethasone has since become a workhorse of Covid treatment, saving perhaps millions of lives at little cost or fanfare. Without that trial, though, it might never have been noticed because of a problem called confounding: when causal effects of different elements can't be considered separately. If doctors give multiple drugs to patients at the same time, who knows which drug works and which one does not? Or, if they choose which drug to give to whom, those more ill may be getting effective drugs, but the severity of their illness could end up masking the positive effect of the drug. Trials allow us to sort through all of this.

Randomized trials are not the only source of useful data. For example, it would have been difficult to quickly determine how transmissible the Delta variant is — a crucial question — without the data collected from close and systematic observation.

If a variant is spreading quickly somewhere, it might be more transmissible, or it could have simply arrived in that area early and gotten a head start. Or it might have just hit a few superspreader events. We've had variants appear, generating alarming headlines, that were later shown to be no more threatening than previous ones.

When I first wrote about the Delta variant in May, it was data from Britain's Public Health Agency that convinced me it was a real menace, worse even than Alpha, whose increased transmissibility had been systematically discerned earlier by the agency. The British had carefully gathered precise information on who had been infected, how and when, to show that people with Delta were infecting about two-thirds more of their close contacts than those infected with the already highly transmissible Alpha, an alarming number that provided a warning of how viciously it could spread throughout the world.

More crucial data about Delta came from Singapore in June showing chains of transmission that included people who had received Pfizer or Moderna vaccines, demonstrating that breakthrough infections were not just happening, but could lead to further transmission. This finding was possible only because of high-quality contact-tracing. The C.D.C. didn't reach this conclusion until the end of July, citing an outbreak in Provincetown, Mass., that included many vaccinated people — though even then, there was no contact-tracing to show if the vaccinated were transmitting to a substantial degree or merely getting infected.

Without sufficient and timely high-quality data, many scientists have had to try to decipher whatever data is available.

For example, much of the debate over whether vaccine efficacy is waning and boosters are needed has centered on Israel's experience because it started vaccinating earlier than many other countries and is now administering boosters. Many charts and graphics from its Health Ministry about vaccine efficacy and booster effects have been floating around recently, leading to a lot of discussion among scientists and consternation on social media, as well as substantial media coverage.

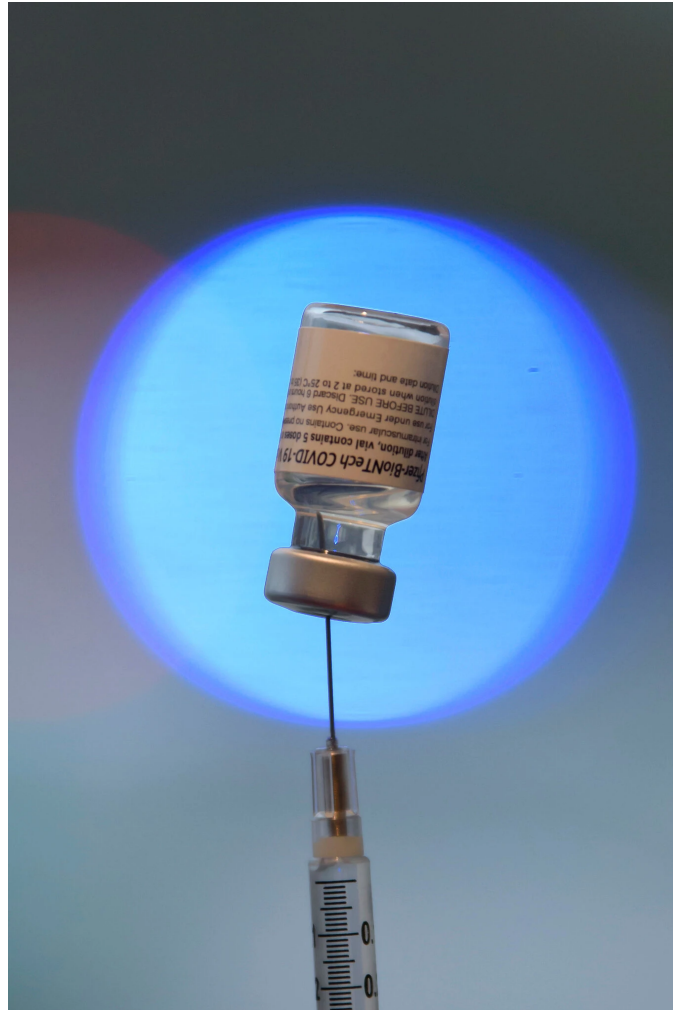
Unfortunately, no raw data, let alone a research paper, was released until weeks after some figures started appearing. That led to scientists squinting at screenshots, trying to reverse engineer graphs. Needless to say, this is less than ideal, not the least because the vaccine and booster data from Israel suffers from confounding. After early reports and charts caused a lot of concern by suggesting the Pfizer vaccine's efficacy may have fallen by as much as 40 percent, an actual preliminary report released weeks later showed that figure was too confounded to be reliable.

In another example, a recent Israeli chart appears to show that a booster shot provides a great deal of protection even a single day after it has been administered — which is essentially impossible. Many things could be going on, including behavior change among those first in line for the booster.

Such confounding can lead to misleading interpretations. It's not that scientists doubt that efficacy can wane over time; the question is one of timing, degree and cause. If vaccination occurred at higher rates, earlier, in Israeli urban areas, and if Delta also had hit urban areas first, a rise in cases might be due to the earlier appearance of the Delta variant in some places, waning protection from vaccination, or a combination of both. If the government responds with boosters, as it is doing, and cases start dropping, is it because of the booster, Delta's natural course, or both?

The best way to answer such questions would have been to systematically collect extensive data and have randomized trials on efficacy and boosters as soon as vaccinations began.

Unfortunately, that has not happened often enough.



Mark Felix/Agence France-Presse — Getty Images

In December, Michael Mina of Harvard's T.H. Chan School of Public Health and I called for a trial, to assess the viability of dose-sparing strategies like delaying second shots, to make vaccines available to more people earlier. Britain, Canada and other countries delayed second shots during the Alpha surge, though there was no trial, so it was harder to pursue such strategies globally even as so many people lost their lives.

To assess the need and effectiveness of boosters, especially for the elderly, a trial could have begun in May or June, when the protective effect of early vaccinations might have begun to wane. By now, we'd have real data rather than a news release from the U.S. Health and Human Services Department announcing that boosters will be available to all vaccinated Americans as early as September, while at the same time saying that is subject to evaluation by the Food and Drug Administration and the Centers for Disease Control and Prevention. If there's data proving the need for boosters, where is it? If not, why did federal officials issue the news release?

All of this is not to say that boosters are useless, or that we should always wait for perfect data before acting, particularly in offering boosters to high-risk groups like the immunocompromised or the elderly. However, announcing that a third Moderna or Pfizer dose will be offered soon even to young, healthy Americans when millions around the world have yet to receive a single dose, requires more than a news release. And ordinary people should not be reduced to trying to decipher such issues by following debates between individual scientists online.

Plus, while extensive data still shows that the vaccines remain remarkably effective against severe disease and hospitalization despite the spread of Delta, social media focuses wildly on vaccinated people with nasty breakthroughs, like those laid up in bed for a week. Even before Delta, we knew some breakthroughs were possible. It's a lack of systematic data that makes these anecdotes harder to interpret and prevents scientists from knowing whether such infections have become more common and dangerous.

Misinformation, which has caused so much damage, thrives under conditions of confusion and uncertainty, particularly when the relevant authorities lose credibility and aren't seen as timely. To this end, systematic and extensive data collection is an investment as necessary as ones for vaccines and therapeutics.

It's not surprising that some of the best data has come from Britain. Britain has a national health care system that makes bigger trials and systematic data collection easier. This epidemiological rigor also speaks to the vision of British scientists who started planning early. It's sadly not a coincidence that the United States, with our fractured, privatized, bureaucratic and bloated health care systems, is so lagging.

Clearly, the Trump administration's negligence and incompetence have put us in a difficult spot, and this is not a problem a new administration can solve in a few months. However, it is one that continues to hinder our pandemic response. For example, compounding the C.D.C.'s failure to track all breakthrough infections, many states that send data to the agency can't determine how many of their hospitalized Covid patients had been vaccinated, Politico has found, making it hard to assess how dangerous breakthrough infections can be.

As encouraging as it was to hear that the C.D.C. was starting a Center for Forecasting and Outbreak Analytics, the bedrock of such efforts is having high-quality data.

In the absence of a more systematic effort, we may even need ad hoc efforts like the remarkable Covid Tracking project, begun by The Atlantic last year when the administration failed to produce data on hospitalizations and cases. The project assembled hundreds of volunteers to make calls around the country and aggregated the data itself.

To cut through this fog of pandemic more effectively, we need to invest in a national infrastructure to coordinate and encourage systematic data collection. "To scent out the truth," as Clausewitz advised.

The Times is committed to publishing a diversity of letters to the editor. We'd like to hear what you think about this or any of our articles. Here are some tips. And here's our email: letters@nytimes.com.

Follow The New York Times Opinion section on Facebook, Twitter (@NYTopinion) and Instagram.

Zeynep Tufekci (@zeynep) is an associate professor at the University of North Carolina, the author of "Twitter and Tear Gas: The Power and Fragility of Networked Protest" and a New York Times Opinion columnist. [@zeynep](#) • [Facebook](#)