

# World view



By Natalie Dean

## Coordinate vaccine studies to stem confusion

**Incompatible research designs will obscure essential answers about vaccine effectiveness. It's time to plan together.**

**M**ore and more COVID-19 vaccines are rolling out safely around the world; just last month, the United States authorized one produced by Johnson & Johnson. But there is still much to be learnt. How long does protection last? How much does it vary by age? How well do vaccines work against various circulating variants, and how well will they work against future ones? Do vaccinated people transmit less of the virus?

Answers to these questions will help regulators to set the best policies. Now is the time to make sure that those answers are as reliable as possible, and I worry that we are not laying the essential groundwork. Our current trajectory has us on course for confusion: we must plan ahead to pool data.

Many questions remain after vaccines are approved. Randomized trials generate the best evidence to answer targeted questions, such as how effective booster doses are. But for others, randomized trials will become too difficult as more and more people are vaccinated. To fill in our knowledge gaps, observational studies of the millions of vaccinated people worldwide will be essential.

Investigators are setting up these studies. One approach is the test-negative design: inexpensive studies that draw from people with symptoms who seek testing. By comparing vaccination rates in those who test positive and those who test negative, we can estimate how effective the vaccine is. This is how we assess the influenza vaccine each year.

Several large test-negative studies are being planned for COVID-19 vaccines. The US Centers for Disease Control and Prevention is conducting a study across multiple sites, with more than 500,000 health-care workers in total. Other studies are being run by the US Department of Veterans Affairs and various government agencies, private health-care providers and academic medical centres. Similar efforts are under way in other countries. There will be several hundred studies at least, and I worry that coordination and cross-consultation will be inadequate.

Imagine what will happen when these studies generate results, each with their own populations, eligibility criteria, validation procedures and clinical endpoints. Differences in study design will cloud answers and prevent cross-cutting conclusions. If we don't want our final answers to be a jumble, we must act now to consider how data can be compared and combined.

The first step is to post study protocols online, on individual websites, as preprints or in journals. This will let trial planners draw on others' insights, jump-starting

**There is rarely a cure for messy data."**

an exchange of ideas to improve designs. For example, experience gained from monitoring influenza-vaccine effectiveness can inform approaches to data collection and validation. The World Health Organization (WHO) intends to maintain a table linking to public protocols, and researchers should proactively make sure that their design is listed.

The next step is to develop and publicize expert consensus on best practices. The WHO has convened a working group on post-introduction vaccine-effectiveness studies, with a report due to be published imminently. This will provide invaluable resources for investigators setting up cohort, case-control and test-negative studies.

But there will be more work to do in translating these recommendations into functional protocols, particularly in countries without extensive influenza-surveillance systems. The WHO, its regional partners and other agencies should disseminate guidance as well as providing technical support and access to data-management consultants, epidemiologists and statisticians. This could include virtual seminars and online training sessions.

Perhaps most importantly, we must coordinate now on plans to combine data. We must take measures to counter the long-standing siloed approach to research. Investigators should be discouraged from setting up single-site studies and encouraged to contribute to a larger effort. Funding agencies should favour studies with plans for collaborating or for sharing de-identified individual-level data.

Even when studies do not officially pool data, they should make their designs compatible with others. That means up-front discussions about standardization and data-quality thresholds. Ideally, this will lead to a minimum common set of variables to be collected, which the WHO has already hammered out for COVID-19 clinical outcomes. Categories include clinical severity (such as all infections, symptomatic disease or critical/fatal disease) and patient characteristics, such as comorbidities. This will help researchers to conduct meta-analyses of even narrow subgroups. Efforts are under way to develop reporting guidelines for test-negative studies, but these will be most successful when there is broad engagement.

There are many important questions that will be addressed only by observational studies, and data that can be combined are much more powerful than lone results. We need to plan these studies with as much care and intentionality as we would for randomized trials.

Unless we act now to ensure the quality and consistency of this research, we will be stuck with muddy findings, trying to look backwards to work out how or whether studies can be compared. There is rarely a cure for messy data. Working out data standards up-front takes time, but will bring essential knowledge. To save lives and livelihoods, share protocols now.

**Natalie Dean** is a biostatistician specializing in infectious-disease epidemiology at the University of Florida in Gainesville. e-mail: nataliedean@ufl.edu