ASIAN BUREAU OF FINANCE AND ECONOMIC RESEARCH

Research Digest

JUNE 2022

Helping Developing Countries without Good Data Infrastructure Measure Vaccine Effectiveness for their Populations

***Abhiroop Mukherjee**, The Hong Kong University of Science and Technology, **George Panayotov**, The Hong Kong University of Science and Technology, **Rik Sen**, University of New South Wales, **Harsha Dutta**, The Hong Kong University of Science and Technology, **Pulak Ghosh**, Indian Institute of Management, Bangalore*

You might have heard of vaccine inequality – the fact that while richer countries are providing third and fourth doses of vaccines to people, many poorer countries still do not have adequate access to one full course of vaccines. But another critical problem is plaguing policy making in many of these countries at the same time: their public health data infrastructure is not good enough to allow them to measure Vaccine Effectiveness (VE) using standard methods.

For example, *view-hub.org*, a comprehensive database of 249 COVID VE studies (as of 5th May 2022), contains only 16 large-scale studies (50,000+ subjects) from low/middle income countries. Further, these studies are conducted in those LMICs which have better data – allowing for standard VE measurement – like Brazil, which accounts for 8 out of these 16. Much of Asia and Africa is left out, creating an information void for public policy in some of the most populous countries in the world.

As a result, these countries are struggling to understand whether and how well vaccines are working for their populations, a key step in deciding the allocation of their limited vaccine doses, and especially important under the continued threat from newer, less vaccine-responsive variants.

In our recent paper in ***Science Advances***, we show how to measure real-world VE using public health records, even in these data-deficient countries.

Don't we know the effectiveness of popular vaccines already?

Yes and no. We know VE from clinical trials, which were mostly conducted before new variants like delta or omicron. We also have some VE data on real-world effectiveness with these variants, but they mostly come from studies conducted in developed countries, like the UK, USA or Israel.

Why do we need VE estimates for individual countries? Is it not enough that we have estimates from developed countries?

Not really enough. VE could differ between populations. Also, different countries could have different levels of pre-existing exposure to Covid-19, and this can affect VE. For example, much of India had some exposure to Covid before 2022 (indicated by "sero-surveys"), but this was not the case in China. As a result, the same vaccine could have different VEs in these two different places.

What is public health data? What is the use of calculating VE from such data?

This is data typically collected by health authorities/governments, and therefore cover a large population. It is useful to calculate VE using large datasets for statistical reasons – e.g., so we know that the estimate we are getting is not driven by some specific segment of population, like doctors and nurses, who may have different VE from the rest of us. Also, with omicron, hospitalization rates are

often low; so if we do not start with a large enough number of people in our data, we will end up with an insufficient number of observations to estimate VE against hospitalization with confidence.

But then why can't all countries measure VE from their own public health data, just like UK or Israel?

The key piece of information that is missing in many countries is the link between data on vaccination and data on infections (or hospitalization/death). In other words, the health authorities in these countries may know how many people in the total population were vaccinated and how many got infected (or hospitalized, etc.), but they often don't know whether a particular person who tested Covid-positive was vaccinated or not. Standard VE measures cannot be calculated absent such a link.

What can be done in data-deficient contexts when such linking information is missing?

This is the problem our paper addresses, using identification methods popularized by last year's economics Nobel Prize winners. We demonstrate how to use a simple technique, called Regression Discontinuity Design (RDD), to estimate VE even with limited data. The method exploits the fact that many governments choose age-based eligibility cutoffs for vaccinations – e.g., those aged 45 and above are eligible for vaccinations, whereas those aged 44 or below are not. This means that many more 45-year-olds in the country will be vaccinated than 44-year-olds; but the average 45-year-old is likely to be very similar to the average 44-year-old health-wise. As a result, absent vaccines – or with an ineffective vaccine – you would expect an equal number of 44-year-olds and 45-year-olds infected or hospitalized. But the more effective the vaccine is, the fewer 45-year olds should end up infected or hospitalized, simply because more of them are vaccinated. This is the key idea.

Can you provide a little more detail about these RDD measures?

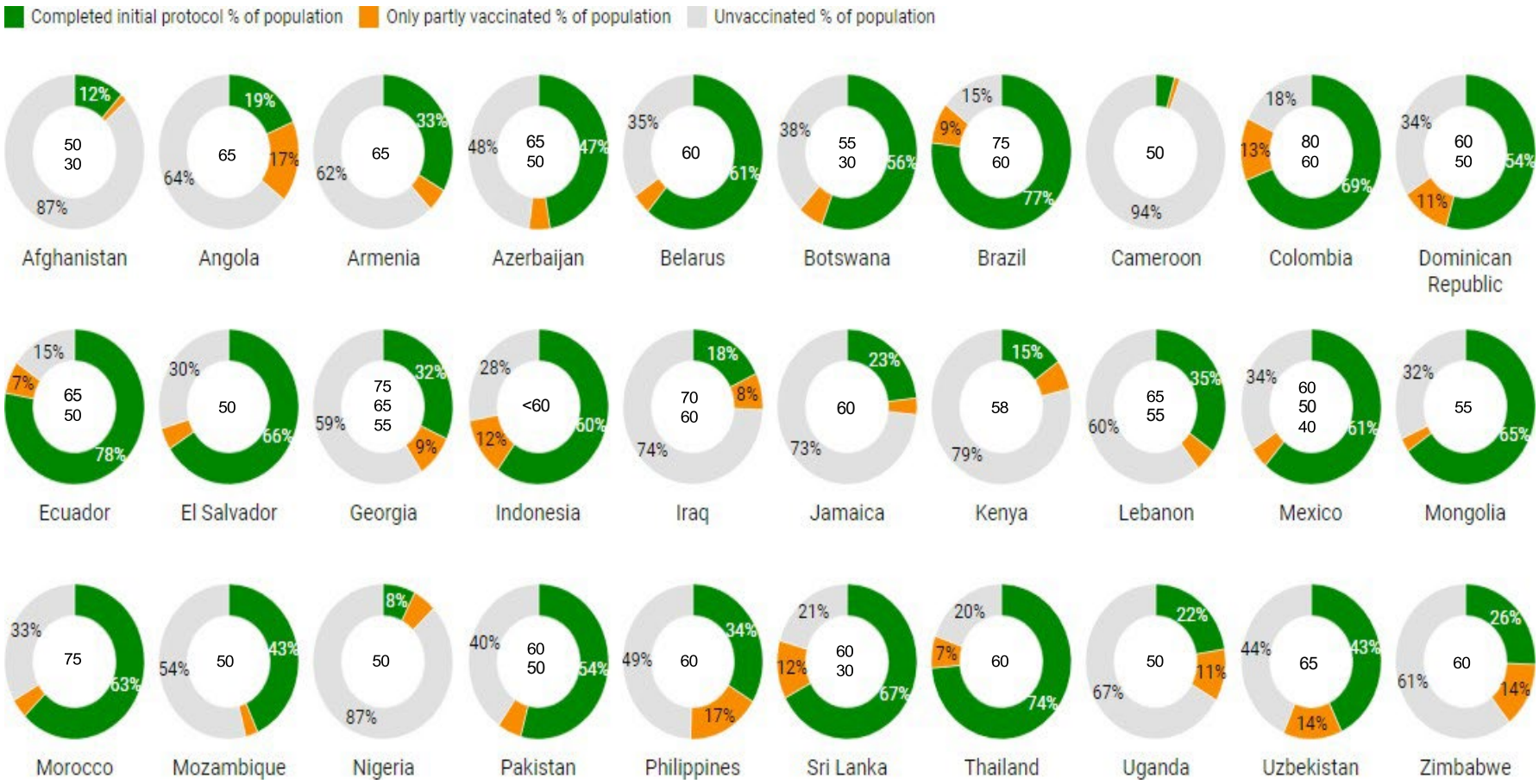
We apply two RDD-based measures of VE. The first measure can be used to evaluate VE against COVID infections, which we cannot link in any way to vaccinations. Our second measure requires knowledge of the vaccination status only for those who get hospitalized (i.e., the breakthrough hospitalizations); when this information is available it provides more robust VE estimates than the first one. Both measures are easy to use, and can be implemented using just an excel spreadsheet.

What do you find using your measure?

We find that even during India's delta-driven second wave, vaccines helped a lot. Our data consists mostly of those who received one dose of the AstraZeneca/Serum Institute *Covishield* vaccine, and we find VE of 55.2% (95% CI: 44.5-65.0%) against symptomatic disease, 80.1% (63.3-88.8%) against hospitalizations, and 85.5% (24.8-99.2%) against ICU/CCU/HDU admissions or deaths. These estimates are based on the largest dataset that has been used to estimate VE, covering 8,755,414 COVID vaccinations, 8,179,635 tests, and 141,800 hospitalizations. Our estimates are also higher than first-dose VE estimates for the same vaccine from the UK against the same variant. This is due, perhaps, to higher pre-existing infection-induced immunity in India.

So is this RDD method applicable beyond India? In how many countries is this method applicable?

Here is a picture showing where different countries are in their vaccination progress, from May 2022. Inside each circle we indicate the age-based eligibility cutoff for each country. Note that almost all countries used such eligibility – and are still using such eligibility criteria – for Covid vaccines. So this chart can serve as an indication of countries where our methods can add value in the Covid context. Most importantly, these RDD methods can be used to estimate VE for any vaccine – and not just Covid vaccines – as long as they are allocated to people using some sort of age-based eligibility criterion. So these methods are widely applicable.


Completed initial protocol % of population
Only partly vaccinated % of population
Unvaccinated % of population
Afghanistan
Angola
Armenia
Azerbaijan
Belarus
Botswana
Brazil
Cameroon
Colombia
Dominican Republic
Ecuador
El Salvador
Georgia
Indonesia
Iraq
Jamaica
Kenya
Lebanon
Mexico
Mongolia
Morocco
Mozambique
Nigeria
Pakistan
Philippines
Sri Lanka
Thailand
Uganda
Uzbekistan
Zimbabwe