



16 May 2022

Dear Editor,

Please find enclosed, our manuscript entitled “Low testing rates hamper effectiveness of genomic surveillance programs for early detection of novel SARS-CoV-2 variants: a mathematical modelling study”, which we would be grateful for your consideration for publication in *The Lancet*.

Genomic surveillance has been at the forefront of the global COVID-19 pandemic response for monitoring the evolution of SARS-CoV-2, and has served as the primary tool for identifying variants of concern (VOCs). However, the reliability of genomic surveillance to provide timely information on the emergence and spread of VOCs requires worldwide contribution of SARS-CoV-2 genomic data, as evidenced by the emergence of VOCs such as Alpha in the United Kingdom, Gamma in Brazil, Delta in India and Omicron in Southern Africa. Yet, to date, high income countries (HICs) have produced ~20 times more genomic surveillance data than low- and middle-income countries (LMICs). The problem with this disparity is obvious when looking at the emergence of the Omicron VOC – Omicron had already spread through the majority of the population of South Africa and to at least five other countries before being first described in early December 2021. The consequences of low testing and sequencing rates are now looming larger as many countries around the world begin to dismantle parts of their surveillance programs in the post-crisis phase of the pandemic.

To enhance genomic surveillance capacities, much focus, and thus investments have been channelled towards boosting sequencing volumes in LMICs. Genomic surveillance guidelines from public health agencies, including the World Health Organization (WHO) and other academic groups have largely focused on estimating the number of cases to be sequenced to ensure robust variant detection and prevalence monitoring. However, all of these guidelines fundamentally assume high levels of COVID-19 diagnostic testing such that the sequenced samples are representative of the viral diversity circulating the population. These guidelines are infeasible in most LMICs given their low COVID-19 testing rates (i.e. mean ~ 27 tests per 100,000 people per day (tests/100k/day) relative to >800 tests/100k/day in HICs between 2020 and March 2022) which consequently leads to substantial bias in the samples collected for sequencing.

To identify meaningful surveillance targets for LMICs that account for both testing and sequencing needs, we developed an agent-based modelling framework and simulated extant virus/VOC epidemics (i.e. SARS-CoV-2 wild-type/Alpha variant and Delta/Omicron variant) in a prototypical LMIC under a range of diagnostic tests availability and different genomic surveillance strategies for variant detection. We found that under the mean LMIC test availability of 27 tests/100k/day, current sequencing sample size guidance provided by WHO and other academic groups would either lead to weeks-to-months delays in new variant detection or failure to detect novel VOCs at all owing to the inherent spatiotemporal bias resulting from low testing rates.

Crucially, we found that diagnostics testing rates play a far more important role in the early detection of novel SARS-CoV-2 variants and monitoring of variant growth than the amount of genetic sequencing. Increasing testing rates to at least 100 tests/100k/day and sequencing 5-10% of collected specimens represents a critical juncture to ensure the prompt detection of novel variants (i.e. within one month since its introduction and <25% of circulating variant proportion) and accurate measurements of their circulating prevalence. Importantly, other than population-wide surveillance efforts, this could just as well be achieved through sentinel surveillance systems if selected sentinel sites provide sufficient geographical coverage of the population (i.e. mean ~10 samples for sequencing per 1,000,000 people per week).

We found that for countries with low testing capacities, investment priorities should be given to increasing diagnostic tests availability and access to testing. Sequencing quantities should only be of concern after a vital amount of diagnostic capacity is already in place and even then, excessive boost in sequencing capacities would only yield diminished returns in variant detection and monitoring at the expense of resource wastage.

While our results were derived from modelling surveillance outcomes in LMICs, our findings on the impact of diagnostic testing on genomic surveillance programs are pertinent to HICs as well, especially since multiple countries are beginning to scale down surveillance efforts in the current post-crisis phase of the COVID-19 pandemic. In turn, our work will be of key interest to the readers of The Lancet, including governments, clinicians, epidemiologists, global, and public health experts working on pandemic mitigation and pathogen surveillance.

Thank you for your consideration and we look forward to your decision.

Sincerely,

Alvin X. Han, Brooke E. Nichols and Colin A. Russell on behalf of all authors