Reviewers’ comments in blue; Response in black.

Your Article, "SARS-CoV-2 diagnostic testing rates determine the sensitivity of genomic surveillance programs" has now been seen by 3 referees. You will see from their comments below that while they find your work of interest, some important points are raised. We are interested in the possibility of publishing your study in Nature Genetics, but would like to consider your response to these concerns in the form of a revised manuscript before we make a final decision on publication.  
  
In brief, two referees think that your work is an important one that merits publication, while the third sounds more circumspect at this stage.  
  
Reviewer #1 provides a largely positive report; they don't ask for substantial further work, but seem to request further discussion of your findings.  
  
Reviewer #2 thinks this is "an impressive and important contribution", but they have a number of technical concerns. To our reading, these seem reasonable and do not seem unduly hard to address.  
  
Reviewer #3, as noted, is the most skeptical. This appears to be not due to any major technical or methodological concern (though they do suggest the model is overly complex - perhaps suggesting the advance presented by your model needs to be better explained?), but can be summed up in the following quote: “From a public health perspective, the article, in the reviewer’s view, lacks important insights that would allow translating these results into practice.”  
  
Across these three reviews, we think there is sufficient support for a potential future publication. We think the requests made by referees seem largely reasonable and doable. We would, in particular, highlight the suggestion that how these results translate into public health policy needs to be discussed in much more depth (broadly shared by all reviewers), as well as clarifying the technical details of your model (Reviewers #2, #3).

We thank the editor and all reviewers for taking the time to review our manuscript and providing feedback to improve our manuscript. We have addressed all of the reviewers’ comments (see point-by-point response below) and revised the manuscript accordingly.   
  
Reviewer #1

The paper describes the need for timely detection of new SARS-CoV-2 variants by expanding the testing rates, especially in LMIC that currently have less that 27/100K people/day. The authors make a compelling argument that low testing rates and spatiotemporal biases delay the time to detection of new variants by weeks to months.

We thank the reviewer for carefully considering our manuscript.

1. The manuscript is very well written, and the analysis are generally robust and provide statistical evidence to support a view that is intuitive and generally known within the community, which is that low test availability affects the detection of new variants introduced into a population as well as accurately measuring the proportion of circulating variants. Indeed, the “guidance for surveillance of SARs-CoV-2 variants” manual published by the WHO in August 2021 highlighted the specificity and sensitivity of sequencing strategies, giving recommendations on sampling strategies to increase the specificity and sensitivity of variant detection.

We thank the reviewer for the positive feedback of our work.

1. The authors highlight how non-uniform sampling results in spatiotemporal bias and this delays detection of VOCs. It would be useful to include a short overview of the various sampling strategies that have been used in published reports from LMICs. Due to the peculiar challenges with storage and logistics in many LMICs, sampling strategies can also be purposive and convenient, especially when there is a collaboration between sequencing institutions and the sentinel sites. There are interesting examples of sampling positive samples from passengers arriving at airports, which could provide an early detection strategy for imported new variants before they spread in the community (see Genetic diversity of SARS-CoV-2 infections in Ghana from 2020-2021 - PubMed ([nih.gov](http://nih.gov)).

We agree and have now provided a high-level overview on genomic surveillance efforts and challenges in LMICs, as well as how our work can help address some of these challenges:

Line 351: “*During the initial phase of the pandemic in 2020, due to limited testing and sequencing capacities, many LMICs were initially focused on genomic surveillance efforts at points of entry at country borders to deter introductions27–29. Over time, especially after the emergence of VOCs, SARS-CoV-2 genomic surveillance gradually expanded to include community surveillance as many LMICs enhanced their sequencing capacities4,29–31. This was done either by establishing regional sequencing networks to maximize available resources, investing in local sequencing capacities or partnering with global collaborators31–33. Sequencing turnaround time has also improved from an average of ~170 days in 2020 to ~30 days in 2021 across the African continent, albeit with substantial variation among countries31. While sequencing capabilities have expanded in LMICs, obtaining spatiotemporally representative samples remains a key challenge31. Our work shows that the sensitivity of genomic surveillance programs is highly dependent on diagnostic testing rate and that a mean testing rate of 100 tests/100k/day at sentinel sites that are geographically spread out across the community is a good basis for monitoring virus variants. While a reflexive PCR test after a positive Ag-RDT diagnosis is currently performed to obtain samples suitable for sequencing (and is possible in many tertiary facilities in LMICs), this presents additional cost and logistical barriers. Recent studies showed that SARS-CoV-2 sequencing can be performed using materials obtained from Ag-RDTs performed at point-of-care34–36. Importantly, whole genomes can be recovered up to eight days after testing, providing opportunities for sequencing to be performed on samples performed through self-testing as well34.*”

1. A bit more attention could also be given to the impact of time between sample collection and sequencing as well as turn-around time, considering difficulties in obtaining reagents and servicing equipment in LMIC.

We agree and we now state that this is a particularly prevalent issue faced by many LMICs that impacts turnaround time:

Line 245: “*Third, reducing turnaround time from samples referred to sequencing output results in a 1:1 decrease in time to new variant detection regardless of the proportion of sequenced samples, test availability or sampling coverage (Fig. 5). These gains require scale up in sample transport networks, access to sequencing machinery, trained personnel, and/or increases in numbers of sequenced samples to make the most efficient use of each sequencing run12.* ***Furthermore, LMICs also often face high costs and extended delivery delays of laboratory reagents and consumables that were sometimes further exacerbated by recurring travel bans during the acute phase of the pandemic****5,13,14.*”

Reviewer #2

This manuscript seeks to address an important challenge in the public health response to SARS-CoV-2 using a detailed simulation of varying testing and sequencing strategies on SARS-CoV-2 surveillance outcomes. Diagnostic testing and genomic surveillance are crucial to not only understanding the burden of infection, but also viral transmission and evolutionary dynamics. The authors note that the diagnostic testing and sequencing capacity of LMICs is far below that of HIC and global testing and sequencing rates are falling as public health officials shift priorities in responding to the SARS-CoV-2 pandemic. The authors simulate COVID-19 Alpha and Delta/Omicron infection waves in a population of 1,000,000 people with characteristics similar to Zambia. In each of these scenarios, they simulated varying diagnostic testing and genomic surveillance strategies to examine the impact of diagnostic testing rates, spatiotemporal bias, and volume of samples selected for genomic sequencing on time to detection and prevalence estimation for emerging variants. They conclude that increased testing capacity, representative sampling, and sequencing 5-10% of positive specimens would have a meaningful impact on SARS-CoV-2 surveillance outcomes, particularly in LMICs. The comprehensive analysis undertaken in this well-written manuscript is an impressive and important contribution to public health. However, there were several areas where I had concerns or additional clarifying information is needed.

We thank the reviewer for the careful consideration of our manuscript.

1. A clearer explanation of tertiary facilities would be greatly beneficial to readers. For example, the authors state that 20% of healthcare facilities were assumed to be tertiary facilities (line 313-314). How does this translate to the assumed testing volume at tertiary facilities?

How are tertiary facilities assumed to be distributed spatially?

We have revised the methods section to describe how tertiary facilities were modelled more clearly:

Line 399: “*The simulated number of healthcare facilities (i.e., community clinics and tertiary hospitals) where individuals with mild symptoms seek symptomatic testing and have their virus specimens collected was based on an empirical clinic-to-population ratio (i.e.* ***one healthcare facility for every 7,000 individuals on average****)38,39.* ***Although PATAT does not explicitly simulate the spatial location of individuals, contact networks and healthcare facilities are ordered to approximate localized community structures (i.e. the closer the number order of a facility, the closer they are in the same neighborhood) that is most illustrative of urban centers. Households are proximally ordered and distributed around these facilities based on an empirical distance-structured distribution that correlates with probabilities of symptomatic individuals seeking testing at clinics*** *(Extended Data Table S1).*”

Line 489: “*Twenty percent of healthcare facilities were assumed to be tertiary facilities based on empirical data collected from Zambia38,39. We assumed that tertiary facilities provide testing for mild symptomatic individuals as well as hospitalized patients with severe symptoms.* ***Given that healthcare facilities were proximally ordered, we randomly selected tertiary facilities in each independent surveillance simulation (see below) but ensured that the selected facilities were not consecutively ordered. In sum, all tertiary facilities accounted for a median 18.4% (interquartile range = 17.7-19.1%) of total testing volume across all simulations****.”*

1. I missed the level of confidence for time to detection. Further, the number of days to detection in Figure 3 seems quite low given the assumed sampling volume. In the population-wide strategy with 27 tests/100k/day, the figure shows that if 20% of positive specimens were selected per day for sequencing, the first specimen of the emerging variant would be sampled on day 14 (~1.6% prevalence based on Figure 2). Assuming 20% test positivity, this means that ~11 samples per day would be selected for sequencing. Using binomial sampling theory, this scenario seems to have a low probability of occurrence unless there is an assumption I am missing.

Figure 3 plots the “expected day when the first variant specimen to be sequenced is sampled since variant introduction” as computed from the 1,000 independent random surveillance simulations. We have revised the plot to include the standard deviation of the day of first variant specimen collection to reflect variability in detection outcomes across simulations.

Importantly, testing rate (in the model and in reality) is not static. Many LMICs could only meet testing demand outside of periods when infections peaked but struggled to maintain sufficient testing capacity during the peak of outbreaks (Salyer et al., The Lancet, 2021). This is also known to us through our FIND co-authors working on testing distribution with local public health officials in many LMICs. As such, the reported 27 tests/100,000 people/day is the “*mean* testing rate” in LMICs over time. It is not the fixed number of tests performed per day. As stated in the Methods section, we had assumed that test stocks were replenished weekly on top of any unused tests from the previous week. This thus largely reflected the situation in most LMICs where testing demand was largely fulfilled at the start and end of an epidemic wave but difficult to maintain during peak infections:

Line 459: “*We varied levels of Ag-RDT stocks per day (i.e., 27, 100, and 200-1,000 (in increments of 200) tests/100k/day), running 10 independent epidemic simulations for each testing rate. Given the start of a week on Monday, we assumed that a week’s worth of tests are delivered to healthcare facilities every Monday and unused Ag-RDTs in the previous week are carried forward into the next week. If test stocks for a particular week were exhausted before the end of the week, testing for the rest of that week ceased.*”

In other words, for 1,000,000 people, there could be as many as ~1890 tests available and ~380 diagnoses (assuming 20% test positivity) at the start of a week. With 20% of positive specimens sequenced (~75 sequences), by binomial sampling, the probability of obtaining at least one variant sequence at 1.6% variant proportion is ~70%.

1. The number of days to detection in the one tertiary facility scenario for 27 tests/100k/day and 100 tests/100k/day seems to differ by ~1 week between Figure 3 and Figure 5. Where is this discrepancy coming from? This seems notable given that a conclusion drawn from this analysis is that increasing testing capacity from 27 tests/100k/day to 100 tests/100k/day can speed up detection by 1-2 weeks.

Figure 3 shows the “expected day when the first variant specimen to be sequenced is sampled since variant introduction” whereas the plot in Figure 5 assuming negligible turnaround time depicts the “expected day when the first variant sequence is generated”. We have now made this clearer in the figures and corresponding captions.

These timing would differ by ~1 week because we “*assumed that positive specimens sampled within each week for sequencing are consolidated into a batch before they are referred for sequencing. Turnaround time refers to the time between collection of each weekly consolidated batch of positive specimens to the acquisition of its corresponding sequencing data*” (line 507 and caption of Fig. 5).

1. Did the authors account for correlated samples in monitoring the observed variant proportion?

The observed variant proportion is computed from the sequencing pool of samples which were randomly sampled from whichever source they were collected from (i.e. population-wide or varying number of sentinel tertiary facilities). As such, in the population-wide strategy where sequencing specimens were randomly sampled from the entire community, the observed variant proportion would be less impacted by spatiotemporal biases from highly correlated samples that are more likely to emerge from sampling at a single sentinel facility.

1. A deeper discussion of how these findings could be used by a public health department in planning resources (how to select tertiary facilities, etc) and how influential the Zambia-specific parameters are on these findings would be beneficial.

See response to reviewer 1, point 2 for discussion on how to apply our findings to public health policy.

Even though we performed our simulations using demographic parameters from Zambia, our findings can be broadly applied to other countries where VOCs emerged, as described in the main text and the Supplementary Notes. We have also compared our findings to Germany, high-income country example, as suggested by reviewer 3.

Line 285 in main text: “*Despite performing our simulations using demographic parameters from Zambia, the emergence and detection of each VOC to date represents interesting case studies for the work described here (Supplementary Notes). For example, at the time of first detection of the Omicron variant, in South Africa in November 2021, the daily SARS-CoV-2 testing rate was 51 tests/100,000 people/day9, which was among the highest testing rates in Africa. However, the Omicron variant was only detected 6-8 weeks after its likely emergence4. At that point, Omicron had already infected a substantial portion of the population in Gauteng, South Africa (i.e., the estimated circulating variant proportion was >80% by mid-November)4. Not only had the variant already spread across the rest of South Africa and to neighboring Botswana4, Omicron samples were also collected in multiple other countries, including Hong Kong15, Denmark16, and the Netherlands17 before the initial reports on the identification of the Omicron variant. This situation is consistent with our modelling findings, where novel variant detection is possible with <100 tests/100k/day but only after the new variant has spread widely across the population.*

*In another example, Germany randomly sequenced ~60-70 sequences per week (i.e. <1% of cases sequenced per day) in December 202018. During this time, testing rates in Germany averaged at ~300 tests/100k/day9. Germany was able to detect the Alpha variant one week before WHO declared the lineage a variant-of-concern in mid-December 202018. The Alpha variant likely emerged in the UK in mid-September 202019 and rapidly proliferated across the country before it was reported in December 202020. Our analyses showed that the expected time before the first Alpha variant specimen was sampled for sequencing since its introduction is >4 weeks (i.e. around November 2020) at Germany’s testing and sequencing rate. This falls in line with the likely period of Alpha’s introduction into Germany, similar to the period estimated for its European neighbors such as the Netherlands21.”*

1. In the assessment of the performance of current guidance (lines 84-100), the authors determine that all three approaches that they examined were insufficient when testing rates were low due to poor representativeness. It is a helpful exercise to contextualize the methods in the current analysis within the currently available literature. However, more information (in addition to what is presented in Table 1) on how these other approaches differ from the current analysis, particularly using the population-wide strategy, would be beneficial to the reader for a head-to-head comparison.  
     
   Additionally, based on the information presented in Table 1, I am particularly unclear about how the authors implemented the approach from Wohl et al given the critical considerations listed in the table. In the second point, the authors state that the method presented in Wohl et al assumes asymptomatic patients are tested as well. However, the linked spreadsheet (and the manuscript) show that the method can be applied to different scenarios with varying levels of asymptomatic testing. Also, in the first point, the authors state that the method assumes the observed variant proportion is representative of the circulating proportion. However, this paper provides details on a correction factor to convert between observed prevalence and actual prevalence.

We have now provided further context of the three guidance analyzed in the main text and a full section in the Supplementary Notes, detailing the mathematical background and assumptions made in each method, specifically pointing out how they did not incorporate biases arising from low testing rates and spatial nonuniformity in sampling coverage in their sample size calculations.

Line 104 in main text: “*We used recommended sample sizes from three prominent guidance: (1) The World Health Organization and European Centre for Disease Prevention and Control computed sample size using the binomial method7,8; (2) By subsampling genomic surveillance data generated in Denmark in 2020-2021 when the country was testing at >2,000 tests/100k/day on average, Brito et al. suggested that sequencing 0.5% of all detected cases with a turnaround time of 21 days would result in a 20% of variant detection before reaching 100 cases5; (3) Wohl et al. formulated a novel framework computing sequencing sample size by modeling the biological and logistical processes that impact sampled variant proportions6. Critically, all three methods did not consider how low testing rates and spatial nonuniformity in sampling coverage impact sampled variant proportions, and in turn, speed of variant detection. The assumptions and mathematical background, specifically how they fail to factor in spatiotemporal bias in sample size estimation, of each guidance are detailed in Table 1 and Supplementary Notes.*”

In particular for the approach from Wohl et al., we provided explanations in the Supplementary Notes as to why even though the probability of asymptomatic testing is included in their correction factor, it only corrects for biases in the observed variant proportion arising from relative differences in diagnostic sensitivities, sample qualities and conditional asymptomatic and symptomatic testing probabilities between the two viral variants. It does not factor in distortions in the observed variant proportion from other sources of bias, including: (1) stochastic effects arising from low testing volumes and unevenness of daily test stock availability, (2) unconditional probability of asymptomatic testing, and (3) spatial biases when only a subset of samples from sentinel sites are sent for sequencing.

Reviewer #3

The authors present the results of a relatively complex modelling study with which they simulate different testing and sequencing strategies and based thereon compute the temporal delay in detecting novel variants, as well as the accuracy in estimating the frequency of a novel variant.

We thank the reviewer for the thorough consideration of our manuscript.

1. B & E: Originality & significance; reliability of conclusions:   
   The approach is quite complex for deriving the ‘not too surprising’ results: (i) More testing and (ii) less biased testing shorten the time to detection of novel emerging variants, as well as their abundance determination.

We respectfully disagree with the reviewer’s opinion that our results were “not too surprising”. As described in our paper, current guidance largely focused on estimating the number of samples to sequence for variant detection. While the WHO and ECDC acknowledged that testing coverage must be evenly distributed to reduce bias, to our knowledge, there are no evidence-based advice on the volume of testing and amount of sentinel surveillance needed to achieve different SARS-CoV-2 variant detection targets. Importantly, the interplay between testing rates, sampling bias and sequencing proportions and the degree to which testing frequencies impact surveillance outcome were not previously well understood and appreciated by the audiences whom we presented this work to, including our WHO co-authors as well as numerous key figures throughout the public health community. To this end, our work has already provided a much-needed evidence base for a variety of public health decisions and recommendations.

1. From a public health perspective, the article, in the reviewer’s view, lacks important insights that would allow translating these results into practice. E.g., what is the trade-off between the increased expenditure of more testing and the public health gains? For that, the authors should evaluate the costs/cost increase of the expanded testing (and sequencing), and quite importantly define the broader goal of the genomic surveillance program. This goal will determine the requirements: I.e. is the goal the containment of a VOC (in a region/country)?; Is it to inform vaccine development?; or to be able to adapt diagnostic assays, e.g. primer sets?, etc.…

We agree with the reviewer that a comprehensive evaluation on cost effectiveness of expanded testing is an important step towards translating our results to practice. However, such analyses require considerations on both healthcare expenditures and societal costs, infrastructure building and implementation under different economies and community structures, all of these within the scope of attaining various public health goals, some of which as listed by the reviewer. This is beyond the scope of this work which is to critically evaluate what could be achieved by current guidance and elucidate the barriers preventing current guidance from achieving the desired surveillance outcomes in terms of variant detection and prevalence estimation. This does not imply that our work lacked “important insights” to meaningfully inform public health decisions (see response to point 1).

We continue to collaborate with our co-authors from FIND on modelling and cost effectiveness studies of expanded testing and surveillance frameworks with the aims of achieving different public health goals. In fact, in a companion preprint, we have used the agent-based model presented in this work to estimate the impact of expanded testing and strategies in reducing SARS-CoV-2 transmissions (<https://www.medrxiv.org/content/10.1101/2022.06.16.22276516v1>).

1. A discussion on the feasibility of various testing strategies and possible barriers would be required: I.e., scaling-up population-wide testing may become infeasible as SARS2 becomes endemic; Is it feasible to sequence on site (in Zambia)? What are the current barriers?

See responses to reviewer 1, points 2 and 3.

We are not recommending large-scale population-wide testing but to improve/maintain testing rates to adequate levels at sentinel facilities that are geographically representative of the population.

1. C. Methodology:  
   On the technical side, the reviewer wonders, whether the model’s complexity, which also complicates the analysis (and leaves significant carbon imprint for simulation), is really needed to derive at the conclusions presented.

The complexity in our model simulating where and when individuals are infected, tested, and in turn, where and when samples were collected are all needed for us to characterize the interplay between testing rates, sampling biases and sequencing proportions. Furthermore, as mentioned in our response to point 2, the model and simulations performed in this work have been used to interrogate the impact of testing in reducing SARS-CoV-2 transmissions.

1. Technical comments:  
   Page 17, 19: What do you mean by ‘…running 10 bootstrap simulations for each testing rate.’  
   • Do you mean …”running 10 stochastic simulations”? To my knowledge bootstrap-resampling refers to sth. else than to numerically realize (simulate) a stochastic process.

We have now replaced “bootstrap” with “independent”.

1. Numerics: Details about the numerical realization of the simulations is missing. From my understanding, you perform parallel updates (i.e. fixed time step of 1 day)?   
   • Details on the implementations are missing. Which distributions did you sample from?  
   • Did you assess the appropriateness of the procedure and the chosen time step in terms of accuracy with exact (Gillespie or hybrid) simulations?

Type of contact network: What type of network is it? Static, temporal or adaptive (i.e.: [https://doi.org/10.1007/978-3-642-01284-6](https://eur04.safelinks.protection.outlook.com/?url=https%3A%2F%2Fdoi.org%2F10.1007%2F978-3-642-01284-6&data=05%7C01%7Cx.han%40amsterdamumc.nl%7C7386dbe495eb4cda975108da81e5ef2a%7C68dfab1a11bb4cc6beb528d756984fb6%7C0%7C0%7C637965121794065173%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C3000%7C%7C%7C&sdata=3OQ4btEBK5tevzgBkkv74bRBW7Zl9EwBv3bF7lYC1h0%3D&reserved=0)). From the supplementary methods, the model appears adaptive to the reviewer (individuals change their contact behavior upon diagnosis). Importantly, in this case, parallel updating will be numerically incorrect (the question is to what extent?).

Full details of our model can be found in the Supplementary Notes and Table S1. This includes full descriptions of the distributions that we had sampled from and the parameters used.

We performed sequential updates in each time step and this is now explicitly clarified in technical details of the model in the Supplementary Notes:

“*The sequential operations during each timestep follow the following order: (1) update the disease progression of infected individuals, (2) update the status of isolated/quarantined agents, (3) application of community testing strategies and (4) computation of transmission events within contact networks.”*

As such, while we assumed adaptive contact networks with individuals changing their contact patterns by isolating after a positive diagnosis, we would not run into numerical errors associated with parallel updating.

Agent-based models with sequential updates in time step of a day have also been developed and validated by others to simulate the SARS-CoV-2 transmissions. Some prominent examples include Covasim (Kerr et al., 2021; <https://doi.org/10.1371/journal.pcbi.1009149>), López et al., 2021 (<https://www.science.org/doi/10.1126/sciadv.abd8750>), Faucher et al., 2022 (<https://www.nature.com/articles/s41467-022-29015-y>).

1. G. References/state of art/significance  
   Several recent advances have been made in the field of genomic surveillance of SARS2 that need to be discussed:

What about more cost-efficient, quicker and easier to implement/already implemented methods in combination with sequencing, such as:  
a. Spike gene target failure (SGTF) assays?  
b. Waste-water sampling & sequencing, e.g. [https://doi.org/10.1038/s41564-022-01185-x](https://eur04.safelinks.protection.outlook.com/?url=https%3A%2F%2Fdoi.org%2F10.1038%2Fs41564-022-01185-x&data=05%7C01%7Cx.han%40amsterdamumc.nl%7C7386dbe495eb4cda975108da81e5ef2a%7C68dfab1a11bb4cc6beb528d756984fb6%7C0%7C0%7C637965121794065173%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C3000%7C%7C%7C&sdata=yJksWXk0iiWeeirL6Bs%2BooqGsxA00%2BAQ4HIlOYY49wU%3D&reserved=0)   
c. Genome-based incidence estimation to assess e.g. case ascertainment [https://doi.org/10.1038/s41467-021-26267-y](https://eur04.safelinks.protection.outlook.com/?url=https%3A%2F%2Fdoi.org%2F10.1038%2Fs41467-021-26267-y&data=05%7C01%7Cx.han%40amsterdamumc.nl%7C7386dbe495eb4cda975108da81e5ef2a%7C68dfab1a11bb4cc6beb528d756984fb6%7C0%7C0%7C637965121794065173%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C3000%7C%7C%7C&sdata=TN2qHBuoJz6pgZHss6idc4NqGknxqZDqXdq7ehSaB2E%3D&reserved=0) to inform adaptive testing (maybe even sequencing) strategies?

We have now expanded our discussion on the recent advances in genomic surveillance through (a) SGTF assays and (b) wastewater surveillance in the Discussion:

Line 327: “*While we find that routine representative sampling is vital for monitoring SARS-CoV-2 evolution, additional surveillance systems, including targeted surveillance of particular populations and settings (such as immunocompromised individuals or unusual events), and wastewater sampling, could enable increased variant detection sensitivity22. In particular, recent advances in wastewater sequencing and deconvolution methods to resolve multiple viral lineages in mixed wastewater samples enabled detection of emerging variants before they were captured by clinical genomic surveillance23–25. However, sequence quality is often poor in wastewater samples and in turn, these methods depend on a priori knowledge of the lineage-defining mutations of VOCs and variants-of-interest, which are currently still identified based on significant upsurges in clinically diagnosed cases. Furthermore, centralized wastewater management systems, which these methods rely on for accurate determination of relative lineage prevalence, are currently non-existent in many LMICs. Substantial investments, coordination and time are needed to enable local sanitation infrastructures suitable for wastewater surveillance26. Detection of genetic markers such as S-gene target failure in PCR assays may also provide faster notification of viral lineages with these specific mutations. However, whole genome sequencing is still needed for unambiguous genotyping of SARS-CoV-2 samples. Ultimately, clinical diagnostic testing and surveillance will remain the core mode of SARS-CoV-2 surveillance in most countries.*”

For reference provided in (c), this is a novel approach for estimating incidence using SARS-CoV-2 sequencing data by leveraging on soft selective sweeps to infer effective population size as a correlate of underlying incidence. In turn, this depends on sufficient and random testing and genomic surveillance to accurately capture temporal changes in mutation frequencies among the viruses in circulation. Our work thus provides testing targets and genomic surveillance strategies necessary to facilitate the use of this incidence estimation approach.

1. A comparison or discussion of modeling results with actual results from targeted surveillance programs (~70 sequences per week), such as e.g.: [https://doi.org/10.1093/cid/ciac399](https://eur04.safelinks.protection.outlook.com/?url=https%3A%2F%2Fdoi.org%2F10.1093%2Fcid%2Fciac399&data=05%7C01%7Cx.han%40amsterdamumc.nl%7C7386dbe495eb4cda975108da81e5ef2a%7C68dfab1a11bb4cc6beb528d756984fb6%7C0%7C0%7C637965121794065173%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C3000%7C%7C%7C&sdata=chBaapKlJW70NRWOJGVe92Jp2Tq2Wc1G8I54%2Bf%2FwlI4%3D&reserved=0)

We have now included this comparison in the Discussion:

Line 299: “*In another example, Germany randomly sequenced ~60-70 sequences per week (i.e. <1% of cases sequenced per day) in December 202018. During this time, testing rates in Germany averaged at ~300 tests/100k/day9. Germany was able to detect the Alpha variant one week before WHO declared the lineage a variant-of-concern in mid-December 202018. The Alpha variant likely emerged in the UK in mid-September 202019 and rapidly proliferated across the country before it was reported in December 202020. Although our simulations were performed for a Zambian population, the expected time before the first Alpha variant specimen was sampled for sequencing since its introduction is >4 weeks (i.e. around November 2020) at Germany’s testing and sequencing rate. This falls in line with the likely period of Alpha’s introduction into Germany, similar to the period estimated for its European neighbours such as the Netherlands21.”*

1. Currently, I cannot see how the results would translate into national public health action. This would require to adress points raised above (A & E). Also, the novelty of the insights is not clear.

We hope that our responses to points 1 and 2 above (as well as the responses to the other reviewers) make how this work helps to shape policy and investment strategies clearer.

1. It may be of advantage to differentiate between the de novo emergence of a variant and its introduction from the outside. This may imply different public health objectives, as well as the feasibility to obtain of these objectives (e.g. such as containment) with the help of genomic sequencing. Also, the turnaround time (sample-to-sequence) is critical to the feasibility of some of these objectives (due to rapid exponential growth of SARS2).

We agree and have now indicated the differentiation between *de novo* emergence and external introduction:

Line 268: “*Our findings serve to inform expectations of genomic surveillance initiative and should be interpreted according to the public health objectives of each program. If the objective is to serve as an early warning system for the de novo emergence of new variants before they are likely to have spread widely, then all factors above can be considered essential and could require substantially more than 100 tests/100k/day. Critically, determining that a new variant is a threat requires not only detection of the variant itself but also the capacity to reliably monitor changes in its prevalence and potential clinical impact on short timescales. The results presented here also inform the design of programs for the sensitive and reliable detection of changes in variant prevalence. Otherwise, if the objective is to detect for introduction of novel variants from overseas, some of the factors above may be relaxed depending on the public health objectives. For instance, if the aim is to attempt containment, all factors should still be considered to promptly detect and monitor the spread of the variant. However, if the aim to ensure sufficient time for control strategies to be enacted, less samples could be sequenced or turnaround time could be longer, for example, so long as the mitigation strategies remain useful when implemented.*”

Additional discussion on turnaround time is now included. See response to reviewer 1, point 3.

1. Ag-RDT availability may be decoupled from samples being available for sequencing, as Ag-RDT samples cannot be used directly in sequencing, unlike PCR samples. According to the manuscript it is proposed to verify positive Ag-RDT samples with PCR. Is this even feasible in e.g. Zambia (availability of technical equipment) and what would be the investment required to make it feasible?

As elaborated in point 2, estimating the economic investment is highly complex and outside of the scope of this work. However, we have now discussed the feasibility of using Ag-RDT samples for sequencing:

Line 364: “*While a reflexive PCR test after a positive Ag-RDT diagnosis is currently performed to obtain samples suitable for sequencing (and is possible in many tertiary facilities in LMICs), this presents additional cost and logistical barriers. Recent studies showed that SARS-CoV-2 sequencing can be performed using materials obtained from Ag-RDTs performed at point-of-care34–36. Importantly, whole genomes can be recovered up to eight days after testing, providing opportunities for sequencing to be performed on samples performed through self-testing as well34.”*

1. y-axis label: “No of days since …”. Do you mean “average number of days …” (from XY number of simulations). If so, report error-bars and number of simulations.

We have changed the y-axis label, plotted standard deviations, as well as stated number of simulations in the caption.

1. It may also be legitimate to re-evaluate whether a (compute-intense and difficult-to-analyze) stochastic AB-modelling approach is required.

See response to points 1 and 4 above.

1. The paragraphs the “Emergence of SARS-CoV-2 variants of concern” can be shortened to the material relevant to this work

We opted to keep these paragraphs in the Supplementary Notes as they serve to provide additional, detailed background on how testing and sequencing capacities impacted the detection of other SARS-CoV-2 variants of concern besides Omicron BA.1 in countries where they likely emerged from.

1. You should state early on what kind of diagnostic test you are referring to. As late as on page 7, you state that it is rapid antigen testing.

We corrected this by explicitly stating in the Introduction that rapid antigen tests were the primary SARS-CoV-2 diagnostic tool used in our simulations.

Line 78: “*We assumed that clinic-based professional-use Antigen Rapid Diagnostic Tests (Ag-RDTs) form the basis of testing given persistent reports that polymerase chain reaction (PCR) tests are poorly accessible for detection of symptomatic cases outside of tertiary medical facilities in many LMICs11*”

1. Finddx testing rates report very different things in different countries: In most European countries, for example, the numbers stated largely refer to PCR (and usually not Ag-RDT) tests, which are entering the reporting pathway.

This is correct and we are cognizant of this given that these data are collated by our co-authors from FIND. In fact, this suggests that testing rates from most high-income European countries were likely higher than reported, which further widens the testing equity gap with low- and middle-income countries.

1. Page 15:  
   I think the most informative study on Ag-RDT test sensitivity vs. virus load & time of infection, is actually the human challenge study ([https://doi.org/10.1038/s41591-022-01780-9](https://eur04.safelinks.protection.outlook.com/?url=https%3A%2F%2Fdoi.org%2F10.1038%2Fs41591-022-01780-9&data=05%7C01%7Cx.han%40amsterdamumc.nl%7C7386dbe495eb4cda975108da81e5ef2a%7C68dfab1a11bb4cc6beb528d756984fb6%7C0%7C0%7C637965121794065173%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C3000%7C%7C%7C&sdata=dT6k%2F6sw2doIXoeoEBZbW6ak6GjaCZv7DXlMLjFxpRg%3D&reserved=0)), not reference 18.

We cited and used the Ag-RDT sensitivity parameters provided in reference 18 (now ref. 40; Brümmer et al., 2021) as they were pooled from a meta-analysis of 133 articles that studied the results from >100,000 tests performed on 61 different Ag-RDTs. As for parameters on viral load and periods of infection for wild-type SARS-CoV-2 and variants-of-concern, these were not based on reference 18 but from multiple sources, including Linton et al., 2020, Kang et al., 2021 and Hay et al., 2022, which recruited between 34 and 2,875 infected participants.

However, the parameters inferred from the human challenge study cited by the reviewer was only based on 18 infected individuals. Additionally, the challenge virus was sampled in 2020, genotyped as a B.1 lineage virus with the D614G mutation, which is akin to the wild-type SARS-CoV-2 virus.

Nonetheless, the estimated parameters in this human challenge study are similar to those that were parametrized for the wild-type SARS-CoV-2 virus in our study: (1) The human challenge study found that within-host viral load peaked around 4-6 days after inoculation on average. We assumed that within-host viral load of symptomatic individuals infected with the wild-type SARS-CoV-2 virus would peak upon symptom onset at 5.6 days after infection (Table S1); (2) Median time to first detection by rapid antigen tests in the challenged participants was ~4 days after inoculation. We assumed that the latent phase period of symptomatic individuals to be around 4.5 days on average (Table S1); (3) Viable virus was recovered up to ~8 days after symptom onset in the human challenge study. We also assumed that the recovery period after symptom onset was 8 days on average.

1. You write “…can lead to unreliable estimates …. Regardless of the proportion of samples sequenced.”  
   o This claim is not entirely supported by your results (Fig. 4). At least not in the stated generality.

We have now rephrased this sentence to more accurately represent our results:

Line 26: “…*can lead to unreliable estimates of variant prevalence even when the proportion of samples sequenced is increased.*”