Building toward useful SARS-CoV-2 models in Africa

Belinda Archibong[a] and C. Jessica E. Metcalf[b,c,1]

> All models are wrong; some models are useful.
> George Box

In early 2020, the trajectory of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic on the African continent was much discussed. Slow growth in mortality even prompted speculation that the continent might be "spared" as a result of an overall younger population, effective government responses (1), or perhaps even protection via cross-immunity from other coronaviruses. In 2021, it is increasingly clear that significant and disruptive morbidity and mortality have occurred in at least some settings, and novel variants are of increasing concern. However, it is also clear that understanding the trajectory of the pandemic on the continent requires recognizing vast variation in factors that could shape how bad the infection will be for people who are infected, but also how fast the virus will spread (2). Ssentongo et al. (3) tackle this question using statistical models to model case trajectories across the African continent (Fig. 1*A*), providing week-ahead forecasts of case numbers.

This analysis falls within a rich and active area of research into forecasting within infectious disease biology, at the statistical end of a spectrum that ranges from purely statistical through to fully mechanistic models (4). The development of predictive models hinges on data availability and accuracy, limited in many settings. Ssentongo et al. meet this challenge by leveraging some of the wealth of openly available data that has emerged in the wake of this pandemic, and focusing their analysis at the national scale. Covariates, including policy interventions, climate effects, the human development index, and the context in terms of numbers of cases in neighboring countries, are deployed to predict cases a week ahead. The core goal of the analysis is forecasting, and the design is not intended to dissect underlying causal mechanisms. However, there is much interesting context and many interesting future directions are suggested.

Other coronaviruses (also referred to as "common cold" viruses) tend to peak in winter, likely due to an effect of temperature and/or humidity on their transmission. However, climate drivers will minimally influence the spread of infection when much of the population is still susceptible (2), likely true for SARS-CoV-2 over much of 2020. This, alongside the necessary simplification of aggregating to the scale of countries (as the authors describe), may explain why Ssentongo et al. find that case numbers increase with temperature, a result at odds with experimental data on SARS-CoV-2 (5) and experience with other coronaviruses (6). Nevertheless, variable effects of climatic factors associated with influenza in temperate and tropical regions underscore the need for geographically focused studies (7).

Since viral transmission can only occur between people who are sufficiently close or in poorly ventilated spaces, the human response has played a central role in shaping the peaks and troughs in cases of the pandemic virus. Ssentongo et al. reflect this by using a measure of the stringency of government policies (Fig. 1*A*) provided by the Oxford Blavatnik School of Government (e.g., reflecting school closing, restrictions on gatherings, etc.), as well as an indicator of the testing policy that they provide. While the local contribution to cases declines with testing level, as one might expect if well-functioning isolation systems are in place (8), it increases with the index of stringency. Although it is possible that anticipation of future stringent mobility restrictions might drive population movements that amplify pathogen spread, this counterintuitive finding might also reflect the correlative nature of the analysis. Indeed, many recent findings show that aspects of the stringency index like lockdown measures or school closures (9) do appear to decrease transmission of the disease. Finer-scale analyses and/or investigation of impacts on infections with similar transmission routes [e.g., like measles (2)] could helpfully clarify this.

Inspired by climate science research, infectious diseases modelers are increasingly turning toward "ensemble modeling" for forecasting: Taking the average forecast of multiple models from multiple different

[a]Department of Economics, Barnard College, Columbia University, New York, NY 10027; [b]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544; and [c]School of Public and International Affairs, Princeton University, Princeton, NJ 08540

Author contributions: B.A. and C.J.E.M. wrote the paper.

The authors declare no competing interest.

See companion article, "Pan-African evolution of within- and between-country COVID-19 dynamics," 10.1073/pnas.2026664118.

[1]To whom correspondence may be addressed. Email: cmetcalf@princeton.edu.

Published July 28, 2021.

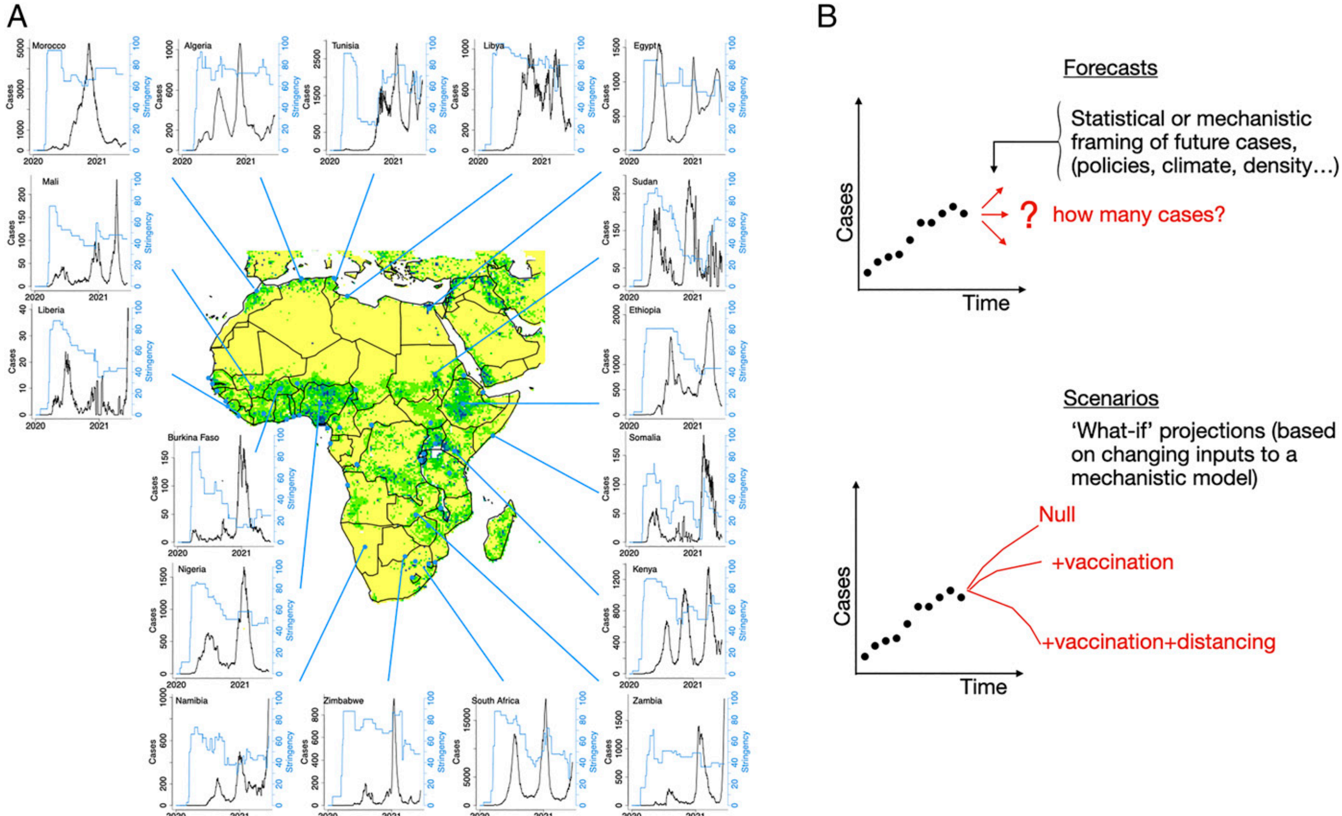

Fig. 1. **Modeling the pandemic on the African continent. (*A*) Time series of cases (black, left axis) and stringency measures (blue, right axis) from exemplar countries (plot titles, and see map, also indicating 2020 population density, from https://www.worldpop.org/). In late June 2021, Uganda, Zambia, and South Africa are experiencing rapid case acceleration. (*B*) Schematic of cases (*y* axis, filled points) over time (*x* axis) illustrating the distinction between forecasting (*Top*), as in Ssentongo et al., where the aim is to forecast the number of cases expected over a chosen time horizon (red arrows show three different forecasts) vs. scenario analysis (*Bottom*), where a mechanistic model is fit to the data, and then the number of cases expected under different control efforts or policies are projected forward by changing the inputs of the model to explore these counterfactuals (here, the schematic illustrates expectations where nothing is done ["null"], or vaccination is introduced, or both vaccination and distancing are introduced). "Ensemble" forecasts or scenario analyses, where multiple models are included, generally improve model performance.**

groups often provides better predictions than any group alone (e.g., https://covid19forecasthub.org). Ssentongo et al. provide a set of forecasts for SARS-CoV-2 on the African continent. As more models are assembled for the continent (10–12), new pathways open to evaluate combinations of predictions to strengthen forecasting nationally and at larger scales.

Forecasts tell us one important part of the policy landscape: Simply knowing whether case numbers are increasing or decreasing (Fig. 1*B*, *Top*) will help public health officials plan the response. However, forecasts do not identify the best policy practices for controlling case numbers. "Scenario analyses" fill this gap (Fig. 1*B*, *Lower*). Scenario analyses compare a range of "what-if" situations, or counterfactuals, asking, for example: What happens to cases if testing doubles relative to if it stays the same? Or if schools are closed? Or if we vaccinate half the population? An integrative forecasting hub (13) encompassing different potential control scenarios (e.g., ref. 14) could help governments evaluate options for control—noting that data quality and the degree to which local and subnational geographies (8, 15) are reflected must be carefully considered in evaluating model results.

As vaccination dose availability continues to be limited on the African continent, and alarming case increases are occurring, possibly associated with novel variants (Fig. 1*A*), any information that can inform policy decisions on the best methods to control the spread of the pandemic has value. There is a growing body of research that finds that not only do epidemics have negative effects on economic development outcomes, worsening numerous indicators from unemployment to gender inequality (16), there are costs (for example to mental health) that are associated with policies like lockdown measures to control the spread of pandemics (17). These costs can be mitigated through properly targeted policy interventions, and a useful predictive model will allow policymakers to effectively factor in these costs to make the optimal decisions in improving societal well-being (17).

Recent papers have highlighted the negative effects of lockdown measures on fiscal revenues of national and subnational governments (18). The reduction in revenues is in concert with increased need for public spending, to mitigate the effects of the pandemic and nonpharmaceutical interventions like lockdown/stay-at-home order measures on vulnerable populations within countries (17–19). While the current scholarly evidence finds that lockdown measures reduce transmission of disease, there is significant heterogeneity in the effects of these lockdown measures on economic outcomes within countries. Some countries, with greater fiscal capacity, are better able to mitigate the costs for populations affected by shutdown enterprises and increased unemployment during the pandemic and lockdown periods, than others.

Constraining pathogen spread via limits on transborder movement may be particularly challenging for countries for which travel is economically important, as Ssentongo et al. suggest, and as genetic evidence of viral mixing between countries indicates (20). Approximately 30% of African countries, 16 out of 54 total, are landlocked countries. These countries had an average per capita gross domestic product of $4,609, 1.5 times less than their nonlandlocked counterparts ($7,061) as of 2019 (in constant 2017 US dollars) by World Bank estimates. Landlocked countries are poorer and more limited in access to trade, especially lucrative maritime trade, than their nonlandlocked counterparts and more likely to rely on open borders for needed trade revenue. Hence, landlocked country governments may face higher costs in attempting to implement lockdown measures while reducing the economic burden of these measures on domestic residents. This is an important fact, highlighted in the paper's finding of positive associations between being landlocked and increased contribution from neighboring countries to SARS-CoV-2 infections. This fact also underscores the necessity of context-dependent policies, as porous borders within Africa pose unique challenges for policymakers seeking to both reduce the spread of disease with lockdown measures, and minimize the costs of both the pandemic and lockdown measures on local populations.

Ssentongo et al. contribute to a growing body of work on the African continent revealing significant heterogeneity in both the transmission patterns of the disease, the accordant policy responses, and the costs associated with those policy responses. This heterogeneity highlights the importance of flexible policy within countries on the continent, and not a one-size-fits-all approach to mitigating the effects of the pandemic. Lessons from past epidemics on the continent, like meningitis epidemics, which are endemic in the African meningitis belt (16), highlight the importance of information sharing and collaborative scientific and policy efforts across countries to combat the spread of infectious disease and minimize the costs of pandemics in the region. Improved predictive modeling, essential for effective policymaking, is only possible with careful data collection and data sharing across countries, and initiatives like the Meningitis Vaccine Project provide a useful blueprint for how this could be done for the SARS-CoV-2 pandemic in Africa.

1 B. Gaye *et al*., Socio-demographic and epidemiological consideration of Africa's COVID-19 response: What is the possible pandemic course? *Nat. Med.* **26**, 996–999 (2020).

2 B. L. Rice *et al*., Variation in SARS-CoV-2 outbreaks across sub-Saharan Africa. *Nat. Med.* **27**, 447–453 (2021).

3 P. Ssentongo *et al*., Pan-African evolution of within- and between-country COVID-19 dynamics. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2026664118 (2021).

4 C. Viboud *et al*.; RAPIDD Ebola Forecasting Challenge Group, The RAPIDD Ebola forecasting challenge: Synthesis and lessons learnt. *Epidemics* **22**, 13–21 (2018).

5 D. H. Morris *et al*., Mechanistic theory predicts the effects of temperature and humidity on inactivation of SARS-CoV-2 and other enveloped viruses. *eLife* **10**, e65902 (2021).

6 S. M. Kissler, C. Tedijanto, E. Goldstein, Y. H. Grad, M. Lipsitch, Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science* **368**, 860–868 (2020).

7 J. D. Tamerius *et al*., Environmental predictors of seasonal influenza epidemics across temperate and tropical climates. *PLoS Pathog.* **9**, e1003194 (2013). Correction in: *PLoS Pathog.* **9**, 10.1371/annotation/df689228-603f-4a40-bfbf-a38b13f88147 (2013).

8 M. Semakula *et al*., The secondary transmission pattern of COVID-19 based on contact tracing in Rwanda. *BMJ Glob. Health* **6**, e004885 (2021).

9 J. Vlachos, E. Hertegård, H. B Svaleryd, The effects of school closures on SARS-CoV-2 among parents and teachers. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2020834118 (2021).

10 S. P. C. Brand *et al*., Forecasting the scale of the COVID-19 epidemic in Kenya. *medRxiv* [Preprint] (2020). https://doi.org/10.1101/2020.04.09.20059865 (Accessed 20 June 2021).

11 J. W. Cabore *et al*., The potential effects of widespread community transmission of SARS-CoV-2 infection in the World Health Organization African Region: A predictive model. *BMJ Glob. Health* **5**, e002647 (2020).

12 E. O. Nsoesie, K. T. L. Sy, O. Oladeji, R. Sefala, B. E. Nichols, Nowcasting and forecasting provincial-level SARS-CoV-2 case positivity using Google search data in South Africa. *medRxiv* [Preprint] (2020). https://doi.org/10.1101/2020.11.04.20226092 (Accessed 20 June 2021).

13 R. K. Borchering *et al*., Modeling of future COVID-19 cases, hospitalizations, and deaths, by vaccination rates and nonpharmaceutical intervention scenarios—United States, April–September 2021. *MMWR Morb. Mortal. Wkly. Rep.* **70**, 719–724 (2021).

14 M. Wagner *et al*., Using contact data to model the impact of contact tracing and physical distancing to control the SARS-CoV-2 outbreak in Kenya. *Wellcome Open Res.* **5**, 212 (2020).

15 S. Abdella *et al*., Prevalence of SARS-CoV-2 in urban and rural Ethiopia: Randomized household serosurveys reveal level of spread during the first wave of the pandemic. *EClinicalMedicine* **35**, 100880 (2021).

16 B. Archibong, F. Annan, Disease and gender gaps in human capital investment: Evidence from Niger's 1986 meningitis epidemic. *Am. Econ. Rev.* **107**, 530–535 (2017).

17 F. Annan, B. Archibong, The value of communication during a pandemic. *SSRN* [Preprint] (2021). https://doi.org/10.2139/ssrn.3772706 (Accessed 20 June 2021).

18 G. Bonaccorsi *et al*., Economic and social consequences of human mobility restrictions under COVID-19. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 15530–15535 (2020).

19 M. Mahmud, E. Riley, Household response to an extreme shock: Evidence on the immediate impact of the Covid-19 lockdown on economic outcomes and well-being in rural Uganda. *World Dev.* **140**, 105318 (2021).

20 J. Lamptey *et al*., Genomic and epidemiological characteristics of SARS-CoV-2 in Africa. *PLoS Negl. Trop. Dis.* **15**, e0009335 (2021).