The importance of local context in COVID-19 models

COVID-19 models have been extensively used to inform public health officials about potential interventions. Nevertheless, careful attention must be taken when extrapolating projections and parameters across different regions, as there is no one-size-fits-all model for the pandemic.

Rosalind M. Eggo, Jeanette Dawa, Adam J. Kucharski and Zulma M. Cucunuba

Infectious disease models are an integral part of public health decision making, and have been crucial tools throughout the COVID-19 pandemic. In early 2020, models estimated the extent of COVID-19 in Wuhan and regional patterns of importations, made projections of potential epidemics and healthcare needs, and provided short-term predictions of case numbers in newly unfolding outbreaks. As countries considered control and mitigation measures, models helped assess likely effects of proposed interventions and determine 'counterfactuals', that is, expected epidemic trajectories if an intervention was not implemented. However, it is important to stress that the quality of such projections depends on within- and between-country variation in transmission, control and burden.

Transmission

The dynamics of directly transmitted respiratory infections depend on multiple factors, including population demography: variation in age and spatial distribution; size and composition of households, schools, and workplaces; and population behavior, often measured as contact rates between age groups and in different settings. Early COVID-19 projections relied on parameter values relating to transmission and disease severity drawn mostly from China — where the epidemic was most advanced and hence most evidence was available — and extrapolated to other cities and countries. Certain factors could be and were modified in projections — for example, by changing modeled age distributions to match country demography[1], and using age-specific mixing rates for the projected country[2] — but other regional or country-level differences are very difficult to extrapolate, an effect compounded by the uncertainties of a new virus and unknown health outcomes.

One example is an epidemic with multiple peaks. These can result if there are strong subdivisions within a population so

that each sub-community has a separate epidemic[3], caused by structure within a city[4], geographics[5] or other factors. National or regional differences in extent or type of community structure may impact projections, and multi-peaked epidemics can also result from interventions. These effects are especially difficult to predict for new pathogens.

Another factor is the frequency and characteristics of different settings that drive transmission[6]. For COVID-19, information has emerged about high-risk settings, such as households, hospitals, congregate settings including long-term care facilities, and overcrowded communities. However, incorporating this level of complexity into transmission models before this information becomes available is challenging. As a result, the dynamics of very detailed models are likely to reflect strong underlying assumptions rather than genuine patterns in (as yet unknown) data.

Severity and impact

To project the extent and impact of COVID-19 in other countries, estimates of severity (for example, age-specific hospitalization and mortality rates) were drawn from early data reported in East Asia[7]. By incorporating these estimates into epidemic models and adjusting for key aspects of demography, these severity estimates have matched experiences in Latin American countries, but there have been settings where to date there is less correspondence between predicted and reported impact, for example in many African countries[8]. Key uncertainties in understanding burden are the extent to which deaths are reported, the prevalence of comorbidities, and how healthcare operates and is utilized. For example, in places where health systems become overwhelmed, higher mortality may result from reduced hospital capacity and quality of care[9]. Such effects could vary from country to country, and even regionally within countries.

Excess mortality data can be useful to better understand deaths and pandemic impacts[10]. However, when using excess or all-cause mortality data, care is needed to understand the context, because interventions such as social distancing and stay-at-home orders could have decreased prevalence of other infectious diseases and thus decreased mortality, and could also have reduced deaths due to other causes (for example, road traffic accidents). Conversely, large epidemics or intensive interventions can disrupt care for non-COVID-19 causes, leading to increased morbidity and mortality that can lag the epidemic or interventions by a considerable period of time, as with deaths from diseases like cancer.

Control

Early responses to SARS-CoV-2 outbreaks varied globally, with many countries implementing lockdown-type interventions, including varying levels of school, work and business closures, stay-at-home orders, curfews and quarantines. In addition, test–trace–isolate strategies and individual-level measures like hand-washing and mask-wearing were promoted to varying extents. Although these interventions appear superficially similar, details of implementation — and the populations they were implemented in — are critical to understanding their effect on transmission and why this differs between countries. In some locations, economic support allowed people to remain at home, while elsewhere, larger informal economies and less support meant more people had to continue working[11]. There have been notable differences across economic strata in ability to adhere to measures like isolation and quarantine across the world, exacerbating health inequalities, as well as creating a more complex epidemiological picture that models may need to include[12]. Even within countries, interventions have been implemented unevenly between regions,

further complicating our understanding of the transmission dynamics.

A major source of uncertainty in modeling, and especially for one-in-a-century pandemics, is understanding specific behavioral responses both to the epidemic and to public health interventions. As awareness of an epidemic rises, and information (and disinformation) spreads, people will inevitably change their normal patterns, for example by decreasing contacts, reducing non-essential travel, and increasing hand-washing, mask-wearing and other hygienic measures. Public health interventions may also be imposed or lifted, and adherence to restrictions can change over the course of the epidemic. These kinds of 'reactive' behavioral responses, especially coupled with intermittent interventions, can give rise to epidemics with multiple peaks. Some early models used different types of interventions introduced at varying times to demonstrate the effect on projected epidemic profiles, but behavioral responses and adherence tend to be very context-specific, limiting generalizability from one area to another. Furthermore, incorporating reactive responses into models is challenging[13] because their magnitude is unknown, difficult to predict, and behaviors may not track information about underlying infection level, but rather news sources, and responses may grow and wane over time. This pandemic has provided new insights into behaviors, adherence to interventions and the role of media that will be critical for planning in future pandemics.

Surveillance data

Judging model predictions against observed surveillance data relies on being able to accurately measure infection outcomes in a population. For all infections there will be under-reporting, and for a new virus this could be considerably worse, resulting from variable case definitions, testing quality and capacity, and testing behaviors. The percentage of cases that are detected (that is, the ascertainment ratio) varies as testing changes, but also is subject to time-dependent and location-dependent variation. Testing may be used for a variety of purposes: to confirm recovery among previously infected individuals; for diagnosis among those acutely ill; as part of routine hospital surveillance; or as part of screening protocols for essential workers or other individuals in the community. Depending on the extent, timing and duration of each approach, a testing strategy could substantially change the ascertainment ratio, and this would need to be incorporated into country-specific models. In the UK, for example, positive COVID-19 cases were

initially detected at hospitalization, but in Kenya, long-distance truck drivers were initially prioritized for screening, before community testing was scaled up in both countries. In Latin American countries, a variety of molecular, immunological and rapid antigen tests are used, making comparisons difficult. Moreover, low- and middle-income countries have low access to widespread diagnostic testing[14]. Models can include an explicit 'observation model', which allows changes in case ascertainment through time and can allow incorporation of multiple data streams with different ascertainment ratios. Fitting models to incomplete or unrepresentative data streams is very challenging, and can result in wide uncertainty bounds if it can be achieved at all. Issues with data availability as well as ascertainment may mean it is difficult to rigorously evaluate forecasts. Population-based seroprevalence studies can overcome some of these difficulties, validate models and make more realistic impact evaluations and projections.

Novel data sources

SARS-CoV-2 emerged into a world more connected and technologically advanced than previous epidemics. New sources of information have been used since early 2020, such as when Baidu made within-China travel volumes live and publicly available to help link human mobility data to control efforts. Similar datasets are available in many countries from apps, mobile phones, and traffic and pedestrian measurements, among others[15]. There are also a range of self-reported monitoring methods, from questionnaires, symptom-reporting apps, and even 'smart' thermometers that can try to provide real-time monitoring[16]. However, technological data, especially those requiring smartphones or personal devices, frequently only measure a non-random segment of the population, often under-representing certain age or socioeconomic groups[17]. In interpreting these data and incorporating them into models, scientists must address this differential representativeness.

These biases are not only an issue for passively collected data, but also for intervention systems that use smartphones, such as digital contact tracing. There is widely differential take-up of these apps even where they are available, either due to trust and privacy concerns about use of the data, or due to the technical requirements for devices that can support them. Therefore, the utility of these data for analysis or epidemic mitigation may be limited, particularly in low-middle income countries, where smartphone data may be particularly unrepresentative of the overall

population. Despite these limitations, there are some examples of integration between smartphone data and epidemiological models that have provided insights into social heterogeneities in infection and potential ways to address risk disparities[18].

Outlook

It is unrealistic to expect that models will be able to include a range of setting-specific intricacies and complexities at the start of a pandemic of a newly emerged virus. What models are most useful for and what insights they can provide changes as an epidemic proceeds: early projections may need to be a 'reasonable worst case scenario', for which broad conclusions are most important, and later, as policy questions or needs become more specific, precise mechanisms can be added when there are data to support it. Nevertheless, modelers should strive to convey the uncertainties and to understand and explain how variable the outputs may be when as-yet unestablished parameters of the model are changed. Developing, refining and using models in real-time during a pandemic can also reveal key areas of uncertainty about specific aspects of transmission, providing additional understanding about which assumptions need revisiting, or locations or time periods within which projections are poor.

It is crucial that modeling groups continue to refine their models using the most up-to-date data from observational and interventional studies given the likely duration of the COVID-19 pandemic, potential for further interventions, and the need to plan for vaccines. It is important to tailor models to the population and context under study — as much as is possible and feasible given the aim of the model — in order to provide more accurate and relevant local estimates, both for this pandemic, and for those in the future. ❐

Rosalind M. Eggo[1 ✉], Jeanette Dawa [ID][2,3], Adam J. Kucharski[1] and Zulma M. Cucunuba [ID][4,5]

¹Centre for Mathematical Modelling of Infectious Diseases, London School of Hygiene and Tropical Medicine, London, UK. ²College of Health Sciences, University of Nairobi, Nairobi, Kenya. ³Washington State University, Global Health Programs, Nairobi, Kenya. ⁴MRC Centre for Global Infectious Disease Analysis, School of Public Health, Imperial College London, London, UK. ⁵Faculty of Medicine, Pontificia Universidad Javeriana, Bogotá, Colombia.
✉e-mail: r.eggo@lshtm.ac.uk

References

1. Davies, N. G. et al. *Nat. Med.* **26**, 1205–1211 (2020).
2. Prem, K. et al. Preprint at https://doi.org/10.1101/2020.07.22.20159772 (2020).
3. Watts, D. J, Muhamad, R., Medina, D. C. & Dodds, P. S. *Proc. Natl Acad. Sci. USA* **102**, 11157–11162 (2005).
4. Hoen, A. G. et al. *J. Med. Internet Res.* **17**, e169 (2015).
5. Gog, J. R. et al. *PLoS Comput. Biol.* **10**, e1003635 (2014).
6. Leclerc, Q. J. et al. *Wellcome Open Res.* **5**, 83 (2020).
7. Verity, R. et al. *Lancet Infect. Dis.* **20**, 669–677 (2020).
8. Musa, H. H. et al. *Int. J. Infect. Dis.* **102**, 483–488 (2020).
9. Ñamendys-Silva, S. A., Gutiérrez-Villaseñor, A. & Romero-González, J. P. *Intensive Care Med.* **46**, 2086–2088 (2020).
10. Beaney, T. et al. *J. R. Soc. Med.* **113**, 329–334 (2020).
11. Arnold, J., Garda, P. & Gonzalez-Pandiella, A. *Ecoscope* https://oecdecoscope.blog/2020/06/29/reaching-out-to-informal-workers-in-latin-america-lessons-from-covid-19/ (2020).
12. van Dorn, A., Cooney, R. E. & Sabin, M. L. *Lancet* **395**, 1243–1244 (2020).
13. Verelst, F., Willem, L. & Beutels, P. *J. R. Soc. Interface* **13**, 20160820 (2016).
14. Faust, L. et al. *Microbes Infect.* **22**, 511–514 (2020).
15. Buckee, C. O. et al. *Science* **368**, 145–146 (2020).
16. Whitelaw, S., Mamas, M. A., Topol, E. & Van Spall, H. G. C. *Lancet Digit. Health* **2**, e435–e440 (2020).
17. Grantz, K. H. et al. *Nat. Commun.* **11**, 4961 (2020).
18. Chang, S. et al. *Nature* https://doi.org/10.1038/s41586-020-2923-3 (2020).