
EDITORIALS AND COMMENTARIES

Demystifying COVID-19 pandemic modelling for policymaking

C Raina MacIntyre¹ & David James Heslop¹

¹ University of New South Wales, Australia

Abstract

Mathematical modelling is used widely to inform COVID-19 pandemic policy. Infectious diseases modelling is a long-established science used to estimate future outcomes under various conditions, that can inform policy decisions. Each model depends on assumptions made, the specific modelling methods used and the scenarios explored. Non-modellers can evaluate models using the following principles.

Key words: Mathematical modelling, epidemics, policy

Mathematical modelling is used widely to inform COVID-19 pandemic policy. Infectious diseases modelling is a long-established science used to estimate future outcomes under various conditions, that can inform policy decisions. Each model depends on assumptions made, the specific modelling methods used and the scenarios explored. Non-modellers can evaluate models using the following principles.

When is modelling needed?

Modelling may be the only means of answering some research questions in a rigorous, scientific way and providing forecasts, scenario analysis and evidence based policy advice in near real time. ¹ The randomised controlled clinical trial (RCT) is thought of as the highest level of evidence, but the RCT is appropriate for questions about treatments or interventions. In the case of pandemic planning, however, the question may be around a hypothetical event that has not yet occurred or around interventions that cannot be tested in a RCT. For example, international border closures cannot be tested using a RCT. For some questions, a prospective study is often not feasible if future outcomes or complex interactions between interventions or future scenarios in whole populations are of interest. Policy makers can use modelling to compare a mitigated and unmitigated epidemic of SARS-COV-2, and determine what measures may best mitigate it. A model can look at combined interventions such as vaccines, masks, movement restriction, contact tracing and testing. If lockdowns are undesirable, a model can inform what other, less restrictive measures can control a potential surge.

Modelling of epidemics

Models can look at epidemic, endemic or sporadic infections, and can also be used to look at animal or insect to human transmission, and sexually transmitted infections. Here, we focus on the SARS-COV-2 pandemic which is spread from person to person by the airborne route, which is the most rapid spread with the greatest

epidemic potential. The way epidemics spread depends on the pathogen, the population, how humans behave, mix and move around, and on public health measures. The terms epidemic and endemic are widely misused, but refers to a specific pattern of infection (Box 1).

Epidemic infections are uniquely suited to modelling because humans exist in mutually exclusive states of susceptibility, exposure, infection or immunity. Immunity results from natural infection or vaccination, and immunity can wane, returning people to the susceptible state. A model allocates people to these states, and mathematically determines how they will transition between states - called a Susceptible, Exposed, Infected and Recovered (SEIR) model. These can be very simple or complex, with multiple added compartments such as contact tracing, testing, quarantine, hospitalisation or death. A model can also have parallel streams to account for differences between groups, such as vaccinated and unvaccinated people.

Policy makers can look at the effect of reducing movement and mixing (such as through density limits or a lockdown), test and trace, use masks or vaccinate people, and how this reduces the spread of infection. Modelling allows us to simulate these effects for entire populations and determine how infections may spread in a range of scenarios with and without single or combination interventions. It is a science that allows us to systematically address future spread and compare different disease control options on a population level, with a realistic, simulated, age structured population. For example, if we lift movement restrictions after achieving 70% vaccination of adults, can indoor mask use prevent epidemic growth? SEIR modelling is most commonly used because it is computationally less intensive and can provide rapid outputs to support decision-making during a fast moving COVID-19 surge, such as that occurring now with the Omicron variant of concern.

Box 1. Patterns of infectious diseases

Endemic: a disease that exists permanently in a particular region or population. E.g. Malaria, diabetes. Changes in incidence occur over years, if at all.

Epidemic: An increase in disease that affects many peoples at about the same time in a given region and may spread through one or several communities. It is defined by rate of growth, typically days or weeks, and by the basic reproductive number, R_0 , the average number of secondary cases caused by an infected person in a completely susceptible population. The effective R (R_t) is the R value when immunity or interventions are present to reduce R_0 . If $R > 1$, the number of cases increases (an epidemic may occur); if $R < 1$, the number of cases decreases (infection cannot be sustained and dies out). Therefore $R=1$ is the epidemic threshold. Epidemic infections (in contrast to endemic infections such as malaria) can grow exponentially over days or weeks, which is why epidemic and pandemic diseases require immediate surge capacity in health systems. SARS-COV-2 is a typical epidemic infection, and will always retain this pattern, much like measles.

Pandemic: An epidemic that spreads throughout the world.

Sporadic: Few cases of disease that do not meet definitions above. Typically these are zoonotic infections such as human cases of avian influenza.

There are several different modelling approaches other than SEIR modelling that have been used for SARS-COV-2. Agent-Based Modelling (ABM) is the most common alternative. In an ABM autonomously operating and interacting software entities representing abstracted humans and other artifacts in a system interact over time in a synthetic environment. ABMs are useful when individual behaviour matters, or when you want to examine more complex interactivity such as transmission inside households, on public transport, or the influence of individual decision-making on outcomes. For highly transmissible respiratory epidemic infections, if the question is about population level impacts, the use of a SEIR or ABM does not make a large difference. ABM can be helpful to examine interactions at the micro and meso level within population subgroupings. A barrier to more widespread ABM utilisation is model validation and computational demands of scaling up to large populations.² There are many other modelling methods including network models, with transport network modelling useful to understand how infections can spread through air travel or other travel modes. Mobile phone data can also be used to estimate reduction of movement during the pandemic.³ Economic modelling can utilise the outputs of disease modelling to test the cost effectiveness of competing options in disease control. Geospatial modelling can be used to identify hot spots for infections that vary by location, climate and environmental factors, and phylogenetic modelling can inform control of emerging infections.

Uncertainty and sensitivity to assumptions

Models have inherent uncertainty - where true values of given parameters may not be definitive because data

are not available or different studies provide varying estimates. All the factors that influence future spread of the pandemic are unknown and may change over time. For example, a model created in August 2021 may not include vaccination of children, because it is not yet approved at that time. In the subsequent months, vaccination of children may be introduced, further improving epidemic control and making the model appear to over-estimate case numbers. A model for the Delta variant created in late 2021 will not be appropriate for the Omicron variant, which is newly emerged and has greater uncertainty around model parameters because fewer published studies are available. Usually models test a series of interventions at one time point or simultaneously (such as vaccination, masks and movement reduction). These may change in different combinations due to government response to policy. Sometimes the data to inform an assumption in a model may not be available, or there may be a series of studies with different estimates. For example, when Delta first emerged, estimates of R_0 ranged from 4-8. Estimates of R_0 for Omicron are less certain because there are fewer studies, but several estimates are much higher than Delta.

It is important that any published model outlines which parameters and assumptions are uncertain. Authors should also state which assumptions the outputs are most sensitive to. That is, assumptions that create large changes in model outputs if the values vary. An example may be the infectious period. There may be large differences in model outputs if this is assumed to be 3 days versus 10 days, with 10 days resulting in more transmission. It may be the R_0 – a model that uses the R_0 for the Delta variant will under-estimate the impact of the Omicron pandemic. Uncertainty should be

addressed with a sensitivity analysis of uncertain and/or highly influential parameters – this is done by generating model outputs for a range of possible values for an uncertain parameter to determine the influence of varied estimates.

Dire “predictions” that never eventuate

Models do not make hard and fast “predictions”. They only provide a range of possible scenarios of pandemic spread under different conditions. We often hear that models make incorrect and alarmist “predictions”, and usually this refers to modelling of the worst case scenario of unmitigated spread. Pandemic planning for policymakers is necessarily based on worst case scenarios because the aim is to prevent that scenario ever occurring. Good modelling should produce worst and best-case scenarios to inform optimal decision making and ensure resources for health system surge capacity are adequate. A modelled worst-case scenario rarely occurs because governments act to introduce public health measures when cases start rising, when warned by modelling or when the health system is overloaded and starts failing.

Model validation

A good model should provide some validation. This can be done by fitting the model to observed data, such as notified cases of COVID-19. If the fit is not good, the model can be revised by changing different assumptions until the fit is good. All influential conditions need to be incorporated into a model to fit it well – the degree of reduction of human movement and mixing, testing, tracing, mask wearing, vaccination and any other interventions that will affect spread. If fitting is done without incorporating all such factors, the model will overestimate the impact of included interventions.

How do you assess a model?

To evaluate a model, look at the team, the assumptions used about the pathogen, the population and the interventions being tested, and methodological issues such as uncertainty. The EPIFORGE checklist is recommended to evaluate a modelling study. ⁴ In addition, the following principles can be used to assess the model.

The team

A multidisciplinary team with modellers, medical experts and people with field and other experience is ideal (but not essential) for modelling studies, to ensure that assumptions used in the model are correct.

General methodology

Have the aims clearly been stated? Is the type of modelling approach specified, justified and referenced? Is the model design, verification and validation described? A model schematic or diagram and table of parameters in the model with references for data sources should be provided, as well as differential equations (for

SEIR models) used in the model. The start and end date for the modelled outputs is required and should be justified. The statistical or modelling methods should be appropriate for the research question and referenced. The limitations, biases, uncertainty and their management should be described and addressed. Ideally the model code should be available.

Population

Have the population characteristics, such as age structure, been adequately reflected in the model? This is important because COVID-19 case fatality is highly age dependent, so a country with an ageing population will be more severely affected than a country with a young population. How does the model factor in the mixing of people? Some models assume homogenous mixing, but this is not realistic, because mixing is variable. In SEIR modelling, a who-acquires-infection-from-who matrix may be used, which estimates the degree of contact between people by age group. ⁵

Pathogen

Have the researchers correctly estimated disease parameters such as the incubation period, serial interval and duration of viral shedding? The incubation period matters, as infections with a short incubation period cause epidemic growth more rapidly. Do they assume viral shedding to be a constant function over the infectious period, or use an exponential decline (the latter being more correct for SARS-COV-2 ⁶)? Do they correctly factor in the proportion of asymptomatic and pre-symptomatic transmission (which may influence epidemic growth)? ⁷ Characteristics including the R_0 and incubation period have changed over time, so it is essential that models use data for current variants of SARS-COV-2.

The basic reproductive number, R_0 , can be estimated in some models, while other models use a published estimate of R_0 . R_0 is the number of secondary cases generated from one index case. The lower the R_0 , the easier it is to control a disease. R_0 is determined by characteristics of the organism (infectivity, duration of infectiousness, incubation period, asymptomatic transmission) and population characteristics (demographics, social mixing patterns, population density). Herd immunity is when the entire population is protected, whether they have been vaccinated or not, because the number of susceptible individuals is too small for infection to spread. The population immunity needed for herd immunity can be calculated using R_0 . ⁸ The higher R_0 for Delta⁹ means a higher herd immunity threshold for COVID-19 control. Herd immunity also depends on vaccine efficacy, duration of efficacy and population vaccination rates. ⁷ The early promise of vaccines returning us to “normal” life did not eventuate because few countries have achieved vaccine coverage above the herd immunity threshold for Delta, and countries that have come close have encountered waning vaccine immunity after two doses.

Table 1. Guide to reading modelling studies.

Area to review	Questions
Research question	<ul style="list-style-type: none"> • Is clearly stated? • Appropriate for modelling? • What is the best model methodology for the question?
Research Team	Multidisciplinary expertise? <ul style="list-style-type: none"> • Medical • Virology • Mathematical • Social science • Other relevant disciplines
Methodology	<ul style="list-style-type: none"> • Type of modelling and simulation approach is specified (and referenced) • Model design, verification and validation is described • Statistical methods are appropriate (and referenced) • Limitations/biases and their management are described
Population	Is the population realistic? <ul style="list-style-type: none"> • Age structured • Other demographics • Heterogenous mixing • Population density
Pathogen	Assumptions relevant for dominant variant of concern (VOC) and referenced? <ul style="list-style-type: none"> • R_0 • Incubation period • Latent period • Serial Interval • Duration of viral shedding • Infectiousness function (constant or dynamic)
Interventions	Vaccine efficacy (RCT) or effectiveness (observational) RCT data relevant to dominant VOC and referenced? Are influential real-world interventions accounted for, how is effectiveness estimated and referenced? <ul style="list-style-type: none"> • Masks • Testing and case isolation • Tracing and quarantine of contacts • Reduction in movement (density limits, lockdowns) • Social distancing • Border control
Uncertainty	<ul style="list-style-type: none"> • Have the most influential parameters been identified? • Is there high or low certainty of the values of these? • How certain are assumptions about these (data driven or estimated)? • Has a sensitivity analysis been done for the most influential parameters and parameters that are uncertain?
Other	<ul style="list-style-type: none"> • What is the modelled setting? • Whole population • Population subgroup • Indoors/Household • Outdoors • Transport • Other • Is the geography realistic? • Climate • Weather and airflow • Topography • Animal/vector data

Interventions

What scientific evidence was used to inform interventions tested in the model? The source of vaccine efficacy estimates should be checked, noting that efficacy varies by vaccine. A model of the Delta variant will over-estimate efficacy by using phase 3 clinical trial data from 2020 or other data prior to Delta circulating. Another common way that models over-estimate vaccine impact is by failing to model waning vaccine-induced immunity after two doses.

Models may account for real-world interventions such as testing, tracing, mask wearing and movement reduction, for the most realistic output. Data on effectiveness of non-pharmaceutical interventions is more contentious than RCT data on vaccine efficacy, so should be carefully scrutinised. For example, mask use is a heterogeneous intervention, ranging from low quality cloth masks to N95 respirators, and few studies have estimated the effectiveness of masks in a robust way. The effectiveness of movement reduction is also difficult to quantify, but credible estimates are available.¹⁰

Uncertainty and sensitivity analysis

Have the authors addressed uncertainty? Have they identified the most influential parameters and then done a sensitivity analysis to test the impact of changes in assumptions about those parameters? This is important to demonstrate the uncertainty in the model. A good example is the R_0 . If a new variant of concern emerges and the R_0 is not yet certain, they should use a range of plausible estimates in a sensitivity analysis. The model should also be validated in some way, such as fitting to observed data.

Table 1 provides a guide to evaluating a model.

Conclusion

In summary, good modelling should be multidisciplinary and transparent in clearly outlining the methods, assumptions used and data sources. It should realistically approximate the population, human behaviour, the pathogen and background interventions (such as masks, testing and tracing) that may affect the apparent impact of the intervention of interest (e.g. vaccines). A transparent model can be reproduced by other modellers and produce the same answers. A basic understanding of modelling allows any key stakeholder, regardless of mathematical skills, to read, interpret and

appraise a modelling study using the principles outlined above.

References

1. Anderson, R. M. May, R.M. Directly transmitted infections diseases: control by vaccination. *Science*. **215**, 1053–1060. (1982).
2. Heslop, D. J. et al. Publicly available software tools for decision-makers during an emergent epidemic-Systematic evaluation of utility and usability. *Epidemics* **21**, 1-12, doi:10.1016/j.epidem.2017.04.002 (2017).
3. Badr, H. S. et al. Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study. *Lancet Infect Dis* **20**, 1247-1254, doi:10.1016/s1473-3099(20)30553-3 (2020).
4. Pollett, S. et al. Recommended reporting items for epidemic forecasting and prediction research: The EPIFORGE 2020 guidelines. *PLoS Med* **18**, e1003793, doi:10.1371/journal.pmed.1003793 (2021).
5. Prem, K. A.-et al. Projecting social contact matrices in 152 countries using contact surveys and demographic data. *PLOS Computational Biology* **13**, 1553-7358 (2017).
6. He, X. et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature Medicine* **26**, 672-675, doi:10.1038/s41591-020-0869-5 (2020).
7. MacIntyre, C. R. et al. Modelling of COVID-19 vaccination strategies and herd immunity, in scenarios of limited and full vaccine supply in NSW, Australia. *Vaccine*, doi:10.1016/j.vaccine.2021.04.042 (2021).
8. MacIntyre, R. Vaccination for COVID-19 control and considerations for Australia. *Microbiology Australia* (2021).
9. Liu, Y. et al. The reproductive number of the Delta variant of SARS-CoV-2 is far higher compared to the ancestral SARS-CoV-2 virus. *J Travel Med*, doi:10.1093/jtm/taab124 (2021).
10. Petherick, A. et al. A worldwide assessment of changes in adherence to COVID-19 protective behaviours and hypothesized pandemic fatigue. *Nat Hum Behav* **5**, 1145-1160, doi:10.1038/s41562-021-01181-x (2021).

How to cite this article: MacIntyre CR & Heslop DJ. Demystifying COVID-19 pandemic modelling for policymaking. *Global Biosecurity*, 2022; 4(1).

Published: February 2022

Copyright: Copyright © 2022 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

Global Biosecurity is a peer-reviewed open access journal published by University of New South Wales.