HAL open science

Epidemic models: why and how to use them

Mircea T Sofonea, Simon Cauchemez, Pierre-Yves Boëlle

To cite this version:

Mircea T Sofonea, Simon Cauchemez, Pierre-Yves Boëlle. Epidemic models: why and how to use them. Anaesthesia Critical Care & Pain Medicine, 2022, 41 (2), pp.101048. 10.1016/j.accpm.2022.101048 . hal-03638186

HAL Id: hal-03638186

https://hal.science/hal-03638186v1

Submitted on 12 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

COVID-19 ACCPM Series

Epidemic models: why and how to use them

Mircea T. Sofonea[1,*], Simon Cauchemez[3], Pierre-Yves Boëlle[3]

1: MIVEGEC, Université de Montpellier, CNRS, IRD - Montpellier, France
2: Mathematical Modelling of Infectious Diseases Unit, Institut Pasteur, Université de Paris, UMR2000, CNRS, Paris, France
3 : Institut Pierre Louis d'Epidémiologie et de Santé Publique, Sorbonne Université, INSERM, Paris, France

* corresponding author: mircea.sofonea@umontpellier.fr

ORCID: MTS - 0000-0002-4499-0435, SC - 0000-0001-9186-4549, PYB - 0000-0002-5367-8232.

keywords: Infectious disease modelling, mathematical epidemiology, public health, non-pharmaceutical interventions, COVID-19.

word count: 1631

Acknowledgements

The authors thank the members of the Coordinated Action of the ANRS|MIE *Modelling of Infectious Diseases* for discussion.

Competing Interests

The authors declare no competing interest.

The uncontrolled rise in the incidence of SARS-CoV-2 has prompted the adoption of mitigation measures of an unprecedented magnitude worldwide since March 2020. The dire perspective of an acute infectious disease saturating the health system -- an unexpected situation in developed countries with a healthcare system focusing on chronic diseases -- was highlighted in a report made public by Imperial College scientists at the end of February 2020: Using “epidemic modelling”, they announced the overwhelming of acute care capacities and the need for long term and harsh social distancing measures to fight the disease(1).

Throughout the pandemic, epidemic modelling has been used to characterize risks associated with COVID-19, forecast impact on healthcare(2), support and justify the adoption of measures(3), guide the optimal deployment of vaccines(4), and many aspects of our lives with the virus(5). Yet, given the urgency and the issues at stake in the last 2 years, attention has been mostly focused on the results of modelling, with little opportunity to explain the methodology behind model building. We aim to fill this gap in this short editorial.

“Epidemic modelling” describes a set of approaches where mathematical, statistical, and computational tools are used to study the spread of communicable pathogens in host populations. It uses data and hypotheses describing the demographic processes, environmental characteristics, transmission opportunities, and health consequences of diseases(6). Equipped with suitable equations for evolution, it can be used for logical verification with conceptual “what if…” experiments, quantification, and conjecture based on the construction of forecasts and scenarios(7). As in other scientific domains, these models aim at providing a simplified representation of a real phenomenon, focusing on the subset of properties and processes considered as essential drivers. The construction of these models is therefore the result of structural choices motivated by previous knowledge, data and observation and a set of assumptions based on our understanding of the phenomenon under study.

There are various reasons to model epidemics. First of all, measuring the effect of interventions by direct observation is seldom possible. Contrary to drugs that can be evaluated at the patient level and compared between groups of patients, interventions in epidemics can only be assessed at the population level. Setting aside the difficulty of measuring incidence at the population level, this implies that repeated observation of the same population under different regimes is not possible. In this context, modelling is the only way to compare potential scenarios for evolution. A second difficulty is the non-linear nature of epidemic dynamics(8) and the importance of their stochastic components (beginning and end of a wave, super-propagation)(9) the trajectory of an epidemic cannot be predicted by simple proportionality rules and more sophisticated computations, as performed in models, are required. A third challenge is that unbiased descriptors of an epidemic would require complete knowledge of the whole population at all times, when observation is at best partial. Take for example the positive predictive value of a PCR test: it is well known that this quantity is dependent on the prevalence of the infection in the population as well as on characteristics of the test. However, the prevalence of the disease is generally unknown because it cannot be measured in real-time. A statistical model may be used to reconstruct prevalence using more readily available incidence data and information on the incubation time of the disease. Finally, modelling allows accounting simultaneously for the uncertainty on all parameters describing the disease and its spread, especially for the computation of prediction intervals: while it is relatively easy to produce a point estimate, quantifying its robustness requires advanced analyses(10–12).

Modelling can be used to retrospectively understand the past course of an epidemic, quantify the current situation, or anticipate the future course under different scenarios. Although the same model may be able to answer several of these related questions, it will generally require significant adaptations to accommodate different time and space scales, as

well as changes in the environment and behaviour. Furthermore, considering the potential importance of stated (or sometimes unrecognised) hypotheses in model outputs, it is better if policymakers can be informed by a set of congruent results from independent and, when possible, methodologically complementary models, weighted according to their fit to present or past data, as illustrated by Figure 1, reinforcing shared conclusions and highlighting discrepancies.

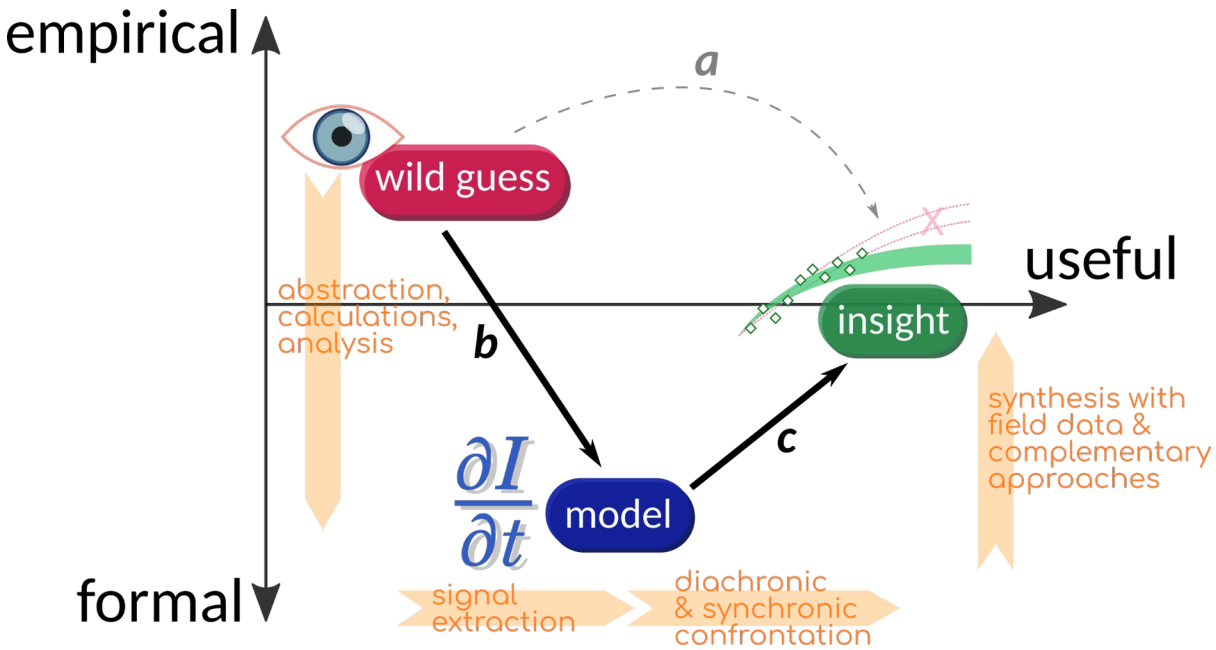

Fig.1. Models as necessary steps between intuition and insight.
Describing a new phenomenon would ideally turn empirical observations and intuition directly into firm knowledge. In the case of an epidemic, such a path (*a*) is difficult to follow because of numerous ethical, spatio-temporal, logistic limitations or concurrent processes. Modelling provides a principled path to extracting knowledge from observational data and hypotheses (step *b*) allowing quantitative manipulation in a formalised space. Longitudinal comparisons (model performance in the past) and transversal comparisons (here/there situations) with field data as well as systematic model explorations (sensitivity analyses) allow us to finally derive robust knowledge (step *c*).

Modelling in infectious disease epidemiology today relies on a variety of methods based on decades of developments at the interface of demography, medicine, and biology. The vast majority of models adopt a “compartmental” description of the disease as a sequence of different stages encountered upon infection to cure or death(6). The paradigmatic example is the “S-I-R” model describing individuals as first susceptible to infection (S), then infected with the disease and contagious (I), and finally removed from transmission by recovery or deaths (R). In such models, the incidence is described as both proportional to the number of susceptible and the number of infected individuals, leading to the non-linear equation for evolution(13).

The solution of these models can be obtained by (partial) differential equations. This approach has produced numerous theoretical and numerical tools. For example, this includes how to compute the reproduction number as a function of other parameters in the model. It is also highly flexible, allowing refining the natural history with several stages, partitioning the population according to relevant characteristics (age, sex, region,...) These models provide a parsimonious (i.e. using few parameters) means to understand the nontrivial behaviour of epidemic pathways. Despite their structural simplicity, experience has shown that such models generate robust results that reinforce their usefulness, even if only at the outset of an epidemic(14). The same models can also be solved by accounting for stochastic effects, for example in the case of large inter-individual variability and one wishes to quantify the impact of this heterogeneity on epidemiologic dynamics, especially in contexts with small numbers of cases (epidemic initiation or extinction, local spread). The logic of these compartmental models is to subdivide the population under study by categories with equivalent epidemiological contribution (by infection status, by age, sex ...). However, this approach does not scale efficiently when the number of relevant characteristics increases. In this case, an “individual-based model (IBM)” solution may be more appropriate. In this approach, mathematical equations for the average population are replaced by individual rules of change that will be solved by computer simulation. Each individual in the population and its interaction with others can be described in relevant detail and rules refined based on these characteristics. IBMs are the preferred approach to explore scenarios that take into account detailed spatial and behavioural mechanisms or put the emphasis on logistics. This collection of approaches makes it possible to study epidemics at various scales: while not all aspects of the COVID-19 pandemic can be covered by a single model, a suite of modelling approaches allows examining all relevant questions.

As always, models will benefit from being based on good data: hence collecting comprehensive, independent, and unbiased data in the early weeks of the spread of an emerging infectious disease, or a new variant, is a major challenge. E.g., in the initial reconstruction of the course of the COVID-19 epidemic in the French population, testing data were too limited while signal extraction from hospital data required external age-stratified infection fatality ratios(10). Importantly, models can be used for short term extrapolation (“forecasting”), from a few days to a few weeks in the case of COVID-19, because of “epidemic inertia”, i.e. the fact that most cases likely to occur or be hospitalized soon are determined by the current incidence. It is a totally different task to use models to explore possible middle or long term scenarios of evolution (projections) as this will rely more heavily on assumptions for change that are widely uncertain, as the successive selection of SARS-CoV-2 variants with changing characteristics illustrate.

The relevance of a model requires confrontation with real data and critical assessment of the underlying hypotheses. A mismatch between the data and the model's projections may challenge the model's assumptions and provide an opportunity to improve it. Importantly, modelling allows exploration of scenarios that could not be amenable to direct experimentation but does not guarantee that such scenarios are desirable or feasible.
Given necessary simplifications, an epidemiological model will have shortcomings that must be looked for, acknowledged, discussed, and alleviated as new data become available. A last, and sometimes mocked, feature of model-based scenario explorations is that they produce “what if…” explorations that can induce change in the very dynamics they are modelling. Unlike in meteorology, where physical processes are unaffected by the prediction of weather models, epidemiological model results have the potential to change the future course of the epidemic. For example, adopting a lockdown based simulated scenario comparisons makes the “no lockdown” strategy counterfactual and voids predictions based on this hypothesis.

The rapidly conducted work since the onset of the COVID-19 pandemic helped provide valuable tools for epidemiological surveillance and health forecasting. Modelling has proved a framework where different fields can more easily communicate through explicit hypotheses and data, all useful qualities that wild guesses and verbal models lack. Claude Bernard was not a supporter of the “numerical approach” in the medical sciences. However, in 1865, he wrote(15) "it is not that I condemn the application of mathematical application in biological phenomena, because it is by it alone that, in science will be constituted; only I have the conviction that the general equation is impossible for the moment, the qualitative study of the phenomena having necessarily to precede their quantitative study." After almost one century and a half of parallel advances in both empirical and modelling approaches to infectious diseases, the COVID-19 pandemic undoubtedly constitutes the opportunity for the junction predicted by Claude Bernard to finally happen.

Références

1. Ferguson N, Laydon D, Nedjati Gilani G, Imai N, Ainslie K, Baguelin M, et al. Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand [Internet]. 2020 mars [cité 13 nov 2020]. Disponible sur: http://spiral.imperial.ac.uk/handle/10044/1/77482
2. Paireau J, Andronico A, Hozé N, Layan M, Crepey P, Roumagnac A, et al. An ensemble model based on early predictors to forecast COVID-19 healthcare demand in France [Internet]. 2021 [cité 22 avr 2021]. Disponible sur: https://hal-pasteur.archives-ouvertes.fr/pasteur-03149082
3. Adam D. Special report: The simulations driving the world’s response to COVID-19. Nature [Internet]. 2 avr 2020 [cité 16 juin 2020];580(7803):316-8. Disponible sur: https://www.nature.com/articles/d41586-020-01003-6
4. Kiem CT, Massonnaud C, Levy-Bruhl D, Poletto C, Colizza V, Bosetti P, et al. Short and medium-term challenges for COVID-19 vaccination: from prioritisation to the relaxation of measures. :45.
5. Haug N, Geyrhofer L, Londei A, Dervic E, Desvars-Larrive A, Loreto V, et al. Ranking the effectiveness of worldwide COVID-19 government interventions. Nat Hum Behav [Internet]. déc 2020 [cité 3 mars 2021];4(12):1303-12. Disponible sur: http://www.nature.com/articles/s41562-020-01009-0
6. Keeling MJ, Rohani P. Modeling infectious diseases in humans and animals. Princeton

University Press; 2008.
7. Djidjou-Demasse R, Selinger C, Sofonea MT. Épidémiologie mathématique et modélisation de la pandémie de Covid-19 : enjeux et diversité. Rev Francoph Lab [Internet]. 1 nov 2020 [cité 19 avr 2021];2020(526):63-9. Disponible sur: https://www.sciencedirect.com/science/article/pii/S1773035X20303154
8. Levy MR, Tasoff J. Exponential-growth bias and overconfidence. J Econ Psychol [Internet]. févr 2017 [cité 20 déc 2021];58:1-14. Disponible sur: https://linkinghub.elsevier.com/retrieve/pii/S0167487016306596
9. Beneteau T, Elie B, Sofonea MT, Alizon S. Estimating dates of origin and end of COVID-19 epidemics. Peer Community J [Internet]. 2021 [cité 19 déc 2021];1. Disponible sur: https://peercommunityjournal.org/articles/10.24072/pcjournal.63/
10. Salje H, Tran Kiem C, Lefrancq N, Courtejoie N, Bosetti P, Paireau J, et al. Estimating the burden of SARS-CoV-2 in France. Science [Internet]. 13 mai 2020 [cité 14 mai 2020];eabc3517. Disponible sur: https://www.sciencemag.org/lookup/doi/10.1126/science.abc3517
11. Sofonea MT, Reyné B, Elie B, Djidjou-Demasse R, Selinger C, Michalakis Y, et al. Memory is key in capturing COVID-19 epidemiological dynamics. Epidemics [Internet]. 1 juin 2021 [cité 19 mai 2021];35:100459. Disponible sur: https://www.sciencedirect.com/science/article/pii/S1755436521000189
12. Zelner J, Riou J, Etzioni R, Gelman A. Accounting for uncertainty during a pandemic. Patterns [Internet]. 13 août 2021 [cité 12 oct 2021];2(8). Disponible sur: https://www.cell.com/patterns/abstract/S2666-3899(21)00153-7
13. Reyné B, Saby N, Sofonea MT. Principles of mathematical epidemiology and compartmental modelling application to COVID-19. Anaesth Crit Care Pain Med [Internet]. 28 déc 2021 [cité 2 janv 2022];101017. Disponible sur: https://www.sciencedirect.com/science/article/pii/S2352556821002241
14. Trapman P, Ball F, Dhersin J-S, Tran VC, Wallinga J, Britton T. Inferring R_0 in emerging epidemics—the effect of common population structure is small. J R Soc Interface [Internet]. 31 août 2016 [cité 24 févr 2020];13(121):20160288. Disponible sur: https://royalsocietypublishing.org/doi/10.1098/rsif.2016.0288
15. Bernard C. Introduction à la médecine expérimentale. Ligaran; 1865.