

KEY CHALLENGES IN MODELLING AN EPIDEMIC - WHAT HAVE WE LEARNED FROM THE COVID-19 EPIDEMIC SO FAR

KLJUČNI IZZIVI PRI MODELIRANJU EPIDEMIJE - DOSEDANJE IZKUŠNJE PRI MODELIRANJU EPIDEMIJE COVID-19

Ivan ERŽEN¹, Tina KAMENŠEK², Miha FOŠNARIČ², Janez ŽIBERT^{2,3*}

¹ National Institute of Public Health, Trubarjeva 2, 1000 Ljubljana, Slovenia

² University of Ljubljana, Faculty of Health Sciences, Zdravstvena pot 5, 1000 Ljubljana, Slovenia

³ University of Primorska, Andrej Marušič Institute, Muzejski trg 2, 6000 Koper, Slovenia

Received: May 25, 2020

Accepted: Jun 1, 2020

Editorial

ABSTRACT

Keywords:

COVID-19 modelling, epidemiological aspects, statistical recommendations, model quality

Abstract: Mathematical modelling can be useful for predicting how infectious diseases progress, enabling us to show the likely outcome of an epidemic and help inform public health interventions. Different modelling techniques have been used to predict and simulate the spread of COVID-19, but they have not always been useful for epidemiologists and decision-makers. To improve the reliability of the modelling results, it is very important to critically evaluate the data used and to check whether or not due regard has been paid to the different ways in which the disease spreads through the population. As building an epidemiological model that is reliable enough and suits the current epidemiological situation within a country or region, certain criteria must be met in the modelling process. It might be necessary to use a combination of two or more different types of models in order to cover all aspects of epidemic modelling. If we want epidemiological models to be a useful tool in combating the epidemic, we need to engage experts from epidemiology, data science and statistics.

IZVLEČEK

Ključne besede:

modeliranje COVID-19, epidemiološki pogled na modeliranje, priporočene statistične metode, kakovost modelov

Izvleček: Matematično modeliranje je lahko koristno za napovedovanje razvoja nalezljivih bolezni, saj s prikazom možnih izidov epidemije pomaga oblikovati javnozdravstvene ukrepe. Za napovedovanje in simulacijo širjenja v času epidemije COVID-19 so bile uporabljene različne tehnike modeliranja, vendar vse niso bile vedno koristne za epidemiologe in odločevalce. Da bi bili rezultati modeliranja zanesljivejši, je zelo pomembno kritično ovrednotiti uporabljene podatke ter preveriti, ali so bili upoštevani različni načini širjenja bolezni v populaciji ali ne. Izdelava dobrega epidemiološkega modela, ki je dovolj zanesljiv in ustreza trenutnim epidemiološkim razmeram v državi ali regiji, je zahtevna, zato je treba pri modeliranju slediti določenim kriterijem. Smiselno bi bilo tudi kombinirati dve različni vrsti modelov. Modeliranje bi bilo tako zanesljivejše, saj bi upoštevalo različne predpostavke. Če želimo, da bodo epidemiološki modeli koristno orodje v boju proti epidemiji, morajo pri modeliranju sodelovati strokovnjaki z različnih področij, predvsem epidemiologije, podatkovne znanosti in statistike.

*Corresponding author: Tel. + 386 1 300 11 76; E-mail: janez.zibert@zf.uni-lj.si

1 INTRODUCTION

The mathematical modelling of an infectious disease can provide an important insight into the stage of an epidemic and its evolution (1, 2). This can help decision-makers to estimate the impact of various intervention strategies. As the COVID-19 epidemic spread, modelling flourished, with experts from different professional backgrounds using a variety of modelling techniques to predict or simulate the epidemic. However, as we have seen, their predictions have not always been useful to epidemiologists and decision-makers in their efforts to combat the epidemic. Indeed, model outputs have frequently been misused to provide sensational news stories that have greatly frightened the public.

Mathematical epidemiological models can be divided into phenomenological, compartmental and individual-based models (1-4). Phenomenological models are data-driven statistical models that use regression analysis, often fitting epidemiological data to exponential or sub-exponential growth observed in the early stages of an epidemic (1). The other two groups of models exploit the insights of the epidemiological dynamics, such as the baseline transmission characteristics of specific pathogens and social contexts. In compartmental models, the population is divided into distinct groups or compartments. For example, in the SIR model (4, 5), these compartments are “susceptible”, “infectious” and “recovered”. While compartmental models usually assume the homogeneous mixing of the population, individual-based models mimic the social network of the population in more detail, for example by incorporating the epidemiological properties of individuals or groups of individuals and their interactions into dynamical networks (6, 7).

2 EPIDEMIOLOGICAL ASPECTS OF MODELLING

A particular challenge of epidemiological modelling is that there are many interventions in the context of an epidemic whose common purpose is to change the rate of spread of the infection. This has an impact on exposure of the population, which means that the course of the epidemic is usually different than that predicted by the models before the intervention. Data from the new epidemic situation should therefore be used. Proper communication and explanation of the new situation to the decision-makers, as well as to the general population, is of the utmost importance.

In the case of COVID-19, the introduction of measures that aimed to slow down or even stop the spread of infection caused a great deal of uncertainty regarding prediction of the further development of the epidemic. Moreover, the course of the infection is not yet fully understood, meaning that it is not possible to assess the impact of an individual measure on the course of the epidemic.

One of the major challenges in modelling epidemics is that the data relies on those who are, in some way, in contact with the healthcare system. The evaluation of the epidemiological situation and the estimation of further development is influenced by the number of people tested, hospitalised and undergoing treatment in intensive care units. If there is unequal access to healthcare for different groups of the population (which is usually the case), modelling predictions can be misleading. Again, it is very important to critically evaluate the data in order to understand the results yielded by an individual model in a specific region.

If the development of an epidemic is to be predicted, an assessment of the number of infected people in the population needs to be carried out. In the case of COVID-19, this could only be done by extrapolating the number of infected people from the number of deaths and the number of patients in intensive care and in hospital. However, this approach is only appropriate if the occurrence of cases in the population is homogeneous. This is usually not the case. We have seen that many people who became seriously ill from COVID-19 came from closed communities such as hospitals, nursing homes and similar. These people do not represent the whole population, and models that rely on the numbers of affected people from those communities are not accurate for the whole population. Clusters of the disease, which is a characteristic of COVID-19 as well, greatly affects the number of patients. However, from the point of view of the potential spread of the disease in the population, clusters are less risky than in cases of the diffuse occurrence of infected people among the population, where the potential for spread is much greater. If the model predictions are to be interpreted properly, it is important to consider whether or not these different ways in which the disease spreads through the population are being taken into account (6, 7).

3 RECOMMENDATIONS FOR RELIABLE EPIDEMIOLOGICAL MODELS

While, generally speaking, the basic epidemiological models are not too demanding to implement or adopt from the literature (1-7) or available software sources, for the above reasons it is a hard task to build an epidemiological model that is reliable enough and suits the current epidemiological situation within a country or region.

When building such a model, the following issues should be considered:

- Data should be derived from reliable, frequently updated sources.
- Data should be delivered in a timely fashion and should be as “deep” as possible, i.e. not just an aggregation of data over a day or across a region, but rather anonymised individual-based data.

- The modelling should take into account the different possibilities of spread among the population locally and among different groups of the population.
- The model should assume the non-homogeneous mixing of population.
- The dynamics of the disease should be included in the building process and should not rely solely on past data.
- The model must consider assumptions that are in line with the epidemiological state of the disease and/or are estimated from the studies or epidemiological findings of the disease, e.g. parameters such as incubation period, infectious time, time to recover, time in hospital, time in ICU, percentages of hospital care, percentages of ICU, etc. should be estimated from the epidemiological state of the disease and properly incorporated into the model.
- The model must satisfy minimal statistical standards such as:
 - evaluation of the model on real data;
 - construction of the model using standard statistical methodology;
 - equipping of the output results with confidence intervals.
- The model should be updated frequently to take account of new insights and new data, which results in more reliable predictions of the future occurrence of the disease.
- The model should be in line with other epidemiological models with the same properties, or should include explanations of why this is not so.
- The model should have publicly available source code and/or explanations of each step of the process.

The model should also be ready to simulate different scenarios of the dynamics of the disease from the present situation to the (near) future.

It might be also necessary to use a combination of two or more different types of model: phenomenological (1) and/or compartmental models (4, 5) to obtain an overview and understanding of the epidemiological dynamics, and individual-based models (6, 7) to take into account the non-homogeneous mixing of the population, i.e., different ways of spreading the disease locally and among different groups of the population.

If a reliable epidemiological model is to be built, it is therefore important to review and use the already developed and available methodology (1-4) and models (5-7) that address the above issues properly.

4 CONCLUSION

Within the large community working on more or less sophisticated mathematical modelling of COVID-19, only some follow the recommended statistical methodology and are suited to the epidemiological situation in individual countries or regions. Some of the models are already being used by decision-makers to track the dynamics of the disease or to estimate the impact of intervention strategies.

We strongly believe that epidemiological models that can be used to combat the epidemic need to engage experts from epidemiology, data science and statistics.

CONFLICTS OF INTEREST

The authors declare that no conflicts of interest exist.

FUNDING

There was no funding.

ETHICAL APPROVAL

This paper does not report on any study so ethical approval was not acquired.

REFERENCES

1. Chowell G, Sattenspiel L, Bansal S, Viboud C. Mathematical models to characterize early epidemic growth: a review. *Physics Life Rev.* 2016;18:66-97. doi: 10.1016/j.plrev.2016.07.005.
2. Li MY. An introduction to mathematical modeling of infectious diseases. Cham: Springer, 2018.
3. Vynnycky E, White R. An introduction to infectious disease modelling. Oxford: University Press, 2010.
4. Keeling M, Rohani P. Modeling infectious diseases in humans and animals. Princeton: University Press, 2008. doi: 10.2307/j.ctvc4m4gk0.
5. Dadlani A, Afolabi RO, Jung H, Sohraby K, Kim K. Deterministic models in epidemiology: from modeling to implementation. Technical report. Gwangju: Institute of Science and Technology, Communications and Sensor Networks Laboratory, 2013. arXiv preprint arXiv: 2004.04675. 2020.
6. Enright J, Kao RR. Epidemics on dynamic networks. *Epidemics.* 2018;24:88-97. doi: 10.1016/j.epidem.2018.04.003.
7. Kang H, Liu K, Fu X. Dynamics of an epidemic model with quarantine on scale-free networks. *Physics Lett A.* 2017;381:3945-51. doi: 10.1016/j.physleta.2017.09.040.