

The Application of Statistical and Mathematical Models to Pandemics--Taking COVID-19 as an Example

Bei Zhang^{1,*}

¹Mathematics, King's College London, Strand London WC2R 2LS, United Kingdom

Abstract. Since the outbreak of a large-scale epidemic can trigger fear and uncertainty, disease and death, and cause a great disruption to daily life while having a significant negative impact on the national and world economies, effective planning and research to deal with the spread of epidemics are of great importance. Since the outbreak of COVID-19 in 2019, applied mathematical-statistical models have been widely used in the control and research of world epidemics. In their response to the COVID-19 pandemic, countries and the World Health Organization (WHO) have used a variety of statistical tools and mathematical models to respond to and combat pandemic diseases. The continuing threat posed by infectious diseases has led to renewed research efforts around the world. This paper analyzed 13 relevant articles from 2010 to 2022, summarized and analyzed the main contributions of statistical and mathematical models in epidemic prevention and control, and put forward suggestions for future research directions, to increase the significance of data-based statistics for the study of infectious diseases, emphasizes the role and advantages of statistical and mathematical models in the control of infectious diseases, and advocates expanding the use of mathematical modelling of epidemics and encouraging increased cross-border collaboration in epidemiology among other related disciplines. This will bring a lot of benefits to future epidemiological research.

1 Introduction

Epidemics are a kind of social difficulties, which not only threaten human life but also cause various social disturbances and have a huge negative impact on the world economy and trade [1]. Take the recent COVID-19 outbreak in 2019 as an example, although it is difficult to specify the economic losses caused by the global COVID-19 pandemic. But according to statistics, the world's collective gross domestic product (GDP) fell by 3.4% in 2020, meaning that the pandemic caused more than \$2 trillion in lost economic output [2]. The cumulative number of confirmed COVID-19 cases worldwide is 603,711,760, including 6,484,136 deaths, according to the WHO.

Due to the huge social and economic losses caused by a large epidemic, effective prevention and control of the wide spread of the epidemic is of great importance. This paper analyzes the important role of mathematical model statistics in the prevention and control of large epidemics from the perspective of using mathematical model statistics and puts forward some suggestions for future research. In this paper, the key terms of statistics, mathematical models and pandemics were searched on several academic platforms, such as Google Academic and Web of Science. 13 articles closely related to the subject

were selected for in-depth analysis, and the relevant experience was summarized, which provided value for future epidemiological research.

2 The Application of Statistical and Mathematical Models to Pandemics

2.1 Introduction of epidemic case

A cholera epidemic in Paris in the 19th century, like the COVID-19 pandemic in 2019, exacerbated economic and social inequality in France. The Industrial Revolution drew large numbers of poor workers to Paris. Because of poor sanitation, slums where the poor congregated became hotbeds for cholera. The rich blame the poor for spreading disease, while the poor believe the rich are trying to poison them. The French society became tense, and as many as 18,000 people died in just six months [3]. And it had a huge impact on the French government at the time. This shows that the outbreak of a pandemic can have a huge impact on social unrest.

And research by IMF staff has found that countries with frequent and severe epidemics tend to experience greater social unrest (See figure 1) [4].

* Corresponding author: beizhang2125@gmail.com

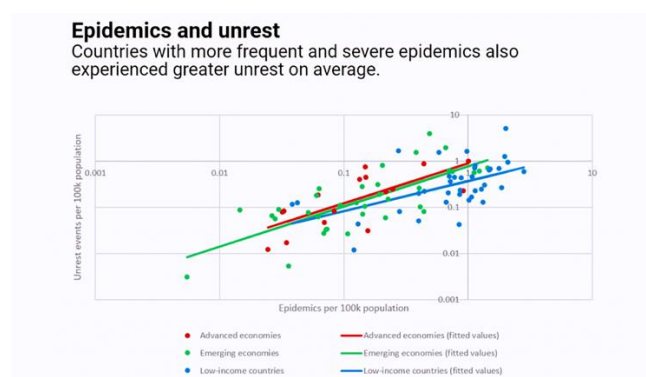


Figure 1. Linear relationship between social unrest and epidemic outbreak [4].

2.2 Mathematical modeling and epidemiological research process

Statistics is a discipline that summarizes and analyzes data prone to random variation [5]. A pandemic is an unpredictable event prone to random variation. Modeling the spread of infectious diseases has a long history of application in assessing epidemiological phenomena [6]. Through the statistics of the large epidemic cases, and the construction of relevant mathematical models, can effectively prevent the outbreak of large epidemic. In epidemiology, the role of mathematical modeling is usually to define the interactions between individuals or groups and the environment, and to translate the rules described by these definitions into equations that can be quantified using mathematical rules. This allows accurate and efficient mathematical data to be used to examine possible future effects.

A mathematical model is an entity that is similar to the system or object under study in some aspects but is simpler or easier to use compared with the research target [7]. Mathematical models are usually used to: 1) better understand the system or object under study; 2) simulate the behavior of the system or object according to similar characteristics; 3) predict the future of the research object (even predict the end); 4) use the predicted results to better control the research system.

There are four key steps in epidemiological research: 1) identifying the research question, designing the study, 2) collecting the data, 3) using statistical knowledge to collate and analyze the data, and 4) applying the results of the study to public health. [8] When we apply mathematical models to epidemiology, we only need to operate in the following four steps: 1) identify the key parts of the model; 2) identify and input 3) start and run the model; 4) Output and interpret the data, and interpret the application of the results

2.3 Case analysis of 2019 COVID pandemic

The outbreak of COVID-19 in 2019 has plunged human society into unprecedented turmoil, and its impact on society has not completely subsided until now. The World Health Organization announced that as of April 20, 2022, 50.4 million confirmed cases had been reported, including 6.7 million deaths directly related to COVID-19 [9]. Mathematical modeling has played an extremely important role in the policy of COVID-19 prevention and control. The corresponding mathematical model can be established to analyze various problems of epidemic prevention and control.

Since the outbreak of COVID-19, with the help of appropriate mathematical models, many important works on the epidemic has been completed in time. Li L, Yang Z, Dang Z, et al., developed an effective model for predicting the number of confirmed cases and deaths of COVID-19 [10]. Arino J, Portet S., proposed a simple model of COVID-19 that is highly sensitive to parameter values describing the proportion of asymptomatic cases [11]. Mizumoto K and Chowell G studied the dynamics of COVID-19 transmission on the Diamond Princess, a large international cruise ship in Japan. In subsequent work on the outbreak, Watson O. J., Barnsley G., Toor J., et al. modeling studies were also conducted on the global impact of COVID-19 vaccine vaccination [12].

Next, this paper focuses on specific examples of the application of mathematical models to COVID-19. Li L., Yang Z., Dang Z., et al. [10] based on the different transmission characteristics of COVID-19 in different stages, a transmission model of COVID-19 was constructed by using Gaussian distribution theory to simulate the transmission process of COVID-19. It is found that the data simulated by the mathematical model fit well with the official data curves of Hubei, non-Hubei regions of China, South Korea, Italy and Iran.

For example, the prediction of the time to complete cure of more than 10,000 cases in Hubei Province, China at that time.

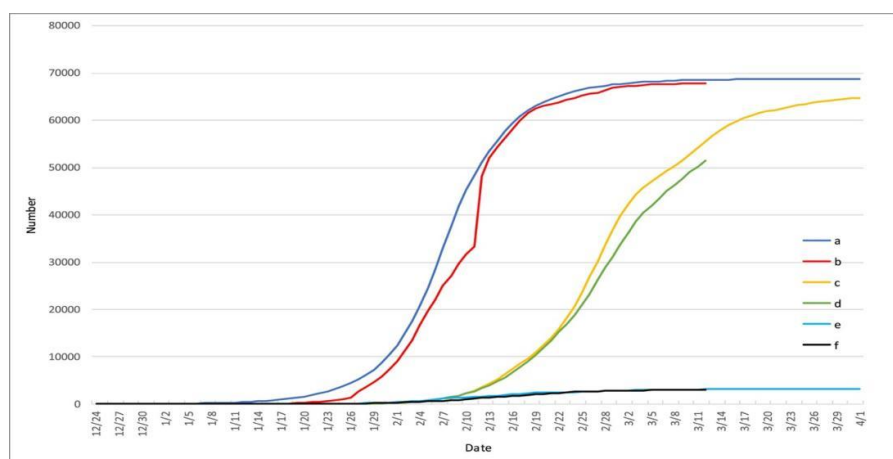


Figure 2. Comparison between the official data of the epidemic and the simulation data of Chinteria province in 2020 (Curve a - Number of simulated infections. Curve b - Number of officially confirmed infections. Curve c - Number of simulated cures. Curve d - Number of cures (official. Curve e - Number of simulated deaths. Curve f - Number of official deaths.) [10]

There are three stages in the simulation. The first phase is before January 23, when Wuhan is not closed; the second phase is from January 23 to February 10; the third phase is after February 10, when Wuhan begins to lock down the community. It is easy to find that the number of

cured deaths in the predicted data is basically consistent with the final official data of Hubei Province.

And their simulation projections for the 2020 Italian epidemic data:

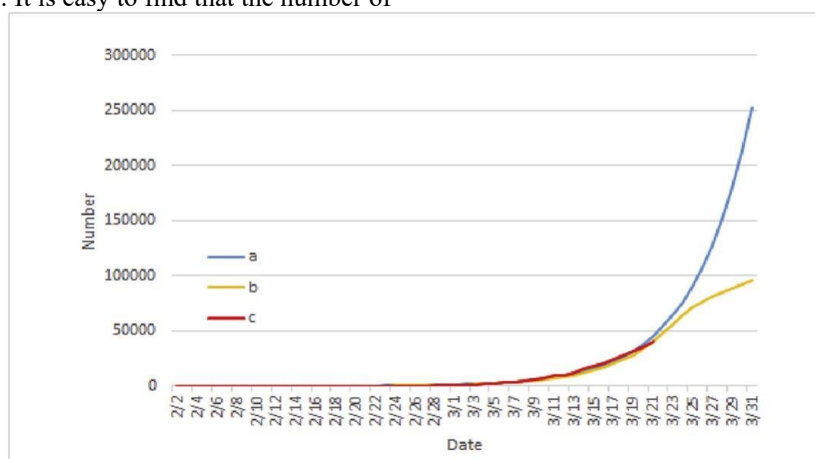


Figure 3. Comparison of Italy epidemic simulation data and official data in 2020. (The curve a-Number of simulated infections (uncontrolled). The curve b-Number of simulated infections (partial control). The curve c-Number of officially confirmed infections.) [10]

From the Figure 3, it can be clearly seen that the number of confirmed cases in Italy will increase dramatically in the short-term starting around March 15th, and even reach 250,000 by the end of March if not controlled in time. The model also suggests that if measures are taken in time to partially contain the outbreak, the number of confirmed cases in Italy could reach 100,000 by the end of March and be largely under control. In fact, Italy began to implement some measures (lockdown) to contain the epidemic on March 8, and the data shown in curve C basically match the model's prediction of the data under partial control. The evolution trend of the existing epidemic data was predicted, and it was found that control had an important impact on the epidemic.

In these two examples, mathematical models are remarkably accurate in predicting pandemics like COVID-19. Therefore, the data predicted by the

mathematical model has full reference significance for the prevention and control of the epidemic.

2.4 A Model Based Study on the Dynamics of COVID-19: Prediction and Control

They built a mathematical model of the coronavirus pandemic from a different perspective. Qualitative analysis is more important than quantitative analysis. To find out the pathological causes of the epidemic and predict the spread of the disease from the most fundamental mode of transmission [13]. The original mathematical model of epidemiology, first seen by Bernoulli [14], is very simple. However, with the progress of computing tools and technology, we can design more complex and efficient mathematical models, which can be better applied in the prevention and control of the epidemic.

A new COVID-19 model was developed based on the reality. It contains five subcategories $S(t)$, Exposed $E(t)$, Hospitalized Infected $I(t)$, Quarantine $Q(t)$ and Recovered or Removed $R(t)$. The object uses the first order differential equation to form the autonomous dynamic system and makes the system flow chart. Beyond that, a very successful mathematical model for studying the qualitative dynamics of COVID-19 is derived through many complex mathematical calculations, such as 1) equilibria A4 (stability analysis of the system, 2) Basic Reproduction number3) equilibria A4) and optimal control problems. And they showed real cases where they could actually apply the mathematical model, real data from three populous states in India, using Mathematica to fit their mathematical model. In addition, in their experiments, they demonstrated that both isolation and media coverage played an important role in reducing the spread of COVID-19.

2.5 Problems in the application of mathematical models in epidemiology

One of the main problems with applying mathematical modeling to epidemiology on a larger scale is that not many epidemiologists have advanced mathematical modeling skills [15]. More epidemiologists with more advanced mathematical training which can perform high-level analyses are still needed. More reliable and reputable epidemiological modelers are not only able to design more efficient and accurate epidemiological models but are also better able to explain to the public, political scientists, the media, and health professionals the applicability and limitations of their models. Such scholars are in short supply and badly needed. This paper argues that the only solution in the short term is to expand cross-border cooperation, allowing more epidemiologists to collaborate with related scholars, including but not limited to mathematicians and computer scientists. This will provide broader selectivity for future studies in epidemiology, which will benefit from such broad collaborations.

3 Conclusion

This paper specifically analyzed several articles on the application of mathematical models in the field of epidemiology and summarized the important role of statistics and mathematical models in preventing and controlling epidemics. It complements the information needed by future scholars for relevant research and broadens the application scope. In particular, the application of mathematical models to the COVID-19 outbreak that has not subsided since 2019. It is concluded that the mathematical model can provide a lot of benefits for epidemiological research, simplify and efficient information processing, and save more energy and time for epidemiologists. Moreover, this study finds that mathematical modeling can predict the future of events with very high accuracy, which is of great significance for the prevention and control of epidemics, sudden diseases with long-term large-scale effects.

However, the limitations of this paper are also very obvious. Firstly, the amount of articles analyzed is limited, which may lead to incomplete information and insufficient thorough analysis of problems. As for the future research direction of epidemiology, the suggestion is to increase the mutual cooperation between epidemiology and other related disciplines. This will enable the staff to better apply more excellent and convenient scientific tools such as mathematical models in the process of studying epidemiology.

References

1. McKay, R.A. & Doroshov, D.B. The Dangers of “Us Versus Them”: Epidemics Then and Now. *J GEN INTERN MED* 36, pp.795–796 (2021). <https://doi.org/10.1007/s11606-020-06368-y>
2. Statista Research Department, Impact of the coronavirus pandemic on the global economy - Statistics & Facts, Sep 13, 2022.
3. Barret P, Chen S, Li N. COVID’s Long Shadow: Social Repercussions of Pandemics [J]. IMF blog, (2021).
4. Kudlick, Catherine J.. Cholera in Post-Revolutionary Paris: A Cultural History, Berkeley: University of California Press, (2020). <https://doi.org/10.1525/9780520916982>
5. Last J. M. A dictionary of epidemiology. 2nd ed. Oxford, Oxford University Press, (1988).
6. Kermack William Ogilvy and McKendrick A. G. 1927A contribution to the mathematical theory of epidemicsProc. R. Soc. Lond. A115700–721. <http://doi.org/10.1098/rspa.1927.0118>
7. Sameni R. Mathematical modeling of epidemic diseases; a case study of the COVID-19 coronavirus [J]. arXiv preprint arXiv:2003.11371, (2020).
8. Chubb M. C., Jacobsen K H. Mathematical modeling and the epidemiological research process [J]. European journal of epidemiology, 25(1): 13-19 (2010).
9. World Health Organization. World health statistics 2022: monitoring health for the SDGs, sustainable development goals [J]. (2022). <https://apps.who.int/iris/bitstream/handle/10665/356584/9789240051140-eng.pdf?sequence=1>
10. Li L, Yang Z, Dang Z, et al. Propagation analysis and prediction of the COVID-19[J]. Infectious Disease Modelling, (2020).
11. Mizumoto K, Chowell G. Transmission potential of the novel coronavirus (COVID-19) onboard the diamond Princess Cruises Ship, 2020 [J]. Infectious Disease Modelling, 5: 264-270 (2020).
12. Watson O. J., Barnsley G., Toor J., et al. Global impact of the first year of COVID-19 vaccination: a mathematical modelling study [J]. The Lancet Infectious Diseases, 22(9): 1293-1302 (2022).
13. Mandal M., Jana S., Nandi S. K., et al. A model based study on the dynamics of COVID-19: Prediction and

control [J]. Chaos, Solitons & Fractals, 136: 109889 (2020).

14. Bernoulli D. Essai d'une nouvelle analyse de la mortalité causée par la petite vérole, et des avantages de l'inoculation pour la prévenir [J]. Histoire de l'Acad., Roy. Sci.(Paris) avec Mem, pp.1-45 (1760).
15. Chubb M. C., Jacobsen K. H. Mathematical modeling and the epidemiological research process [J]. European journal of epidemiology,25(1): 13-19 (2010).