Modeling the transmission dynamics of COVID-19 epidemic: a systematic review

Guan Jinxing, Wei Yongyue, Zhao Yang, Chen Feng

Cite this article as:

View online: https://doi.org/10.7555/JBR.34.20200119

Articles you may be interested in

Hidden Markov model based epileptic seizure detection using tunable Q wavelet transform

Modeling and validating three dimensional human normal cervix and cervical cancer tissues in vitro

Shunts, channels and lipoprotein endosomal traffic: a new model of cholesterol homeostasis in the hepatocyte

Establishment of a rat model with diet–induced coronary atherosclerosis
Modeling the transmission dynamics of COVID-19 epidemic: a systematic review

Jinxing Guan, Yongyue Wei, Yang Zhao, Feng Chen

Department of Epidemiology and Biostatistics, School of Public Health, Center for Global Health, Nanjing Medical University, Nanjing, Jiangsu 211166, China.

Abstract

The outbreak and rapid spread of COVID-19 has become a public health emergency of international concern. A number of studies have used modeling techniques and developed dynamic models to estimate the epidemiological parameters, explore and project the trends of the COVID-19, and assess the effects of intervention or control measures. We identified 63 studies and summarized the three aspects of these studies: epidemiological parameters estimation, trend prediction, and control measure evaluation. Despite the discrepancy between the predictions and the actuals, the dynamic model has made great contributions in the above three aspects. The most important role of dynamic models is exploring possibilities rather than making strong predictions about longer-term disease dynamics.

Keywords: coronavirus, COVID-19, epidemiology modeling, transmission dynamic model, SEIR model, SIR model

Introduction

Since December 2019 when the first severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) cases were reported in Wuhan, China, SARS-CoV-2 has spread to all the Chinese provinces. The disease was officially named COVID-19 by World Health Organization (WHO) on February 11, 2020. Meanwhile, the outbreak of this disease has been announced as a Public Health Emergency of International Concern[1]. According to the WHO observations, the COVID-19 epidemic peaked and plateaued between January 23 and February 2, 2020, in China, and then followed by a rapid decline, paving a path for the elimination of the epidemic in China. However, the pandemic has stirred up a huge wave from European to American, and African countries. As of June 28, there have been over 10 million confirmed cases of COVID-19 and half million deaths caused by it[2], with no trend of moderation. Yet, no efficient antiviral treatments or vaccines against SARS-CoV-2 and COVID-19 has ever been developed.

To prevent the epidemic of COVID-19, it is crucial to understand the early transmission dynamics and estimate the effect of control policies[3]. Transmission dynamic model was widely used to better understand the transmission mechanisms and the factors that are most influential in spread, thereby making more accurate predictions, determining and evaluating control strategies[4].

The authors reported no conflict of interests.

This is an open access article under the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited.
The most commonly used models in COVID-19 are SIR and SEIR. The classic SEIR model includes 4 compartments (Fig. IA): susceptible (S), exposed (E), infectious (I), removed (R). The dynamics of these compartments across time (t) are described by the following set of ordinary differential equations:

\[
\begin{align*}
\frac{dS}{dt} &= -\beta IS /N \quad (1) \\
\frac{dE}{dt} &= \beta IS /N - \alpha E \quad (2) \\
\frac{dI}{dt} &= \alpha E - \gamma I \quad (3) \\
\frac{dR}{dt} &= \gamma I \quad (4) \\
N &= S + E + I + R
\end{align*}
\]

in which \( \beta \) is the transmission rate for infections (defined as the number of individuals that a case can infect per day); \( \alpha \) is the reciprocal of the incubation period; \( \gamma \) is the reciprocal of the infectious period; \( N \) is the total population.

Considering the cases in incubation period or asymptomatic cases, some studies added a transmissibility factor \( \epsilon \), the ratio of the transmission rate of infections to that of asymptomatic cases/ incubation population, in the equation (1):

\[
\frac{dS}{dt} = -\beta(I+\epsilon E)S /N \quad (5)
\]

Other studies added an asymptomatic compartment \( (A) \) (Fig. IB), and the new equation (1) was as follows:

\[
\frac{dS}{dt} = -\beta(I+\epsilon E+\epsilon A)S /N \quad (6)
\]

To estimate the unascertained proportion in the early dynamic, some studies\(^{[5]}\) extended the SEIR model with an unascertained compartment (Fig. IC).

Transmission dynamics modeling is a way to formalize what is known about transmission of COVID-19 and explore the possible futures of a system with nonlinear trend that is almost impossible to do using intuition alone\(^{[6]}\). A series of studies have adapted dynamic models to investigate the crucial epidemiological parameters, predict the trend, estimate the peak, and evaluate control measures, so as to provide critical information for decision making\(^{[7]}\). However, the extent of immunity, the transmission rate of people with no or minimal symptoms and varied contact rates during rapidly changing intervention and control measures remain uncertain, which could have limited the model’s ability to predict the future of the COVID-19 pandemic\(^{[6,8]}\).

Thus, to evaluate these models and to explore a better solution for COVID-19 modeling, we searched PubMed for studies on dynamic model of COVID-19 published or publicly posted from December 31, 2019 to July 13, 2020 (date of the last search), and finally collected 63 studies (Fig. 2).

**Model extension**

Considering the asymptomatic cases, various

![Fig. 1 Illustration of transmission dynamic models.](image)

**Fig. 1** Illustration of transmission dynamic models. A: Basic SEIR model. B: SEIR model modified with an asymptomatic compartment (\( A \)). C: SEIR model modified with an unascertained compartment (\( I_u \)).

![Fig. 2 PRISMA flow diagram.](image)

**Fig. 2** PRISMA flow diagram. A total of 63 studies were finally included in the systematic review. PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses.
control measures, stages of COVID-19 and other factors, some studies extended the basic SIR and SEIR model.

Most studies (84.1%) considered compartment for exposures (latent infections). In addition, three studies extended the compartments based on SEIR model for unascertained infections (11.1%), severity stages (25.4%), and quarantined susceptibles and exposures (27.0%). Moreover, five studies extended the SEIR model to the network structure by accounting for the transmission across cities and regions.

On the other hand, a series of studies extended the model by adding additional parameters to characterize the epidemic. Among the 19 (30.2%) studies including infectibility of latent/asymptomatic infections, 10 studies set a different transmission ability for E versus I. Six studies contemplated unascertained proportions, eight accounted for the in- and out- floating population, and 16 studies made allowance to evaluate the effect of control measures (Fig. 3A and Supplementary Table 1, available online).

In the 63 studies identified, 43 studies estimated the transmission rate and two used a fixed value, while 18 studies failed to report the transmission rate they used. In addition, ten studies estimated the incubation period, although 40 chose a fixed one. Ten studies

**Fig. 3 Overview of studies.** A: The extending models used. B: Objectives of studies. C: The number of studies that used fixed or estimated transmission rate, incubation period, and transmissibility factor (the transmission rate of E/the transmission rate of I). D: The number of fixed and estimated parameters of each study. References were listed in Supplementary.
reported the transmissibility factor, the factor comparing the infectiousness of exposed or asymptomatic infections with infections, six estimated it, while four used a fixed value (Fig. 3C and Supplementary Table 3, available online).

Forty-three studies reported the entire parameter list they used, and the total number of parameters varied from 6 to 29 with different proportion of fitted parameter (Fig. 3D and Supplementary Table 4, available online). The maximum number of estimated parameters was 22 with 0 fixed parameter in the study, while the maximum number of fixed parameters was 23 with 5 estimated parameters.

The objectives of these models included the following three aspects: epidemiological parameter estimation (14 studies), theoretical and actual trend prediction (30 studies), and control measure evaluation (35 studies) (Fig. 3B and Supplementary Table 2, available online).

Estimation of epidemiological parameters

Fourteen of sixty-three studies focused on the estimation of epidemiological parameters, which could help to better understand the transmission potential of COVID-19 and contribute to prediction and decision-making.

Transmission rate

Transmission rate, also called transmission coefficient ($\beta$), refers to the per capita rate at which two specific individuals come into effective contact per unit time[9]. It is a crucial parameter to describe the speed at which an epidemic of a particular disease progresses[4], calculate the basic productive number ($R_0$) and the risk of infection ($\lambda$), and make better prediction of disease. The changes of transmission rate often show the efficiency of quarantine and control measures.

We collected the estimates of transmission rate without control measures in the early dynamic from 13 studies. And the value varied largely among different regions/countries. In China, the estimates of transmission rate with no control measures were all above 1, while in European regions three estimates were all below 1 and in Canada it was 1.7. The estimates in the US varied from 0.5 to 1.4. However, the transmission rates were estimated much higher for some regions of Republic of Korea (6.18) and the whole Republic of Korea (7.06), while one study estimated that in Algeria was 0.41 (Fig. 4A and Supplementary Table 5, available online).

Transmissibility factor

An increasing number of studies emphasized the existence of asymptomatic cases[10–12] and took asymptomatic cases as a compartment in model[13–18]. Asymptomatic cases are potential source of infection[10–12]. Due to their contagiousness and numbers, asymptomatic cases can infect a far greater proportion of the population than would otherwise occur[9].

Totally, ten studies considered the different transmission rate of exposed/asymptomatic carriers versus symptomatic infections. Four studies adopted a fixed value for transmissibility factor ($\epsilon$) at 0.5 (3 studies) or 1 (one study). In addition, all six estimates extracted from four studies had the mean and the 95% confidence interval (CI) lower than 0.7. In one study[19], the estimates ranged from 0.32 to 0.69 (median 0.5), comparable to 0.5, the most commonly used fixed value (Fig. 4C, Supplementary Table 6 and 7, available online).
Unascertained proportion

In the early dynamic of COVID-19, many unascertained infections (with mild or no symptoms) could transmit the virus to a far greater proportion of population than ascertained infections, which is a critical epidemiological characteristic modulating the pandemic potential of COVID-19\(^\text{[20–22]}\). Therefore, some studies extended SEIR model with an unascertained compartment or an unascertained parameter.

Six studies estimated the unascertained proportion, or unreported proportion, and four out of six studies estimated the unascertained proportion of early epidemic in Wuhan ranging from 60% to 99.8% based on data between January 3 and February 2. Another study estimated the ascertained rate of Brazil, Italy and Republic of Korea based on the data up to April 6. Republic of Korea had the highest ascertained rate, 95.6%, meaning that 4.4% cases were unascertained ones in the early dynamic. Besides, one study estimated the ascertained rate of France was 12.5% (95% CI, 8.3%–20%) (Supplementary Table 8, available online).

Incubation period

Understanding the incubation period plays an important role in evaluating the transmission potential, predicting the epidemic trend, and informing the active monitoring or mandatory quarantine period\(^\text{[23]}\). Active monitoring requires the potentially exposed persons to contact local health authorities and report their health status every day. Understanding the length of active monitoring could limit the risk for missing SARS-CoV-2 infections and help the health departments to use the limited resources effectively\(^\text{[24]}\). Furthermore, increasing evidence indicates that SARS-CoV-2 infections could excrete the pathogen and cases during the asymptomatic incubation period are infectious\(^\text{[10–12]}\). Thus, it is crucial to know the length and dispersion in the incubation period for better prevention and control of COVID-19. However, the incubation period of COVID-19 is poorly understood, which could result in a biased prediction\(^\text{[6]}\).

We identified 42 published articles reporting the assumed fixed incubation period, ranging from 2.5 to 10 days (Supplementary Table 9, available online), based on official information or high-quality research. The most commonly assumed incubation period in the model was 5.2 days (11 studies), followed by the 7-day incubation period (7 studies).

Fifteen estimates extracted from ten published studies had the mean (or median) and uncertain incubation period of shorter than 7 days (except one outlier), ranging from 3\(^\text{[25]}\) to 6.67 days\(^\text{[26]}\) (median 4.2 days) (Supplementary Table 10, available online), approximating to SARS-CoV (4.4 days)\(^\text{[27]}\) and the 31 fixed values were also in this range (Fig. 4B).

Trend prediction

The COVID-19 outbreak presents a major challenge to epidemic control in a well-connected and densely populated city and the decision on the time to implement control measures\(^\text{[20]}\). Trend prediction, as a main objective of dynamic model, can provide reference for the prevention and control measures\(^\text{[13]}\).

Short-term prediction of infections

Short-term predictions provide the evidence that policymakers may need to allocate resources or plan interventions\(^\text{[6]}\). Deviations of short-term prediction showed more changeful results, ranging from 216 to 27,578.

We identified 7 studies that made short-term predictions (shorter than 15 days between the date of data acquisition and the date of prediction) and 8 predictions were made. Sixty-two point five percent of predictions had a deviation ratio lower than 50%, which means the prediction is approximated to the actual confirmed cases. Besides, the deviation ratios of 37.5% studies were between 50% and 100% (Fig. 5A and Supplementary Table 11, available online).

Compared with the actual data, short-term predictions are relatively accurate but biased as a
result of data gaps and inherent uncertainties about future human behavior and interventions\cite{6}. A very important reason for short-term prediction is the value for working on emerging detection, prevention, therapy and control programs\cite{4}.

**Long-term prediction of infections**

Deviations of long-term prediction showed more changeable results, ranging from 177 to 2 594 827. We identified 19 studies with 23 long-term predictions, 15 of which were for regions in China, one for the US, four for India, one for Republic of Korea and another two for Italy. 21.7% predictions had a deviation ratio higher than 100%, while 4.3% predictions had a deviation ratio between 50% and 100%. And 74.0% predictions had a lower-than-50% deviation ratio (Fig. 5B and Supplementary Table 12, available online).

**Peak time prediction**

Eight studies predicted the peak time of regions in China. The deviation days between the predicted peak time and actual peak time ranged from 5 to 23. The deviation days of 40% predictions were less than or equal to 5, while 30% predictions had a deviation larger than 10 days. The deviation of other 30% predictions were between 6 days and 10 days (Fig. 5C and Supplementary Table 13, available online).

**Theoretical prediction**

Among the 15 studies that made a theoretical model, 6 modified the contact rate with different assumed values, 1 modified the transmission rate with fixed values and 2 modified the quarantine rate. Two out of fifteen studies fitted the trend of a delay in control measures. Four out of fifteen studies fitted the situations of no control measures or maintaining current measures (Supplementary Table 14, available online).

**Effects of prevention and control measures**

The basic aim of studying the spread of COVID-19, both in time and in space, is to gain a better understanding of transmission mechanisms and the most influential features in that spread, so as to make predictions, and to determine and evaluate control strategies\cite{6}. With no pharmaceutical treatments available, interventions of COVID-19 have focused on contact tracing, quarantine, and social distancing. During the initial pandemic wave, many countries have adopted social distancing measures, and some, like China, are gradually lifting them after achieving adequate control of transmission\cite{16}.

**Actual effect of single control measures**

Seven studies evaluated the actual effect of single control measure, and four evaluated travel ban. The travel ban to Wuhan avoided 13 602 cases\cite{28}, and resulted in 2.91 (2.54, 3.29) days delay of epidemic outbreaks in cities in Hubei (except Wuhan), and 542 000 cases were avoided in China (except Wuhan)\cite{29}. Similarly, travel ban in Europe\cite{30} reduced 10% of the daily cases. In France\cite{31}, it reduced the effective reproductive number ($R_e$) from 3.2 (95% CI, 3.1–3.3) to 0.47 (95% CI, 0.45–0.50). Besides, the effects of social distance strategies\cite{12–33}, clinical diagnosis and universal symptom surveys\cite{34} were also evaluated\cite{29} (Supplementary Table 15, available online).

**Actual effect of comprehensive control measures**

Seven studies evaluated the actual effect of comprehensive control measures. Control measures in Wuhan reduced the effective reproductive number from 3.1 to 0.5\cite{35}, while control measures in China avoided 96% cases in China (except Wuhan)\cite{29} and 99.85% (58 613 197) cases in mainland China\cite{6}. In Italy, control measures reduced 649 000 averted cases (95% CI, 481 000 –1 010 000) in total\cite{32}. In Germany\cite{37} and Bali, Indonesia\cite{38}, control measures reduced the transmission rate. In addition, control measures in Santa Clara County, California reduced the effective reproductive number from 2.88 (2.47, 3.45) to 0.98 (0.85, 1.11)\cite{39} (Supplementary Table 16, available online).

**Theoretical effect of prevention measures**

Nine studies evaluated the theoretical effect of some prevention measures prior to the implementation. Five studies evaluated the opening or releasing strategies, and the effects of wearing facial masks\cite{40} and media reports\cite{41}. Besides, Weitz et al\cite{42} came up with a "Modeling Shield Immunity" and evaluated its effect (Supplementary Table 17, available online).

**Discussion and perspectives**

Dynamic models can mimic the way SARS-CoV-2 spreads and reflect the underlying transmission process. The disease-specific parameters can be modified to test how the pandemic may change under various assumptions and control measures\cite{6}. In COVID-19, dynamic models were used to forecast or simulate future transmission scenarios under various assumptions about parameters governing transmission,
disease, and immunity. A main purpose of epidemiological modeling is to forecast the future incidence of a disease and identify the trend of it. However, even with the best modelling efforts, the course of the epidemic cannot be accurately predicted.

We identified 63 studies related to transmission dynamic model up to July 13. These studies estimated epidemiological parameters, made trend predictions and assessed the effectiveness of prevention and control measures by extending basic dynamic models (SIR or SEIR model).

The dynamic models with extensions are helpful in epidemiological parameter estimation, such as transmission rate, incubation period and transmissibility factor. Transmission rate in early dynamic indicates the potential of further spread and the ability to cause disease without any control measures. In these models, the estimates of incubation period ranged from 3\textsuperscript{[25]} to 6.67 days\textsuperscript{[26]}, comparable to SARS-CoV (4.4 days)\textsuperscript{[27]}, and 31 out of 42 fixed values assumed according to previous studies and official documents were also within this range. Estimates of transmissibility factor were also similar to fixed values. According to experts' advice and the previous studies, 12 studies chose 5.2 days as the incubation period, matching the values of estimated incubation period. If a fixed incubation period must be chosen, 5.2 days may be a reliable choice.

We found that the fixed values of transmissibility factor were 0.5 and 1. In the early studies, the transmission ability of asymptomatic cases remained unknown and some studies assumed transmission ability of asymptomatic cases were the same as that of symptomatic infections. Later, most estimates of transmissibility factor ranged from 0.32 to 0.69, comparable to 0.5, according to expert advice or high-quality studies. As a consequence, 0.5 may be a more reliable value if a fixed transmissibility factor is to be chosen.

In the parameter estimation of dynamic models, the least square, Maximum Likelihood Estimation (MLE), Monte Carlo (MC) and Markov Chain Monte Carlo (MCMC) are the main methods adopted. The most commonly used method is MCMC method. Based on Bayesian Statistics, MCMC method firstly constructs a Markov chain of parameter values. The next parameter combination is chosen by proposing a random move conditional on the last parameter combination and accepting conditional on the matrix of transition probability\textsuperscript{[24]}. Though faster than MCMC, MC method has a larger confidence interval. The estimations of MCMC method are more accurate than MLE, MC and the least square method. But if the initial values and prior distributions are not appropriate, MCMC method are prone to be stuck in areas of low likelihood, which may cost more time. With big data of infectious disease, it is possible to set proper prior distributions and initial values, which makes MCMC method worth using.

In 28 studies, the predictions are not all sufficiently accurate, whether in short-term or long-term predictions. The short-term predictions and longer-term predictions indicate almost all predictions are biased, even when models are made more complicated in order to better approximate actual disease transmission\textsuperscript{[4]}. We noticed that predictions using data of the early transmission period, whether short-term or long-term, showed larger deviations than predictions based on data of well-controlled transmission period. This may be caused by the rapidly changing intervention and control measures in the early transmission period. When the transmission of disease is well controlled, the intervention and control measures are not fast-changing, making the predictions relatively accurate.

Interestingly, we found that most studies using a large number of estimated parameters focused on control measure evaluation\textsuperscript{[16,29,37,40–41,44–45]}, A study\textsuperscript{[46]} focusing on trend prediction with 6 fixed parameters and 14 estimated parameters had a deviation of 113 035 cases compared with actual data. The estimate of transmissibility factor by another study\textsuperscript{[49]} with 10 fixed parameters and 19 estimated parameters was 0.0275, while other estimates or fixed values ranged from 0.32 to 0.69.

The model accuracy is constrained by the present knowledge of the virus. The number of people being or having been infected is the most obvious uncertainty and it is spatially heterogeneous and time-varying. In a dynamic model, uncertainty in a key epidemiologic parameter or a set of parameters. For example, the duration of infectiousness may be presented as a range around a mean trajectory, such as 95% CI, reflecting simulations across the plausible or measured values of a parameter, or as separate simulations. But only a few studies provided complete parameter values and their ranges.

Because transmission dynamic models are simplifications that have usually unknown relationships with actual disease spread, one can never be completely certain about the validity of findings obtained from modeling, such as conceptual results, experimental results, answers to questions, comparisons, sensitivity results and forecasts. To better describe and fit, more factors are put into the model or assumed. But even when models are made
more complicated in order to better approximate actual disease transmission, they are still abstractions. Any model involves trade-offs between simplicity and realism. Identification of the relevant factors is necessary when analyzing a specific disease[4]. Uncertainty about the extent of protective immunity, the extent of transmission and immunity among people and to measure and model contact rates between susceptible and infectious people are the three model parameters that specifically limit our ability to predict the future of the COVID-19 pandemic[6].

Dynamic models are commonly used to estimate epidemiological parameters, make predictions and forecasts, evaluate the intervention and control measures. Several measures are suggested and proved to be effective by models[40,42]. Until we have better data on antibody kinetics and protection against reinfection, models will be useful for exploring possibilities rather than making strong predictions about longer-term disease dynamics[6]. In some regions, the transmission of COVID-19 was well controlled under severe intervention and control measures. However, these measures could not be sustained indefinitely because of the limited financial, material and human resources. Therefore, in the absence of a vaccine or pharmaceutical treatment, it is important to understand long-term consequences of different measures and generate the stepping-down strategies or strategies of mitigation. Evaluating the strategies of mitigation may become another active subject of modeling study.

In the 63 studies we identified, the study populations were limited in China, the United States and European countries. We failed to found a study focusing on Middle East and African populations. This is a major shortcoming of the previous model studies. People of developing countries and regions account for a very large proportion of the world’s population. Due to the insufficient testing ability, the number of COVID-19 cases in African and Middle East may take the lead. Thus, model study in the population of developing country may become a research focus in the future.

Acknowledgments

This study was funded by National Natural Science Foundation of China (82041024 to F.C.). Sponsors had no role in design of the study, collection and analysis of data, or preparation of the manuscript.

References


