

Preliminary prediction of the control reproduction number of COVID-19 in Shaanxi Province, China

LI Zhi-min¹ ZHANG Tai-lei^{1,*} GAO Jian-zhong²
LI Xiu-qing³ MA Ling-juan¹ BAO Xiong-xiong¹

Abstract

Objectives Firstly, according to the characteristics of COVID-19 epidemic and the control measures of the government of Shaanxi Province, a general population epidemic model is established. Then, the control reproduction number of general population epidemic model is obtained. Based on the epidemic model of general population, the epidemic model of general population and college population is further established, and the control reproduction number is also obtained.

Methods For the established epidemic model, firstly, the expression of the control reproduction number is obtained by using the next generation matrix. Secondly, the real-time reported data of COVID-19 in Shaanxi Province is used to fit the epidemic model, and the parameters in the model are estimated by least square method and MCMC. Thirdly, the Latin hypercube sampling method and partial rank correlation coefficient (PRCC) are adopted to analyze the sensitivity of the model.

Conclusions The control reproduction number remained at 3 from January 23 to January 31, then gradually decreased from 3 to slightly greater than 0.2 by using the real-time reports on the number of COVID-19 infected cases from Health Committee of Shaanxi Province in China. In order to further control the spread of the epidemic, the following measures can be taken: (i) reducing infection by wearing masks, paying attention to personal hygiene and limiting travel; (ii) improving isolation of suspected patients and treatment of symptomatic individuals. In particular, the epidemic model of the college population and the general population is established, and the control reproduction number is given, which will provide theoretical basis for the prevention and control of the epidemic in the colleges.

Received: 2020-03-16. Revised: 2020-09-04.

MR Subject Classification: 34A30.

Keywords: COVID-19, control reproduction number, general population, college population, Shaanxi Province.

Digital Object Identifier(DOI): <https://doi.org/10.1007/s11766-021-4065-2>.

Supported by the Fundamental Research Funds for the Central Universities, CHD(300102129201), the Natural Science Basic Research Plan in Shaanxi Province of China (2018JM1011) and the National Natural Science Foundation of China(11701041).

*Corresponding author.

§1 Introduction

Coronavirus disease (COVID-19) [17, 18] infection has spread rapidly to other provinces and neighboring countries since the first case of pneumonia in Wuhan, China. COVID-19 is the third large-scale outbreak caused by highly pathogenic coronavirus. The first two outbreaks were "Severe Acute Respiratory Syndrome" (SARS) in mainland China in 2003 [3] and "Middle East Respiratory Syndrome" (MERS) in Saudi Arabia in 2012 [2]. COVID-19 is an infectious disease caused by a newly discovered corona-virus named the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus strain. As of the morning of February 17, 2020, according to the report of the National Health and Health Commission of the People's Republic of China [10], the total number of confirmed cases was 70639, the total number of death was 1772, the total number of cures was 11025, and the number of existing severe cases was 10644, and the epidemic was spread to many countries.

Mathematical modeling and simulations are helpful to reveal the epidemic law of infectious diseases, determine the key factors of transmission, and predict the epidemic trend, so as to provide theoretical basis and strategies for the prevention and control of infectious diseases. By the propagation mechanism and control strategy of COVID-19, many mathematical models are formulated to study its propagation law. For example, in [1], the transmission network models of bats, hosts, reservoirs and humans simulate the potential transmission from the source of infection to human infection. In [8], a prediction model for the potential domestic and international spread of the outbreak of COVID-19 was studied. In [14], based on the clinical progress of the disease, a deterministic SEIR model was proposed, and the epidemiological status and intervention measures of individuals were studied as well as the control reproduction number of the model was obtained by maximum likelihood estimation. In [11], Peng et al. established SEIRD model and studied the prevention and control measures of COVID-19. In [15], Tang et al. established a stochastic discrete transmission dynamic model. Then by the actual situation of the infected patients and population flow reported in Shaanxi Province, the epidemic transmission chain of Shaanxi Province was obtained. Most of the above papers are based on the data analysis and prediction of the infected population in Wuhan. The estimation of control reproduction number by mathematical model can help to determine the severity and the possibility of outbreak, and provide key information and some control strategies for the determination of disease. The combination of infectious disease dynamics, biostatistics and computer simulations can make the theoretical basis and prevention and control strategy more reliable and practical.

Shaanxi Province is located in the central region of China, bordering Hubei Province, which is the only way to reach the northeast and northwest from Wuhan. With a large population in Shaanxi Province, the epidemic has also been introduced to Shaanxi, and the epidemic has been further spread. Based on the above mentioned papers and the usual methods to study the spread of infectious diseases, this paper aims to study the situation of COVID-19 epidemic in Shaanxi Province, and provide support for the control of COVID-19 epidemic in Shaanxi Province.

§2 Epidemic model of general population

The early transmission of COVID-19 can be described by a *SEQIJR* model. Based on the model of [4], we will describe the transmission mechanism of COVID-19 and the meaning of specific parameters in detail, and then modify the contact rate in the model to a specific function to describe the transmission of COVID-19. The model consists of six sub populations: susceptible $S(t)$, asymptomatic $E(t)$, quarantined $Q(t)$, symptomatic $I(t)$, isolated $J(t)$ and recovered $R(t)$ individuals in a population of $N(t) = S(t) + E(t) + Q(t) + I(t) + J(t) + R(t)$ individuals. The main modeling ideas are as follows:

(i) $S(t)$: The susceptible population increased due to the net inflow of individuals and decreased due to natural death. The susceptible population will also decrease after infection, which is obtained through the contact between the susceptible and the infected. The infected may be symptomatic, asymptomatic, quarantined, or isolated. (ii) $E(t)$: Asymptomatic individuals have been exposed to the virus, but have no clinical symptoms of COVID-19. (iii) $Q(t)$: These are asymptotically infected individuals who are quarantined because of contact with the source of the COVID-19. For simplicity, we assume that all isolated individuals are asymptomatic and who will continue to show symptoms and then move to the isolated class. (iv) $I(t)$: The symptomatic population is produced after the asymptomatic members show the clinical symptoms of COVID-19. And it is diminished by isolation, death caused by disease, rehabilitation and natural death. (v) $J(t)$: These people have developed clinical symptoms and have been isolated through hospitalization and other means. These come from symptoms and isolation. This population is diminished by recovery, death from disease and natural death. (vi) $R(t)$: Symptomatic and isolated individuals recover from the disease, and this population is diminished by natural death. The detailed schematic diagram for the transmission of COVID-19 among the general population is shown in Figure 1.

Accordingly, we obtain the definitions of some parameters in Table 1 and the following *SEQIJR* model:

$$\left\{ \begin{array}{l} \frac{dS}{dt} = A - \frac{\beta_I SI}{N} - \frac{\beta_E SE}{N} - \frac{\beta_Q SQ}{N} - \frac{\beta_J SJ}{N} - \mu S, \\ \frac{dE}{dt} = B + \frac{\beta_I SI}{N} + \frac{\beta_E SE}{N} + \frac{\beta_Q SQ}{N} + \frac{\beta_J SJ}{N} - (\gamma_1 + k_1 + \mu)E, \\ \frac{dQ}{dt} = \gamma_1 E - (k_2 + \mu)Q, \\ \frac{dI}{dt} = k_1 E - (\gamma_2 + d_1 + \sigma_1 + \mu)I, \\ \frac{dJ}{dt} = \gamma_2 I + k_2 Q - (\sigma_2 + d_2 + \mu)J, \\ \frac{dR}{dt} = \sigma_1 I + \sigma_2 J - \mu R. \end{array} \right. \tag{1}$$

Let $N(t) = S(t) + E(t) + Q(t) + I(t) + J(t) + R(t)$ be the total population number at time t . It is clear that $N'(t) = A + B - \mu N - d_1 I + d_2 J$. Therefore, in the disease-free state, the

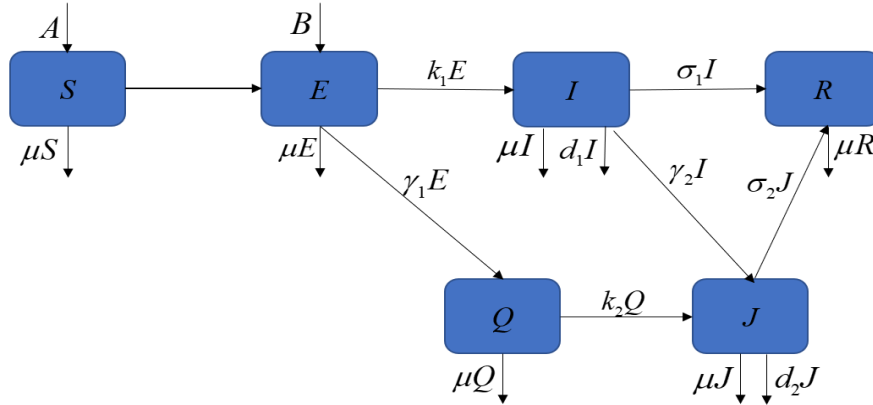


Figure 1. Schematic diagram for the transmission of COVID-19 among the general population.

Table 1. Definitions of some model parameters.

Parameters	Description
A	Recruitment rate of susceptible individuals
B	Recruitment rate of asymptomatic individuals
β_I	Adequate contact rate of an symptomatic individual with others
β_E	Adequate contact rate of an asymptomatic individual with others
β_Q	Adequate contact rate of an quarantined individual with others
β_J	Adequate contact rate of an isolated individual with others
μ	Natural death rate
γ_1	Isolation rate of asymptomatic individuals
γ_2	Isolation rate of symptomatic individuals
σ_1	Recovery rate of symptomatic individuals
σ_2	Recovery rate of isolated individuals
d_1	Disease-induced death rate of symptomatic individuals
d_2	Disease-induced death rate of isolated individuals
k_1	Moving rate from asymptomatic individuals to symptomatic individuals
k_2	Moving rate from quarantined individuals to isolated individuals

total population converges to $\frac{A+B}{\mu}$, that is, $\lim_{t \rightarrow \infty} N(t) = \frac{A+B}{\mu}$.

We adopt the next generation matrix [16] to get an expression for the control reproduction number when control measures are in force as follows (by linearizing the model around the disease-free equilibrium with $B = 0$)

$$\mathcal{R}_0 = \frac{\beta_E}{\gamma_1+k_1+\mu} + \frac{\beta_I k_1}{(\gamma_1+k_1+\mu)(\gamma_2+d_1+\sigma_1+\mu)} + \frac{\beta_Q \gamma_1}{(\gamma_1+k_1+\mu)(k_2+\mu)} + \frac{\beta_J k_1 \gamma_2}{(\gamma_1+k_1+\mu)(\gamma_2+d_1+\sigma_1+\mu)(\sigma_2+d_2+\mu)} + \frac{\beta_J \gamma_1 k_2}{(\gamma_1+k_1+\mu)(\sigma_2+d_2+\mu)(k_2+\mu)}. \tag{2}$$

It is worth noting that the adequate contact rate β_I of symptomatic individuals is constant when there is no artificial control, and when the artificial control starts at time τ , according to

the reference [7], β_I shows an exponential decay, which meets the following requirements

$$\beta_I(t) = \begin{cases} \beta_{I0}, & t < \tau, \\ \beta_{I0}e^{-m(t-\tau)}, & t \geq \tau, \end{cases} \tag{3}$$

where m is the intensity of control over symptomatic individuals.

In the next section, we summarize the reported data of COVID-19 in Shaanxi Province, then adopt the reported data to fit the model (1), and estimate the value of the parameters by least square method and MCMC. Specifically, we employ an adaptive Metropolis-Hastings (M-H) algorithm to carry out extensive Markov-chain Monte-Carlo (MCMC) simulations [5], and to estimate mean values of parameters including adequate contact rate, recruitment rate of susceptible individuals, recruitment rate of asymptomatic individuals and the intensity of control over symptomatic individuals.

§3 Numerical simulation

In this section, firstly, according to the reported data of COVID-19 from the Health Committee of Shaanxi Province, we make statistics on the cumulative number of infected people from January 23, 2020 to February 15, 2020 (see Figure 8 and Table 5 in the Appendix). Then we look up the parameters values with the same meaning as those in the model (1) in some recent COVID-19 research results as the basis of our research. These parameters include μ and τ , as well as initial values $S(0)$, $E(0)$, $I(0)$, $J(0)$ and $R(0)$. For the remaining parameters in model (1), including A , B , β_{I0} , β_E , β_Q , β_J , γ_1 , γ_2 , σ_1 , σ_2 , d_1 , d_2 , k_1 , k_2 and m , as well as the initial value $Q(0)$, we use statistical data to fit the model, and the fitting diagram is shown in Figure 2 and Figure 3. Specifically, the statistical data of the first nine days are used to fit the model to get the estimated value of parameter β_{I0} , as shown in Figure 2. The statistical data of 24 days are used to fit the model to get the estimated value of the remaining parameters, as shown in Figure 3. At the same time, we use MCMC simulations to estimate the average value of these parameters. The detailed parameter values and initial values are shown in Table 2, where the MCMC analysis of parameters β_E , γ_1 , γ_2 , k_1 and k_2 is shown in Figure 4. In numerical simulations, we consider the isolated individuals (including quarantined $Q(t)$ and isolated $J(t)$) through their own protective measures, such as wearing masks, so we assume $\beta_Q = \beta_J = 0$. The following theoretical analysis provides the reasonableness of the assumption.

Lemma 1 [21] Signalling game is a dynamic game with incomplete information. In this game, there are two participants, $i = 1, 2$. Among them, the participant 1 is the signal sender and the participant 2 is the signal receiver. The type of participant 1 is private information, and the type of participant 2 is public information (that is, there is only one type). The specific order of the game is as follows:

- (i) The type of participant 1 is $\theta \in \Theta$, where $\Theta = \{\theta_1, \dots, \theta_k\}$ is the type space of participant 1. Participant 1 knows his type, but participant 2 only knows that the prior probability of participant 1 belonging to θ is $p = p(\theta)$, and $\sum_k p(\theta^k) = 1$.
- (ii) After observing type θ , participant 1 sends signal $m \in M$, where $M = \{m^1, \dots, m^j\}$ is

the signal space.

(iii) When participant 2 observes signal m (but not type θ) sent by participant 1, participant 2 uses Bayes rule to obtain a posterior probability $\tilde{p} = \tilde{p}(\theta|m)$ from prior probability $p = p(\theta)$, and then makes action $a \in A$, where $A = \{a^1, \dots, a^h\}$ is the action space of participant 2. In statistics, the probability before correction is called "prior probability", and the probability after correction is called "posterior probability".

(iv) The payment functions are $u_1(m, a, \theta)$ and $u_2(m, a, \theta)$, respectively. (The payment function is the utility level that participants get from the game).

Lemma 2 [21] The refined Bayesian equilibrium of signalling game includes strategic combination $(m^*(\theta), a^*(m))$ and a posteriori probability $\tilde{p}(\theta|m)$, which satisfies the following conclusions:

$$(i) a^*(m) \in \arg \max_a \sum_{\theta} \tilde{p}(\theta|m) u_2(m, a, \theta);$$

$$(ii) m^*(\theta) \in \arg \max_m u_1(m, a^*(m), \theta);$$

(iii) Participant 2 considers the prior probability $p(\theta)$, the observed signal m and the optimal strategy $m^*(\theta)$ of participant 1, and then according to the Bayes rule, participant 2 gets $\tilde{p}(\theta|m)$.

Lemma 3 [21] All the possible refined Bayesian equilibria of signalling game can be divided into three categories. One of the special equilibria is the semi-separation equilibrium. Semi-separation equilibrium means that some types of senders (participant 1) randomly select signals, while others select specific signals.

Next, we divide the whole population into two categories: the free group and the isolated group (including quarantined $Q(t)$ and isolated $J(t)$). By Lemma 1 and Lemma 2, we get that free group (θ^1) have two choices, i.e. no mask and wear mask, and denote m^1 and m^2 respectively; while the isolated group (θ^2) can only choose to wear mask (m^2) with a probability of 1. By Lemma 3, the case belongs to semi-separating equilibrium. From Lemma 1 and Lemma 2, the isolated people will wear masks. When the free people see the people wearing masks, they will react far away from the people wearing masks. That is to say, the spread of the epidemic among the masked people and the free people will be eliminated, that is, the isolated people and the free people can not be contacted. In detail, according to the payment functions and posteriori probability of the two lemmas, the analysis is as follows:

$$\begin{aligned} u_1(m^1, a^*(m), \theta^1) &= u_1(m^2, a^*(m), \theta^1), \\ u_1(m^1, a^*(m), \theta^2) &< u_1(m^2, a^*(m), \theta^2), \\ \tilde{p}(\theta^1 | m^1) &= \frac{\alpha \times p(\theta^1)}{\alpha \times p(\theta^1) + 0 \times p(\theta^2)} = 1, \\ \tilde{p}(\theta^1 | m^2) &= \frac{(1 - \alpha) \times p(\theta^1)}{(1 - \alpha) \times p(\theta^1) + 1 \times p(\theta^2)} < p(\theta^1), \\ \tilde{p}(\theta^2 | m^2) &= \frac{1 \times p(\theta^2)}{(1 - \alpha) \times p(\theta^1) + 1 \times p(\theta^2)} > p(\theta^2), \end{aligned}$$

here, α is the probability that individuals in the free group choose not to wear masks. Here we mainly focus on the first and third expressions in the above formula and we further get that

m^2 must be the best choice for the isolated group. As a rational person, an isolated individual must wear a mask, so we assume that $\beta_Q = \beta_J = 0$ is reasonable.

Table 2. Parameters and numerical values chosen for the simulation.

Parameters	Description	Value (day ⁻¹)	Source
A	Recruitment rate of susceptible individuals	75.6599	MCMC
B	Recruitment rate of asymptomatic individuals	30.7003	MCMC
β_{I0}	Adequate contact rate of an symptomatic individual with others	0.5	MCMC
β_E	Adequate contact rate of an asymptomatic individual with others	0.0287	MCMC
β_Q	Adequate contact rate of an quarantined individual with others	0	Assumption
β_J	Adequate contact rate of an isolated individual with others	0	Assumption
μ	Natural death rate	1/(75.7 × 365)	[12]
γ_1	Isolation rate of asymptomatic individuals	0.0798	MCMC
γ_2	Isolation rate of symptomatic individuals	0.1053	MCMC
σ_1	Recovery rate of symptomatic individuals	0.0001	MCMC
σ_2	Recovery rate of isolated individuals	0.1788	MCMC
d_1	Disease-induced death rate of symptomatic individuals	0.000001989	MCMC
d_2	Disease-induced death rate of isolated individuals	0.0098	MCMC
k_1	Moving rate from asymptomatic individuals to symptomatic individuals	0.1032	MCMC
k_2	Moving rate from quarantined individuals to isolated individuals	0.2198	MCMC
m	Intensity of control over symptomatic individuals	0.1827	MCMC
τ	The delay time between the start time of artificial control and the time of symptomatic individuals contacting others	9	[6]

Initial values	Description	Value	Source
$S(0)$	Susceptible individuals	21,781,500	[13]
$E(0)$	Asymptomatic individuals	53	[6]
$Q(0)$	Quarantined individuals	600	Assumption
$I(0)$	Symptomatic individuals	3	[6]
$J(0)$	Isolated individuals	3	[6]
$R(0)$	Recovered individuals	0	[6]

Sensitivity analysis is vital to identify key parameters and find effective control strategies for combatting the spread of the disease. It is well known that the control reproduction number \mathcal{R}_0 is a very important parameter in the infectious disease model, which determines whether the epidemic could spread. In model (1), we focus on the parameters $\beta_I, \beta_E, \beta_Q, \beta_J, d_1, d_2, \sigma_1, \sigma_2, \gamma_1, \gamma_2$ in \mathcal{R}_0 . In order to identify the impacts of these parameters on COVID-19 transmission and prevalence, we used the Latin hypercube sampling method and partial rank correlation coefficient (PRCC) (see [9]). From model (1), 2000 samples are randomly generated by assuming a uniform distribution for each parameter based on values from Table 2. We choose all parameters as the input variables, and the value of \mathcal{R}_0 as the output variable. The PRCC

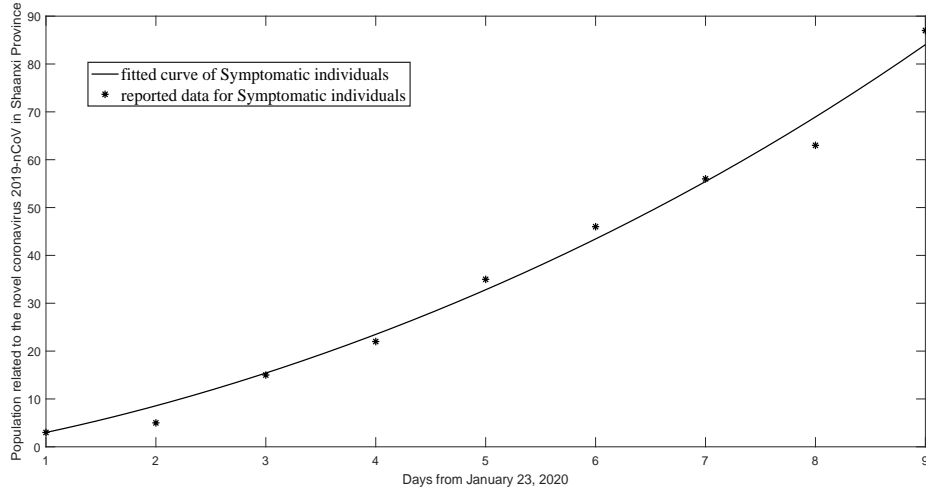


Figure 2. The fitting between the cumulative number of COVID-19 case from January 23, 2020 to January 31, 2020 and the simulation of model (1). And the purpose of this fitting is to estimate β_{I0} without artificial control, which is equivalent to satisfying $m = 0$ in (3).

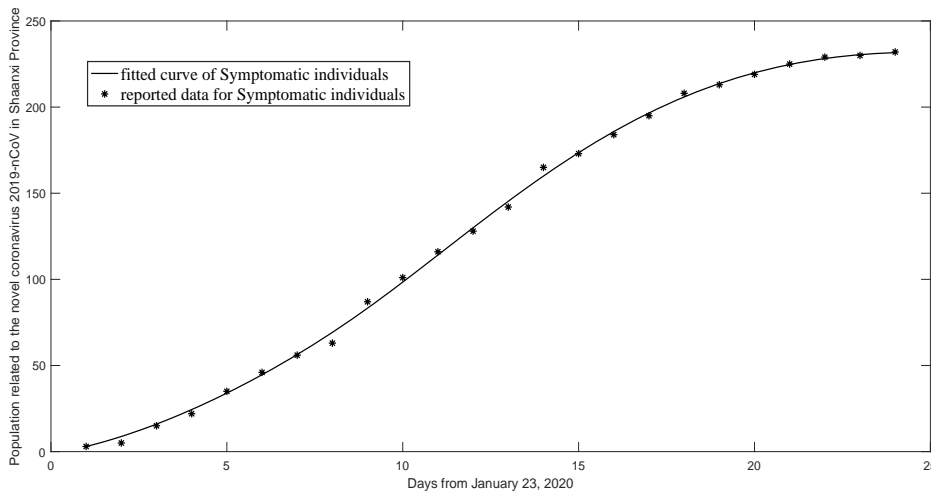


Figure 3. The fitting between the cumulative number of COVID-19 case from January 23, 2020 to February 15, 2020 and the simulation of model (1). Specifically, from January 23 to January 31, we think that $\beta_I(t) = \beta_{I0}$, and artificial control has been started after January 31, that is, from February 1 to February 15, there is $\beta_I(t) = \beta_{I0}e^{-m(t-\tau)}$, where $m = 0.1827$, $\tau = 9$ days.

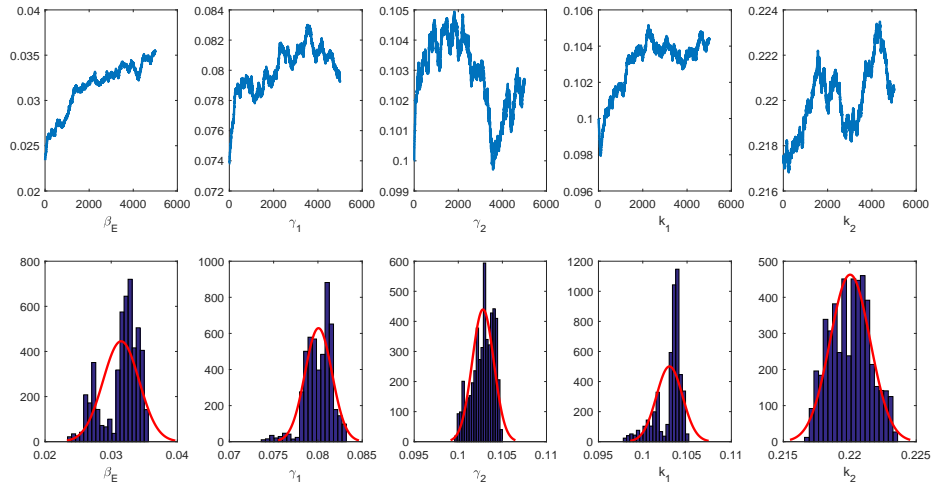


Figure 4. This is MCMC analysis of parameters β_E , γ_1 , γ_2 , k_1 and k_2 . The first row is the random sequence of five parameters, and the second row is the histogram of MCMC corresponding to five parameters.

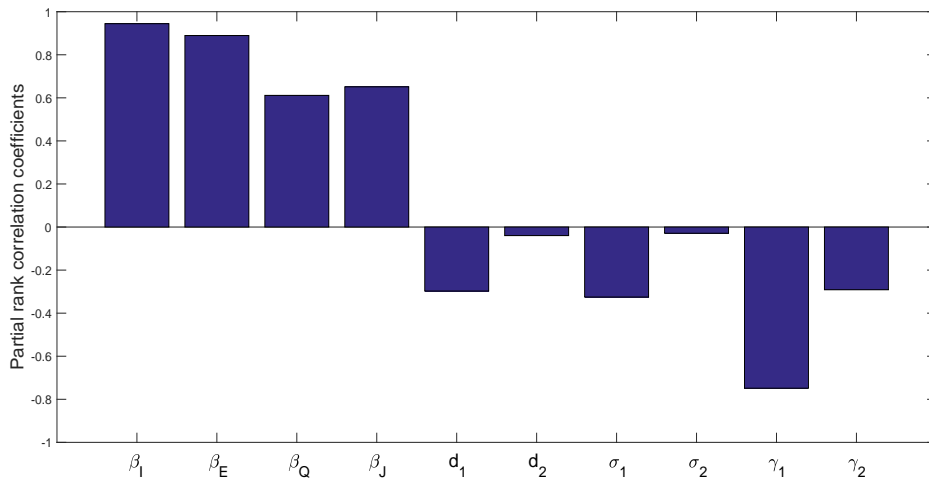
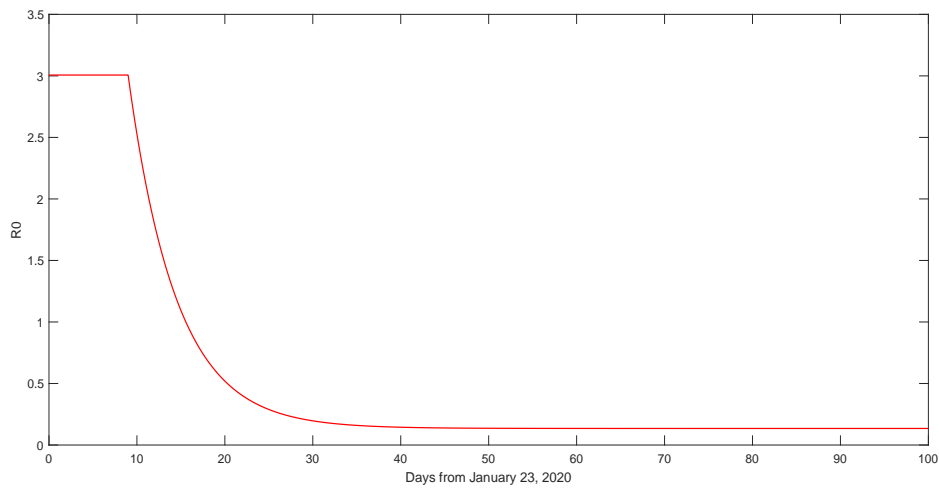


Figure 5. The Partial rank correlation coefficient values for model (1).

Table 3. Partial rank correlation coefficients (PRCCs) for \mathcal{R}_0 and each input parameter.

Input parameter	PRCC	p value
β_I	0.9458	0
β_E	0.8894	0
β_Q	0.6310	< 0.0001
β_J	0.6875	< 0.0001
d_1	-0.3295	< 0.0001
d_2	-0.0567	0.0113
σ_1	-0.3443	< 0.0001
σ_2	-0.0628	0.0051
γ_1	-0.7533	0
γ_2	-0.2705	< 0.0001

Figure 6. We consider that β_I shows an exponential decay, then \mathcal{R}_0 changes with time t .

values of all parameters and corresponding p -values are computed (see Table 3 and Figure 5). The larger PRCCs in absolute value, the more important the parameter in responding to the change in \mathcal{R}_0 . Plus sign or minus sign means the influence is positive or negative respectively. Table 3 and Figure 5 show that β_I , β_E , β_Q and β_J have positive impact upon \mathcal{R}_0 , whilst d_1 , d_2 , σ_1 , σ_2 , γ_1 and γ_2 have negative impact. Hence, from sensitivity we conclude that the most effective approach to reduce the COVID-19 infection is to reduce the parameters β_I , β_E , β_Q and β_J , and increase σ_1 , σ_2 , γ_1 and γ_2 .

In the following, firstly, we focus on β_I shows an exponential decay, so \mathcal{R}_0 changes with time t . According to the parameters values in Table 2, we get the change curve of \mathcal{R}_0 (as shown in Figure 6). We can see from Figure 5 that the control reproductive number will remain unchanged at 3 when $t < \tau$, and gradually decrease from 3 to slightly greater than 0.2 as time increases after $t \geq \tau$.

Next, based on the general population epidemic model (1), we establish the college population and general population epidemic model, and give the control reproduction number.

§4 Epidemic model of college populations and general populations

As a result of the outbreak of COVID-19, schools across the country are still not open as of mid March 2020. Xi'an is the capital of Shaanxi Province, where there are more than 50 universities. At present, when the universities start and the specific prevention and control strategies after the beginning of the universities have been the topic of many people's attention. Therefore, it is necessary for us to study the infectious disease model of college populations and general populations. Based on the paper of Xiao et al. [20], we will embed the general population model (1) into their model. We know that it is impossible to completely isolate the flow between schools and the general populations, so the topic of concern is how the intensity of school closure, the time of school closure and school populations flow affect the spread of COVID-19 for the whole populations. First of all, we assume that the individual in the school is allowed to leave the campus for a period of time and then return, and the general population is not allowed to enter the school. Without losing generality, we assume that the average proportions of individual students exposed to the virus leaving school to enter the general populations is h (school leaving rate).

Table 4. Definitions of some model parameters.

Parameters	Description
β	Propagation probability of contact once
c_u	Contact rate of college population (day^{-1})
c_g	Contact rate of general population (day^{-1})
q_u	Isolation rate of college population (day^{-1})
q_g	Isolation rate of general population (day^{-1})
ϕ_u	The efficiency of personal prevention of college population
ϕ_g	The efficiency of personal prevention of general population
λ	The rate at which isolated susceptible individuals return to susceptible class
δ_1	Migration rate of asymptomatic individuals (day^{-1})
δ_2	Diagnostic rate (day^{-1})
σ_1	Removal rate of infected individuals (day^{-1})
σ_2	Removal rate of treatment individuals (day^{-1})
h	Leaving rate of college population

For college population at time t , let $S_u(t)$, $E_u(t)$, $I_u(t)$ and $R_u(t)$ denote the susceptible population size, the asymptomatic population size, the symptomatic population size and recovered population size, respectively. Some of the symptomatic individuals will be isolated and treated in hospital, i.e. class H_u . Individuals exposed to the virus, in which a proportion of q_u individuals are isolated before they become infected. The asymptomatic individuals in the isolated population will enter S_u^q , and the symptomatic individuals will enter the isolated latent class E_u^q , after which the class E_u^q individuals will enter the hospital for treatment, that is, H_u^q . Individuals in S_u^q return to S_u after a period of isolation, which is usually the longest incubation

period of this infectious disease. For general population at time t , the biological meaning of $S_g(t)$, $E_g(t)$, $Q_g(t)$, $I_g(t)$, $J_g(t)$ and $R_g(t)$ is the same as $S(t)$, $E(t)$, $Q(t)$, $I(t)$, $J(t)$ and $R(t)$ in model (1). And the main modeling ideas are as follows: (i) For college population, we con-

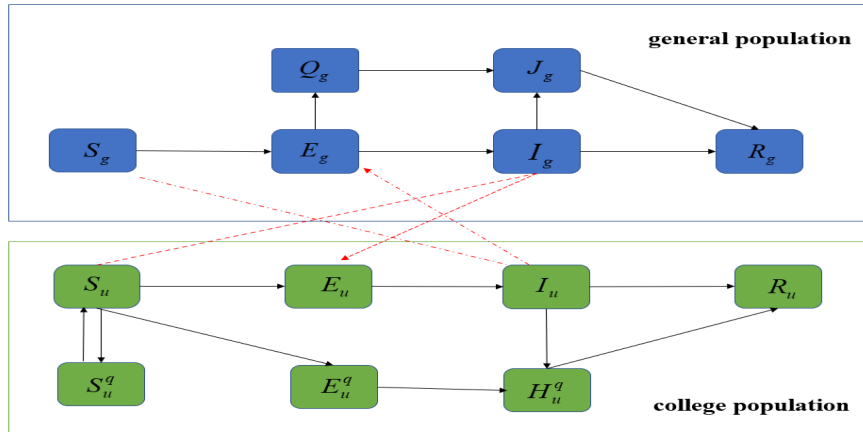


Figure 7. Schematic diagram for the transmission of COVID-19 among general population and college population.

sider isolation of close contacts, isolation of confirmed cases and personal epidemic prevention strategies. If a susceptible individual S_u exposed to COVID-2019 virus is infected, the individual enters latent E_u , then enters symptomatic I_u after the incubation period, until recovery or death enters removed R_u . (ii) When individuals exposed to the virus take self-protection measures, such as wearing masks and personal hygiene, we assume that susceptible people with the proportion of φ_u are protected from infection. In particular, some parameters in the model are denoted as $N_u = (1 - h)S_u + (1 - h)E_u + (1 - h)I_u + S_u^q + E_u^q + H_u^q + (1 - h)R_u$, $N_g = S_g + E_g + Q_g + I_g + J_g + R_g + hS_u + hE_u + hI_u + hR_u$, N_u and N_g denote the total number of college populations and general populations, respectively. And $P_{S_u} = (1 - h)^2[\beta c_u(1 - \varphi_u) + c_u q_u(1 - \beta)]$, $P_{E_u} = (1 - h)^2 \beta c_u(1 - q_u)(1 - \varphi_u)$, $Q_{S_u} = (1 - h)^2(1 - \beta)c_u q_u$, $Q_{E_u} = (1 - h)^2 \beta c_u q_u(1 - \varphi_u)$, $P_{S_g} = (1 - h)[\beta c_g(1 - \varphi_g) + c_g q_g(1 - \beta)]$, $P_{E_g} = (1 - h)\beta c_g(1 - q_g)(1 - \varphi_g)$, $Q_{S_g} = (1 - h)(1 - \beta)c_g q_g$, $Q_{E_g} = (1 - h)\beta c_g q_g(1 - \varphi_g)$, where P_{S_u} is the infection rate of I_u to S_u , P_{E_u} is the infection rate of I_u to E_u , Q_{S_u} is the infection rate of I_u to S_u^q , Q_{E_u} is the infection rate of I_u to E_u^q , P_{S_g} is the infection rate of I_g to S_u , P_{E_g} is the infection rate of I_g to E_u , Q_{S_g} is the infection rate of I_g to S_u^q , and Q_{E_g} is the infection rate of I_g to E_u^q . In addition, compared with model (1), it is obvious that $\beta_I = \beta c_g(1 - \varphi_g)$.

The biological meaning for the parameters are given in Table 4, and the flow diagram between the variables is shown in Figure 7, so the following model is established:

$$\left\{ \begin{aligned}
 \frac{dS_u}{dt} &= -\frac{P_{Su}S_uI_u}{N_u} - \frac{P_{Sg}S_uI_g}{N_g} + \lambda S_u^q, \\
 \frac{dE_u}{dt} &= \frac{P_{Eu}S_uI_u}{N_u} + \frac{P_{Eg}S_uI_g}{N_g} - \delta_1 E_u, \\
 \frac{dI_u}{dt} &= \delta_1 E_u - (\delta_2 + \sigma_1)I_u, \\
 \frac{dS_u^q}{dt} &= \frac{Q_{Su}S_uI_u}{N_u} + \frac{Q_{Sg}S_uI_g}{N_g} - \lambda S_u^q, \\
 \frac{dE_u^q}{dt} &= \frac{Q_{Eu}S_uI_u}{N_u} + \frac{Q_{Eg}S_uI_g}{N_g} - \delta_1 E_u^q, \\
 \frac{dH_u^q}{dt} &= \delta_2 I_u + \delta_1 E_u^q - \sigma_2 H_u^q, \\
 \frac{dR_u}{dt} &= \sigma_1 I_u + \sigma_2 H_u^q, \\
 \frac{dS_g}{dt} &= A - \frac{\beta_I S_g(I_g + hI_u)}{N_g} - \frac{\beta_E S_g E_g}{N_g} - \frac{\beta_Q S_g Q_g}{N_g} - \frac{\beta_J S_g J_g}{N_g} - \mu S_g, \\
 \frac{dE_g}{dt} &= B + \frac{\beta_I S_g(I_g + hI_u)}{N_g} + \frac{\beta_E S_g E_g}{N_g} + \frac{\beta_Q S_g Q_g}{N_g} + \frac{\beta_J S_g J_g}{N_g} - (\gamma_1 + k_1 + \mu)E_g, \\
 \frac{dQ_g}{dt} &= \gamma_1 E_g - (k_2 + \mu)Q_g, \\
 \frac{dI_g}{dt} &= k_1 E_g - (\gamma_2 + d_1 + \sigma_1 + \mu)I_g, \\
 \frac{dJ_g}{dt} &= \gamma_2 I_g + k_2 Q_g - (\sigma_2 + d_2 + \mu)J_g, \\
 \frac{dR_g}{dt} &= \sigma_1 I_g + \sigma_2 J_g - \mu R_g.
 \end{aligned} \right. \tag{4}$$

Then we obtain the disease-free equilibrium $E_0 = (S_u(0), 0, 0, 0, 0, 0, 0, S_g(0), 0, 0, 0, 0, 0)$ of model (4). By the concepts of next generation matrix and reproduction number presented in [16], we get the the basic reproduction number is

$$\begin{aligned}
 \mathcal{R}_0^{ug} &= \max \left\{ A_1, \frac{(A_1 + A_4) + \sqrt{A_1^2 + A_4^2 - 2A_1A_4 + 4A_2A_3}}{2} \right\} \\
 &= \frac{(A_1 + A_4) + \sqrt{A_1^2 + A_4^2 - 2A_1A_4 + 4A_2A_3}}{2}.
 \end{aligned} \tag{5}$$

where

$$\begin{aligned}
 A_1 &= \frac{P_{Eu}}{(1-h)(\delta_2 + \sigma_1)}, \quad A_2 = \frac{P_{Eg}S_u(0)k_1}{(S_g(0) + hS_u(0))(\gamma_1 + k_1 + \mu)(\gamma_2 + d_1 + \sigma_1 + \mu)}, \\
 A_3 &= \frac{\beta_I h S_g(0)}{(S_g(0) + hS_u(0))(\delta_2 + \sigma_1)}, \quad A_4 = \frac{\beta_E S_g(0)}{(S_g(0) + hS_u(0))(\gamma_1 + k_1 + \mu)} \\
 &+ \frac{\beta_I S_g(0)k_1}{(S_g(0) + hS_u(0))(\gamma_1 + k_1 + \mu)(\gamma_2 + d_1 + \sigma_1 + \mu)} + \frac{\beta_Q S_g(0)\gamma_1}{(S_g(0) + hS_u(0))(k_2 + \mu)(\gamma_1 + k_1 + \mu)} \\
 &+ \frac{\beta_J S_g(0)(\gamma_1\gamma_2k_2 + \gamma_2k_1k_2 + \gamma_1k_2\sigma_1 + \gamma_2k_1\mu + \gamma_1k_2\mu + \gamma_1k_2d_1)}{(S_g(0) + hS_u(0))(k_2 + \mu)(\gamma_1 + k_1 + \mu)(\mu + d_2 + \sigma_2)(\gamma_2 + d_1 + \sigma_1 + \mu)}.
 \end{aligned}$$

In particular, when the school leaving rate $h = 0$, we get $A_3 = 0$ and

$$\begin{aligned}
 A_4 &= \frac{\beta_E}{(\gamma_1 + k_1 + \mu)} + \frac{\beta_I k_1}{(\gamma_1 + k_1 + \mu)(\gamma_2 + d_1 + \sigma_1 + \mu)} + \frac{\beta_Q \gamma_1}{(k_2 + \mu)(\gamma_1 + k_1 + \mu)} \\
 &\quad + \frac{\beta_J (\gamma_1 \gamma_2 k_2 + \gamma_2 k_1 k_2 + \gamma_1 k_2 \sigma_1 + \gamma_2 k_1 \mu + \gamma_1 k_2 \mu + \gamma_1 k_2 d_1)}{(k_2 + \mu)(\gamma_1 + k_1 + \mu)(\mu + d_2 + \sigma_2)(\gamma_2 + d_1 + \sigma_1 + \mu)} \\
 &= \mathcal{R}_0,
 \end{aligned}$$

where \mathcal{R}_0 is the control reproduction number of the general population model (1), and so we get

$$\mathcal{R}_0^{ug} = \frac{(A_1 + A_4) + \sqrt{A_1^2 + A_4^2 - 2A_1A_4}}{2} = \max \{A_1, A_4\} = \max \left\{ \frac{P_{Eu}}{(1-h)(\delta_2 + \sigma_1)}, \mathcal{R}_0 \right\}. \tag{6}$$

Specifically, the college population model is as follows

$$\left\{ \begin{aligned}
 \frac{dS_u}{dt} &= -\frac{P_{Su}S_uI_u}{N_u} + \lambda S_u^q, \\
 \frac{dE_u}{dt} &= \frac{P_{Eu}S_uI_u}{N_u} - \delta_1 E_u, \\
 \frac{dI_u}{dt} &= \delta_1 E_u - (\delta_2 + \sigma_1)I_u, \\
 \frac{dS_u^q}{dt} &= \frac{Q_{Su}S_uI_u}{N_u} - \lambda S_u^q, \\
 \frac{dE_u^q}{dt} &= \frac{Q_{Eu}S_uI_u}{N_u} - \delta_1 E_u^q, \\
 \frac{dH_u^q}{dt} &= \delta_2 I_u + \delta_1 E_u^q - \sigma_2 H_u^q, \\
 \frac{dR_u}{dt} &= \sigma_1 I_u + \sigma_2 H_u^q,
 \end{aligned} \right. \tag{7}$$

we can get that the control reproduction number of the college population model (7) is $\mathcal{R}_0^u = \frac{P_{Eu}}{(1-h)(\delta_2 + \sigma_1)}$.

From the above analysis, when the school leaving rate h is 0, it follows that the college populations and the general populations are isolated from each other, simultaneously, the model (4) becomes two independent models (1) and (7).

§5 Discussion

Based on the transmission characteristics of COVID-19 epidemic in Shaanxi Province, on the one hand, we consider that both susceptible and asymptomatic individuals with potential contact have external input, especially the population input from Wuhan region, on the other hand, we take into account the policy and medical situation of Shaanxi Province, the government publicizes that individuals with potential contact are isolated for 14 days by themselves, and the symptomatic individuals are hospitalized for observation and treatment in time. So we set up the *SEQIJR* model to describe the epidemic situation in Shaanxi Province. In the light of the COVID-19 infection reported by Health and Health Committee of Shaanxi Province, we

counted the confirmed cases from January 13, 2020 to February 15, 2020, and then we used the statistical data to fit the model (1). Some parameters of the model can be found in some official websites and academic papers, and the average values of other parameters are obtained by least square method and MCMC.

From the analysis of the epidemic model (1) of the general population, the control reproduction number of the epidemic spread in Shaanxi Province is obtained, that is, the control reproduction number remained at 3 from January 23 to January 31, then gradually decreased from 3 to slightly greater than 0.2. Then we discuss some biological implications and focus on the impact of some key model parameters, and combine with sensitivity analysis, the strategies of controlling COVID-19 are given, i.e., (i) reducing the infection rate, in particular, measures such as wearing masks, and restrict travel and paying attention to personal hygiene can be taken, and under special circumstances, urban blockade strategy is adopted; (ii) increasing isolation of individuals exposed to the virus and treatment of symptomatic individuals and (iii) increasing medical level and medical equipment to improve treatment rate. In Section 4, considering the actual problems of when and how to prevent and control the epidemic in colleges, based on the epidemic model of general population, a high-dimensional epidemic model of general population and college population is established, and the control reproduction number is given, which will provide a theoretical basis for the prevention and control of infectious diseases in colleges and universities.

§6 Appendix

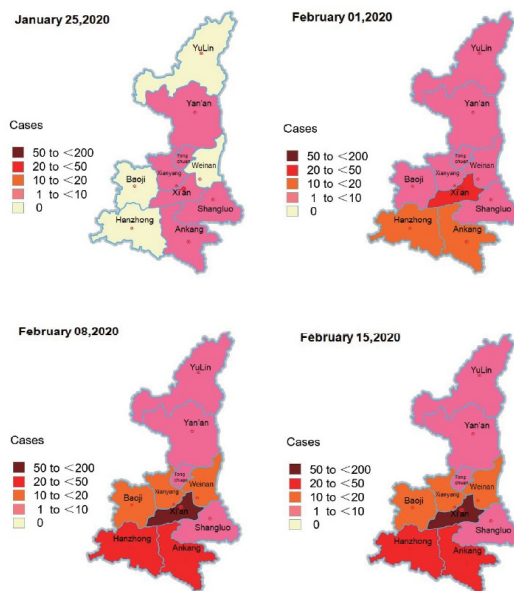


Figure 8. Heat-map showing the spreading of COVID-19 infection in Shaanxi Province.

Table 5. Cumulative total of reported the COVID-19 cases in Shaanxi Province of China [6].

Date	1.23	1.24	1.25	1.26	1.27	1.28	1.29	1.30
Shaanxi Province	3	5	15	22	35	46	56	63
Date	1.31	2.1	2.2	2.3	2.4	2.5	2.6	2.7
Shaanxi Province	87	101	116	128	142	165	173	184
Date	2.8	2.9	2.10	2.11	2.12	2.13	2.14	2.15
Shaanxi Province	195	208	213	219	225	229	230	232

References

- [1] T Chen, J Rui, Q Wang, et al. *A mathematical model for simulating the transmission of Wuhan novel Coronavirus*, bioRxiv, 2020, <https://doi.org/10.1101/2020.01.19.911669>.
- [2] G Chowell, S Blumberg, L Simonsen, et al. *Synthesizing data and models for the spread of MERS-CoV, 2013: key role of index cases and hospital transmission*, *Epidemics*, 2014, 9: 40-51.
- [3] D Feng, S J De Vlas, L Fang, et al. *The SARS epidemic in mainland China: bringing together all epidemiological data*, *Tropical Medicine and International Health*, 2009, 14: 4-13.
- [4] A B Gumel, S Ruan, T Day, et al. *Modelling strategies for controlling SARS outbreaks*, *Proceedings of the Royal Society B: Biological Sciences*, 2004, 271(1554): 2223-2232.
- [5] H Haario, M Laine, A Mira, et al. *DRAM: efficient adaptive MCMC*, *Statistics and Computing*, 2006, 16(4): 339-354.
- [6] Health Committee of Shaanxi Province, Available from: <http://sxwjw.shaanxi.gov.cn/col/col2/index.html>.
- [7] P E Lekone, B F Finkenstädt. *Statistical inference in a stochastic epidemic SEIR model with control intervention: Ebola as a case study*, *Biometrics*, 2006, 62(4): 1170-1177.
- [8] G Luo, M L Mchenry, J J Letterio. *Estimating the prevalence and risk of COVID-19 among international travelers and evacuees of Wuhan through modeling and case reports*, *PLOS ONE*, 2020, 15.
- [9] S Marino, I B Hogue, C J Ray, et al. *A methodology for performing global uncertainty and sensitivity analysis in systems biology*, *Journal of Theoretical Biology*, 2008, 254(1): 178-196.
- [10] National Health and Health Commission of the Peoples Republic of China, Available from: <http://www.nhc.gov.cn/xcs/yqfkdt/202002/18546da875d74445bb537ab014e7a1c6.shtml>.
- [11] S Peng, Y Shan. *Beware of asymptomatic transmission: Study on 2019-nCoV prevention and control measures based on extended SEIR model*, bioRxiv, 2020, <https://doi.org/10.1101/2020.01.28.923169>.
- [12] Shaanxi Provincial Bureau of Statistics, Available from: <http://tjj.shaanxi.gov.cn/site/1/html/126/131/138/18679.htm>.
- [13] Shaanxi Provincial People's Government, Available from: <http://www.shaanxi.gov.cn/sxxw/sxyw/115352.htm>.

- [14] B Tang, X Wang, Q Li, et al. *Estimation of the transmission risk of the 2019-nCoV and its implication for public health interventions*, Journal of Clinical Medicine, 2020, 9(2): 462.
- [15] S Tang, B Tang, N L Bragazzi, et al. *Stochastic discrete epidemic modeling of COVID-19 transmission in the Province of Shaanxi incorporating public health intervention and case importation*, medRxiv, 2020, <https://doi.org/10.1101/2020.02.25.20027615>.
- [16] P Van den Driessche, J Watmough. *Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission*, Mathematical Biosciences, 2002, 180(1-2): 29-48.
- [17] World Health Organization, COVID-19, Available from: <https://www.who.int/csr/don/12-january-2020-novel-coronavirus-china/en>.
- [18] World Health Organization, COVID-19, Available from: <https://www.who.int/news-room/detail/12-02-2020-world-experts-and-funders-set-priorities-for-covid-19-research>.
- [19] P Wu, X Hao, E H Y Lau, et al. *Real-time tentative assessment of the epidemiological characteristics of novel coronavirus infections in Wuhan, China, as at 22 January 2020*, Eurosurveillance, 2020, 25(3).
- [20] Y Xiao, Y Zhou, S Tang. *Principles of Biomathematics*, Xi'an Jiaotong University Press, 2012.
- [21] W Zhang. *Game theory and information economics*, Shanghai People Press, 2004.

¹School of Science, Chang'an University, Xi'an 710064, China.

²School of Ecological Environment, Northwestern Polytechnical University, Xi'an 710072, China.

³College of Economics and Management, Shanxi Normal University, Linfen 041004, China.

Email: t.l.zhang@126.com