

COVID-19 Propagation Prediction and Assessment Method with Imported Cases and Infection Generations: Shanxi Province as a Case

Juan Zhang

Shanxi university

Gui-Quan Sun

Shanxi university

Mingtao Li

Taiyuan University of Technology

Wenyi Zhang

Center for Disease control and prevention of PLA

Lili Chang

Shanxi University

Zhongwei Jia

Peking University

Ricardo J Soares Magalhães

The university of queensland

Juping Zhang

Shanxi University

Huarong Ren

Shanxi University

Xin Pei

Taiyuan University of Technology

Xiaochun Cao

Shanxi University

Fenfen Zhang

North University of China

Zi-Ke Zhang

Zhejiang University

Guirong Liu

Shanxi Univeristy

Meiping Yao

Shanxi University

Zhen Jin (✉ jinzhn@263.net)

Research article

Keywords: COVID-19, Integral model, Propagation chain, Duration time, Final scale, Isolation assessment

Posted Date: March 5th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-16298/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

COVID-19 propagation prediction and assessment method with imported cases and infection generations: Shanxi province as a case

Juan Zhang^{a,b}, Gui-Quan Sun^{a,b}, Mingtao Li^c, Wenyi Zhang^d,
Lili Chang^{a,b}, Zhongwei Jia^e, Ricardo J Soares Magalhães^f,
Juping Zhang^{a,b}, Huarong Ren^{a,g}, Xin Pei^c, Xiaochun Cao^{a,b},
Fenfen Zhang^{h,i}, Zi-Ke Zhang^{j,k}, Guirong Liu^g, Meiping Yao^g,
Zhen Jin^{a,b,*},

^a*Complex Systems Research Center, Shanxi University, Taiyuan, Shan'xi 030006, China*

^b*Shanxi Key Laboratory of Mathematical Techniques and Big Data Analysis on Disease Control and Prevention, Taiyuan, Shan'xi, 030006, China*

^c*College of Mathematics, Taiyuan University of Technology, Taiyuan, Shanxi 030024, China*

^d*Center for Disease Control and Prevention of PLA, Beijing 100071, China.*

^e*National Institute of Drug Dependence, Peking University, Beijing 100871, China*

^f*School of Veterinary Science, The University of Queensland, Gatton Qld 4343, Australia*

^g*School of Mathematical Sciences, Shanxi University, Taiyuan, Shanxi 030006, China*

^h*Data Science And Technology, North University of China, Taiyuan, Shan'xi 030051, China*

ⁱ*Applied Mathematics, North University of China, Shuozhou, Shan'xi 036000, China*

^j*College of Media and International Culture, Zhejiang University, Hangzhou 310028, China*

^k*Alibaba Research Center for Complexity Sciences, Hangzhou Normal University, Hangzhou 311121, China*

Abstract

When everyone focuses on 2019 coronavirus disease (COVID-19) in Hubei province, the epidemic in other province cannot be ignored, which also has an impact on the epidemic in the whole country. The most distinctive epidemic characteristic in

all regions except Wuhan is that the most of confirmed cases are imported cases from Wuhan, and the propagation chain is relatively clear. Based on detailed contact tracing information of confirmed cases, combined with first-order outbreak response measures, we establish a disease transmission dynamical model to describe the infection propagation chain among the human population. Using Shanxi province as a case, modeling results indicate that the epidemic peak in Shanxi province occurred in February 2. In addition, our model suggests that according to the current development trend, COVID-19 will disappear in February with the final epidemic number of approximately 175 cases. It is verified that the most effective outbreak control measures in Shanxi include home isolation of people, surveillance and isolation of second-generation cases, contact tracing and management of contacts. With the end of the holiday, if the average number of contacts per person per day is less than 6, it has little impact on the incidence of COVID-19, and even if third- and fourth-generation cases occur, the epidemic will be under control, no later than late March with a final outbreak size of 220 cases. However, if the average number of contacts per person per day is greater than 6, the number of COVID-19 cases will continue to be reported resulting in another epidemic peak. Through the forecast and evaluation of COVID-19 in Shanxi, it is verified that the model with infection generations is more accurate to describe the spread mode and can be extended to regions with imported cases.

Key words: COVID-19, Integral model, Propagation chain, Duration time, Final scale, Isolation assessment

INTRODUCTION

On December 8th, 2019, the first case of unexplained pneumonia was reported in Wuhan city, Hubei province of the Republic of China. Due to long incubation period(95% confidence interval[CI], 4.1 to 7.0, even longer) before clinical symptoms [1], with the travel and homecoming of individuals who travel to affected areas by airplane and high-speed rail, the new coronavirus is coming fast and widespread. From January 13, confirmed cases started to appear in other countries and other Chinese provinces, respectively. The high infectivity and household clustering of COVID-19 infections raised a lot of health concern and panic among people. From January 25, all provincial government initiated a first-level emergency response. These included mandatory registration of all people arriving from Wuhan city, their family members and related contacts, and the registered individuals with home quarantine for at least 14 days, suspension of bus and taxi services, closure of all entrances and exits of

* Corresponding author.
Email address: jinzhn@263.net (Zhen Jin).

expressways leading to the city surveillance for new COVID-19 cases and contact tracing and management of contacts. In addition, all grass-roots organizations, government organs, enterprises and institutions needed to strengthen personnel management and health monitoring services. All of these measures are designed to improve early detection, reporting, quarantine and treatment of COVID-19 cases. Issued on January 26, State Council issued a directive for the Spring Festival holiday to be extended until February 10.

With the series of prevention and control measures implemented locally, the outbreak in all province except Hubei province seems to be under control in February 10. However, with the end of the Spring Festival, people started to work which means that contact between people in the workplace and public transport may result in the increase in risk of transmission of SARS-CoV-2. As a result a number of questions arise including how many COVID-19 cases are being missed and not isolated? How much impact did the end of the holiday have on the incidence of SARS-CoV-2 transmission? Will a second outbreak peak occur as a result for the people resuming their work duties after the Spring Festival? To adequately answer these questions, mathematical models can be constricted to describe the transmission dynamics of COVID-19. As we know, the elementary segments of epidemic of infectious disease includes three aspects: the source of infection, transmission routes and characteristics of the susceptible population. For Shanxi province, the source of infection are mainly imported cases from Wuhan city (which is called as the first-generation cases), and the main transmission route is close contact of imported cases with local people (which belongs to the second-generation cases). So, different from the traditional modeling idea, we establish a dynamics model with infection generations to describe the propagation chain of COVID-19. In addition, outbreak control interventions such as surveillance for new cases and contact tracing and management of contacts can also be well reflected in the model.

Disease dynamics models originate from the general Kermack-McKendrick model (SIR, SIS) and has been extended to various model types and specifications. Dynamical model is based on transmission mechanism of infectious diseases and addressed the limitation of small sample data allowing the parameterization of infectious disease transmission to predict the trend of the infection and assess the impact of prevention and control measures. A number of models have been put forward which have modelled the transmission dynamics of COVID-19 [2–4]. These models established differential equations that assumed that the random variable of the residence time of an individual in a compartment is exponentially distributed. In fact, by observing actual COVID-19 case data, we note that the normal, or log-normal or gamma distributions are more appropriate. Therefore, there is a need to develop modeling frameworks that consider the residence time as a normally distributed parameter and present an integration modeling method that accounts for the spread of SARS-CoV-2 between infection generations.

In this study, using reported COVID-19 cases data from Shanxi province, we establish an integration model to forecast the epidemic trend, peak time, peak number of COVID-19 cases, and estimate the duration of the epidemic. In addition, we aim to assess the effect of outbreak response interventions including surveillance, tracing and isolation of confirmed cases and their contacts on these estimates. By above research in Shanxi province, we verify the rationality and efficiency of the model.

MATERIALS AND METHOD

Materials

Data source: The official reported confirmed cases and their detailed track information [5]. Family population obtained from National Bureau of Statistics [6].

Prerequisites

- a) Since the number of confirmed cases is very small compared to the total human population in Shanxi province, the susceptible group is not considered in model.
- b) Once human are infected by the SARS-CoV-2 virus, they may undergo several days before clinical symptoms, but they has infection during this period. So, it is assumed that once human are infected, they are infectious.
- c) It has been known that infected individuals can spread the virus without showing any symptoms. Asymptomatic infection after the incubation period is not considered in model.

Dynamical Model

In this paper, we take infectious cases as research objects. Once human are infected by the SARS-CoV-2 virus, they first undergo an incubation period with several days. After incubation period, when showing clinical symptoms, people will spend some days on home medication before going to the hospital. When going to hospital, they need spent one to three days to take Nucleic acid test to be confirmed. According to whether or not they are identified and confirmed, we divide the infectious cases into two classes: non-confirmed cases and confirmed cases. The non-confirmed cases refer to the population who has been infected and infectious, but has not been confirmed by medical

institutions. After taking Nucleic acid test and being confirmed, the non-confirmed cases become confirmed cases. Once confirmed, the cases will be being treated in isolation in hospital and separate from the infection of the epidemic. The number of non-confirmed cases at time t is denoted by $I(t)$. According to the propagation chain, they are divided into three subclasses: the first-generation, the second-generation and the third-generation, whose numbers at time t are denoted by $I_1(t)$, $I_2(t)$, $I_3(t)$. The first-generation non-confirmed cases refer to the imported non-confirmed cases from Wuhan city. The second-generation non-confirmed cases refer to local non-confirmed infected by the first-generation cases, and they can be divided into two parts: family non-confirmed cases and non-family non-confirmed cases. Family non-confirmed cases are the infected but no confirmed family members of the first-generation cases. Non-confirmed non-family cases refer to population that are generally infected in public places or public transport by the first-generation non-confirmed cases. The third-generation non-confirmed cases refer to local cases infected but not confirmed by the second-generation non-confirmed cases, which has not been reported in Shanxi province. For the first-generation non-confirmed cases, the import number at time τ is $A(\tau)$, who are confirmed at time t with probability $G(t - \tau)$. The daily number of close contacts per first-generation non-confirmed case is $B = C_1 + C_2$, where C_1 is the contact number with family members, C_2 is the contact number with non-family members. λ is the probability of post-exposure infection. At time τ , the import number of the second-generation non-confirmed cases is $\lambda B I_1(\tau)$. For family member and non-family member who are infected at time τ , their confirmed probability at time t are $F_1(t - \tau)$ and $F_2(t - \tau)$, respectively. Since all people from Wuhan city and their family members were complied with home quarantine for at least 14 days, the effective contact number of the family members with other people is 0. The daily contact number of each non-family member with other people, whose proportion is $\sigma = \frac{C_2}{C_1 + C_2}$, is C_3 . If exists, the import number of the third-generation non-confirmed cases at time τ is $\lambda \sigma C_3 I_2(\tau)$, who are confirmed at time t with probability $F_3(t - \tau)$. The cumulative numbers of the first-generation confirmed cases, the second-generation confirmed cases and the third-generation confirmed cases at time t are denoted by $Q_1(t)$, $Q_2(t)$ and $Q_3(t)$, respectively. The transmission of the virus between these subpopulations can be seen in Fig.1, and the corresponding dynamical model is system (1). Where, $G(t_1)$, $F_1(t_2)$, $F_2(t_3)$, $F_3(t_4)$ are accumulation normal distribution functions of the individual residence time t_1, t_2, t_3, t_4 . $t_1 \sim N(\mu_1, \sigma_1)$, $t_2 \sim N(\mu_2, \sigma_2)$, $t_3 \sim N(\mu_3, \sigma_3)$, $t_4 \sim N(\mu_4, \sigma_4)$. $A(t)$ is a piecewise function.

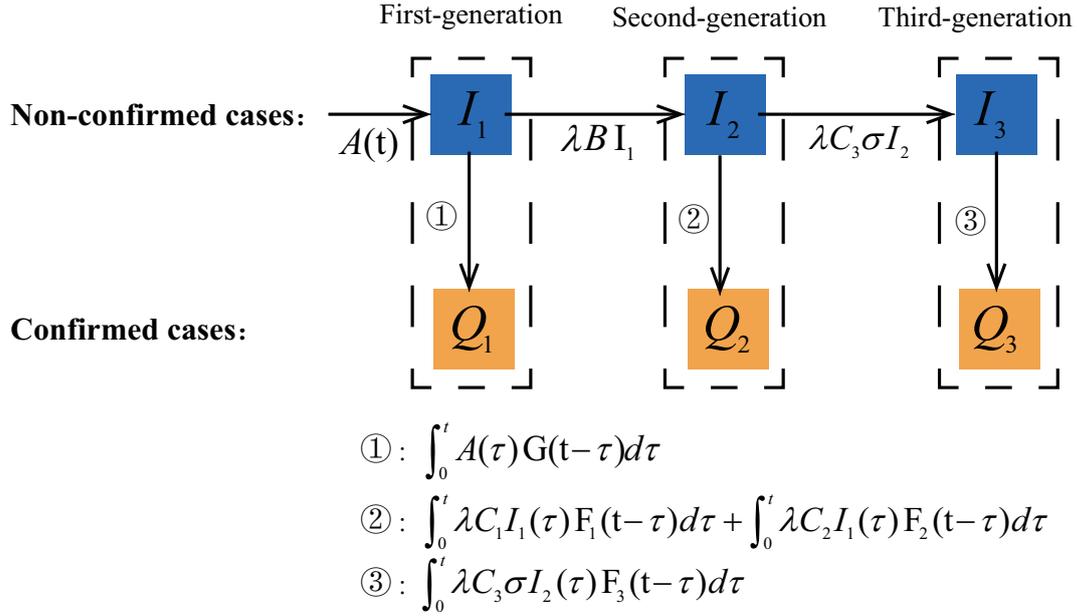


Fig. 1. Transfer diagram on the dynamical transmission of coronavirus among sub-populations.

$$\begin{cases} I_1(t) = I_1(0) + \int_0^t A(\tau)[1 - G(t-\tau)]d\tau, \\ I_2(t) = I_2(0) + \int_0^t \lambda C_1 I_1(\tau)[1 - F_1(t-\tau)]d\tau + \int_0^t \lambda C_2 I_1(\tau)[1 - F_2(t-\tau)]d\tau, \\ I_3(t) = I_3(0) + \int_0^t \lambda C_3 \sigma I_2(\tau)[1 - F_3(t-\tau)]d\tau, \\ Q_1(t) = Q_1(0) + \int_0^t A(\tau)G(t-\tau)d\tau, \\ Q_2(t) = Q_2(0) + \int_0^t \lambda C_1 I_1(\tau)F_1(t-\tau)d\tau + \int_0^t \lambda C_2 I_1(\tau)F_2(t-\tau)d\tau, \\ Q_3(t) = Q_3(0) + \int_0^t \lambda C_3 \sigma I_2(\tau)F_3(t-\tau)d\tau. \end{cases} \quad (1)$$

Parameter values

In this subsection, we give parameter values in model (1) (see Table 1) and describe the implementation process of parameter estimation, which is implemented by the function *fminsearch* in the optimization toolbox in MATLAB.

a) We apply actual time of arrival of the first-generation non-confirmed cases from February 14 to 25, to give $A(t)$ (see Fig.2).

b) The time from the arrival time in Shanxi province to confirmed time is tested to obey normal distribution, by analyzing which we can obtain the 95% confidence interval of μ_1 and σ_1 . By applying $Q_1(t)$ to fit the first-generation cumulative confirmed cases, we estimate the values of μ_1 and σ_1 . The fitting result can be seen in Fig.3.

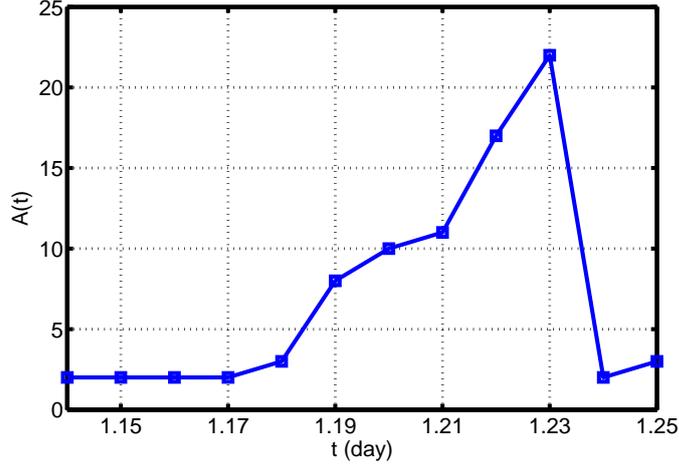


Fig. 2. $A(t)$ with time.

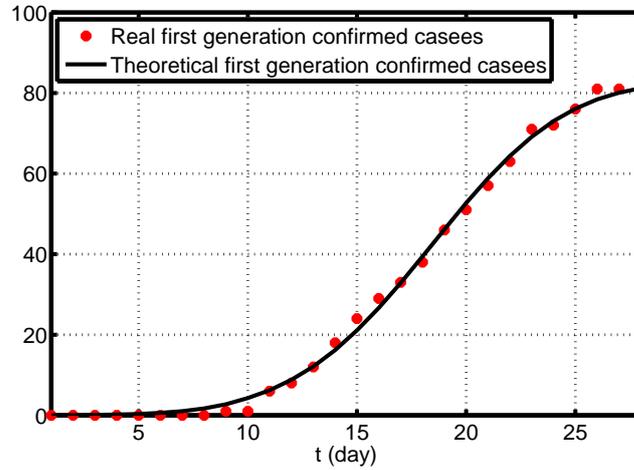


Fig. 3. The fitting result of the first-generation cumulative confirmed cases. Where the red dots are actual reported cases from January 14 to February 10. Black solid line is the theoretical cases of $Q_1(t)$.

c) C_1 is obtained by analyzing family members in Shanxi province. μ_2 ($\mu_2 = \mu_3$), σ_2 are obtained by analyzing time from the contact time of the second-generation cases with the first-generation cases to confirmed time, which obeys normal distribution. By applying $\int_0^t \lambda C_1 I_1(\tau) F_1(t - \tau) d\tau$ to fit the second-generation cumulative confirmed family cases, we obtain the value of λ and adjust the value of σ_2 . The fitting result can be seen in Fig.4.

d) By applying $\int_0^t \lambda C_2 I_1(\tau) F_2(t - \tau) d\tau$ to fit the second-generation cumulative confirmed non-family cases, we obtain the mean value of σ_3 and C_3 . The fitting result can be seen in Fig.5.

e) By applying $Q(t) = Q_1(t) + Q_2(t)$ to fit the total cumulative confirmed cases, we slightly adjust the mean value of λ from 0.0129 to 0.0149. The fitting result

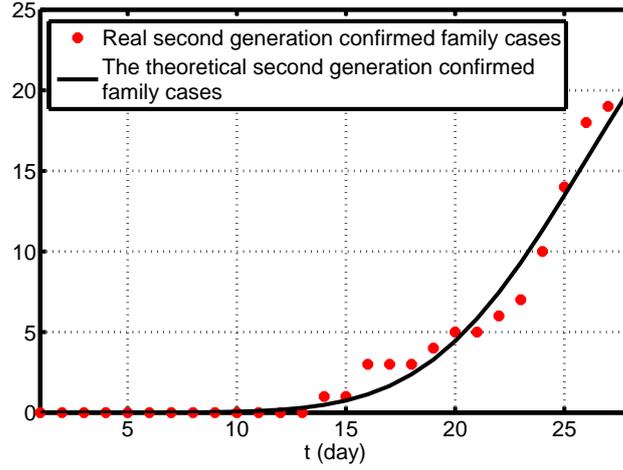


Fig. 4. The fitting result of the second-generation cumulative confirmed family cases. Where the red dots are actual reported cases from January 14 to February 10. Black solid line is the theoretical cases of $\int_0^t \lambda C_1 I_1(\tau) F_1(t - \tau) d\tau$.

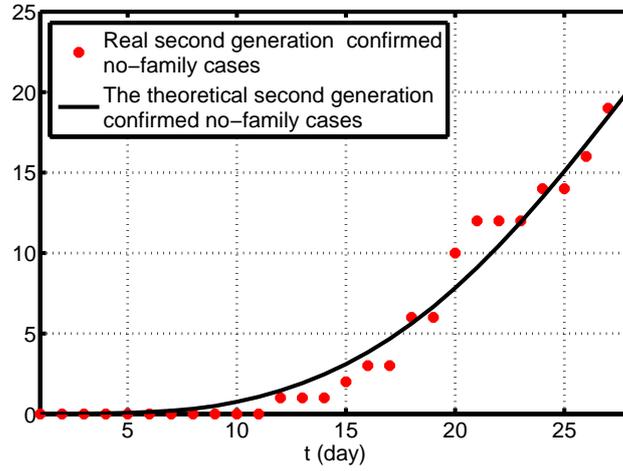


Fig. 5. The fitting result of the second-generation cumulative confirmed non-family cases. Where the red dots are actual reported cases from January 14 to February 10. Black solid line is the theoretical cases of $\int_0^t \lambda C_2 I_1(\tau) F_2(t - \tau) d\tau$.

can be seen in Fig.6.

RESULTS

Based on above parameter values, the prospect of the epidemic in Shanxi province can be seen in Fig.7 and Table 2. The inflection point (peak time) of the epidemic has occurred in February 2. The disease will disappear in

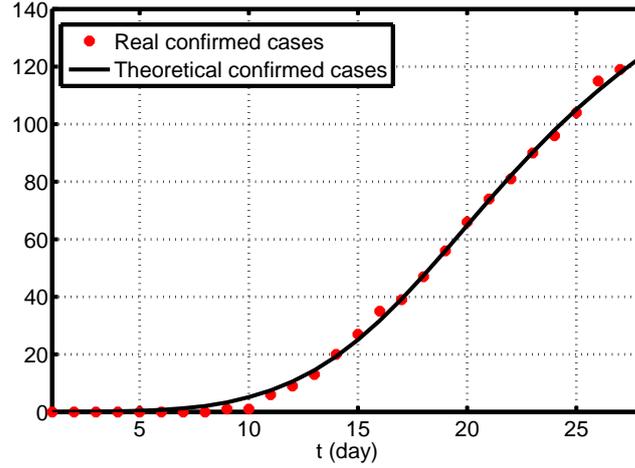


Fig. 6. The fitting result of the total cumulative confirmed cases. Where the red dots are actual reported cases from January 14 to February 10. Black solid line is the theoretical cases of $Q(t)$.

Table 1

Values of parameters and variables in model (1).

| Parameters | Value | 95% confidence interval | resource |
|------------|---------|-------------------------|---------------|
| $A(t)$ | / | / | Real data |
| μ_1 | 11.1410 | [9.4257, 11.5382] | Estimated |
| σ_1 | 4.3896 | [4.1968, 5.7102] | Estimated |
| μ_2 | 15.75 | [13.5139, 17.9861] | Dada analysis |
| σ_2 | 3.7703 | [2.4931, 5.9756] | Estimated |
| μ_3 | 15.75 | [13.5139, 17.9861] | Dada analysis |
| σ_3 | 8.2934 | / | Estimated |
| λ | 0.0149 | / | Estimated |
| C_1 | 2.9431 | [2.9169, 2.9693] | Dada analysis |
| C_2 | 3.5919 | / | Estimated |
| C_3 | 0 | / | Assumed |
| $I_1(t)$ | 1 | / | Real data |
| $I_2(t)$ | 0 | / | Real data |
| $I_3(t)$ | 0 | / | Real data |
| $Q_1(t)$ | 0 | / | Real data |
| $Q_2(t)$ | 0 | / | Real data |
| $Q_3(t)$ | 0 | / | Real data |

late February or the beginning of March and the mean final scale of cases will reach 175. It is inferred that there are about 20 non-confirmed family cases and about non-confirmed non-family cases that have not yet been found and confirmed, respectively. According to the current surveillance and control measures, the first-generation non-confirmed cases and second-generation non-confirmed family cases can be easily tracked and isolated. However, the situation of the second-generation non-confirmed non-family cases is complicated, and it's hard to track them all. Therefore, 20 non-confirmed non-family cases will be current risk source.

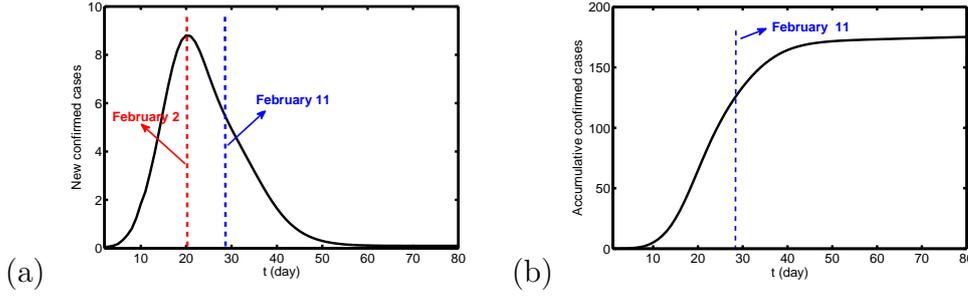


Fig. 7. The prediction result of the total confirmed cases in current situation. (a) New cases. (b) Cumulative cases.

Table 2

Prediction result of model (1).

| Index | Value |
|---------------------------------|-----------------|
| Peak time | 20(February 2) |
| Peak value of new cases | 9 |
| Final scale of cumulative cases | 175 |
| Duration time | 44(February 26) |

On January 10, Chinese new year break was over. Some groups had to start to work. In this case, due to the existence of non-confirmed non-family cases, the contact between individuals in workspace and public transport may increase the risk of transmission of the epidemic. How much impact does it have on the incidence of disease? If it causes third-generation infections, will there be a secondary outbreak? We estimate the prospect of the epidemic with the end of holiday and having third-generation infection, considering that the second-generation non-confirmed non-family cases have different contact numbers with other people (see Fig.8 and Table 3). Let $\sigma = \frac{C_2}{C_1 + C_2} = 0.5498$, $\mu_3 = \mu_2, \sigma_3 = \sigma_2$. It is concluded that it has no effect on the peak time and peak value. When the contact number increases by 2, the final scale will increase 6. When the contact number increases to 10, the mean final scale will reach 208,

and the duration of the epidemic will extend to the middle of March. So, as long as the second-generation non-confirmed cases are managed well, the end of holiday has little effect of the progress of the disease.

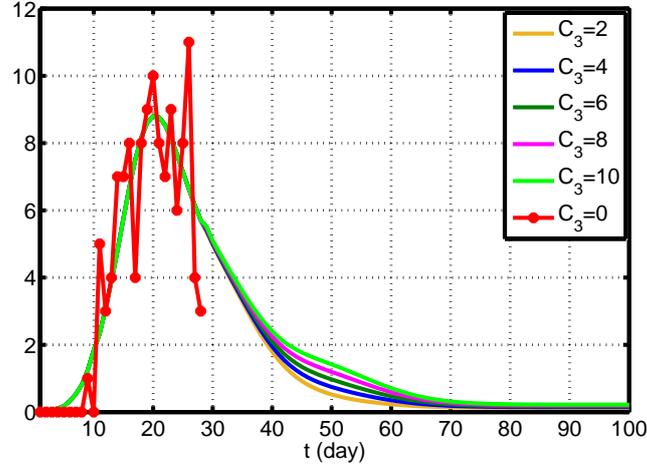


Fig. 8. The prediction result of the total cumulative confirmed cases with the end of holiday.

Table 3

Prediction result of model (1) with the end of holiday.

| Cases | Peak time | Peak value(new case) | Final scale | Duration time |
|------------------|----------------|----------------------|-------------|---------------|
| $C_3 = 2$ | 20(February 2) | 9 | 183 | 51(March 3) |
| $C_3 = 4$ | 20(February 2) | 9 | 189 | 56(March 8) |
| $C_3 = 6$ | 20(February 2) | 9 | 196 | 60(March 12) |
| $C_3 = 8$ | 20(February 2) | 9 | 202 | 62(March 14) |
| $C_3 = 10$ | 20(February 2) | 9 | 208 | 64(March 16) |
| $C_3 = C_4 = 2$ | 20(February 2) | 9 | 199 | 53(March 5) |
| $C_3 = C_4 = 4$ | 20(February 2) | 9 | 220 | 66(March 18) |
| $C_3 = C_4 = 6$ | 20(February 2) | 9 | / | / |
| $C_3 = C_4 = 8$ | 20(February 2) | 9 | / | / |
| $C_3 = C_4 = 10$ | 20(February 2) | 9 | / | / |

If there appears the third- and fourth-generation infection, let $\sigma = \frac{C_2}{C_1 + C_2} = 0.5498$, $\mu_4 = \mu_3 = \mu_2$, $\sigma_4 = \sigma_3 = \sigma_2$. As long as the quarantine measures were strictly enforced such that the mean contact number C_3, C_4 are both less than 6, even if there appears the third-generation cases and the fourth-generation cases, the epidemic will be under control, and its end time can be no later than late March, and the mean final scale is no more than 220 (see Fig. 9).

If the mean contact number per person per day is more than 6, the disease cannot disappear in the short term and can cause another peak. With this, more generations of cases need to be considered in model.

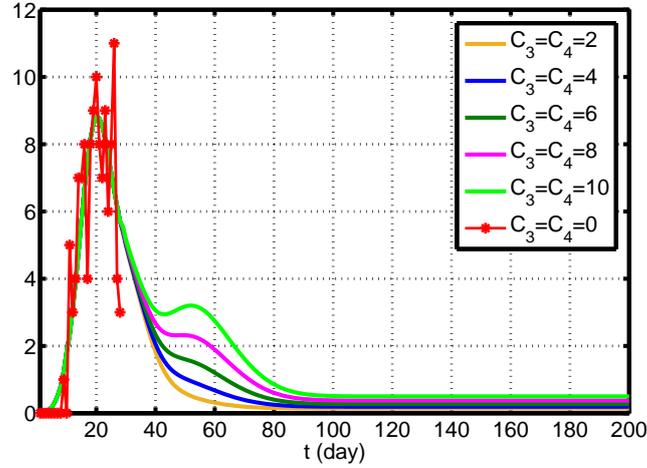


Fig. 9. The theoretical result of the total cumulative confirmed cases with having the third-generation and fourth-generation infection.

DISCUSSION AND CONCLUSION

In this section, we discuss the epidemic trend of Shanxi province under different conditions.

Case I: More than half of the first cases from Wuhan arrived in Shanxi after January 21. If the first-level emergency response was moved up from January 25 to January 22, and the individuals from Wuhan city were absolutely isolated from family members and other people, the epidemic will weaken by half (see Fig.10 and Table 4). Our results demonstrate that the epidemic will disappear in January 31 with the final size of 85 cases.

Case II: If people from Wuhan city are allowed to live with their family members, but are absolutely isolated from other people ($C_2 = 0$), the epidemic will disappear in February 20 with the mean final size of 124 cases(see Fig.11). If isolation measures are not enforced, first-generation non-confirmed cases have no stronger isolation awareness and freely enter public places. Assuming that the SARS-CoV-2 infection can only be spread to second-generation, with different contact numbers of the first-generation cases with non-family members, the trend of disease can be seen in Fig.9. When contact number increases by 2, the peak time has little change, and the peak value will increase 1, and the mean final size of the outbreak will increase 20-30 cases. If the contact number

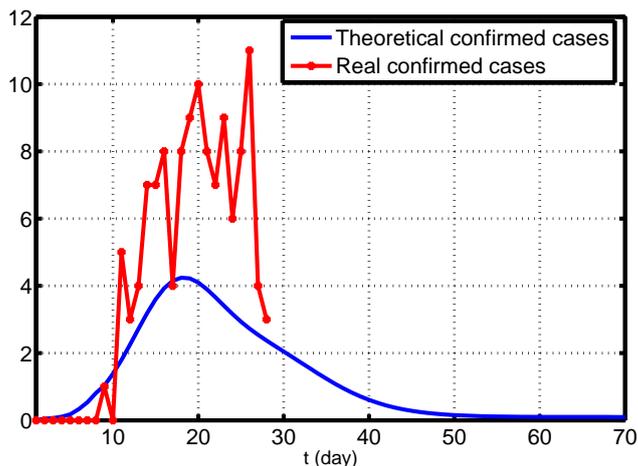


Fig. 10. The theoretical result of the total cumulative confirmed cases with Case I.

Table 4

Prediction result of model (1).

| Cases | Peak time | Peak value(new case) | Final scale | Duration time | |
|------------|----------------|----------------------|-------------|-----------------|-----------------|
| Case I | 18(January 31) | 4 | 85 | 37(February 19) | |
| $C_2 = 0$ | 20(February 2) | 7 | 124 | 38(February 20) | |
| $C_2 = 2$ | 20(February 2) | 8 | 153 | 41(February 23) | |
| Case II | $C_2 = 4$ | 20(February 2) | 9 | 182 | 44(February 26) |
| $C_2 = 6$ | 21(February 3) | 10 | 211 | 46(February 28) | |
| $C_2 = 8$ | 21(February 3) | 11 | 239 | 47(February 29) | |
| $C_2 = 10$ | 21(February 3) | 11 | 268 | 49(March 1) | |

increases to 10, that is the degree of activity of first-generation non-confirmed cases increases to 2.5 times as the actual situation, the mean final size will be approximately 270 cases and the duration of the epidemic will extend to the beginning of March.

Our results demonstrate that the early detection, early reporting and early quarantine of second- generation cases is key for effective outbreak control. With second-generation non-confirmed cases be well managed, if third-generation cases only appear, the effect of end of holiday is small. If fourth-generation cases also appear, contact tracing and isolation measures should be strictly performed to decrease the number of contacts per person per day be less than 6, otherwise the number of COVID-19 cases will continue to be reported resulting in another epidemic peak. Overall our results can provide theoretical guidance to understand the development and control of the disease. More importantly, our model provides important insights into the impact of imported cases while accounting for the relatively small data volume and can be extended to all provinces except Hubei province. Since asymptomatic infection

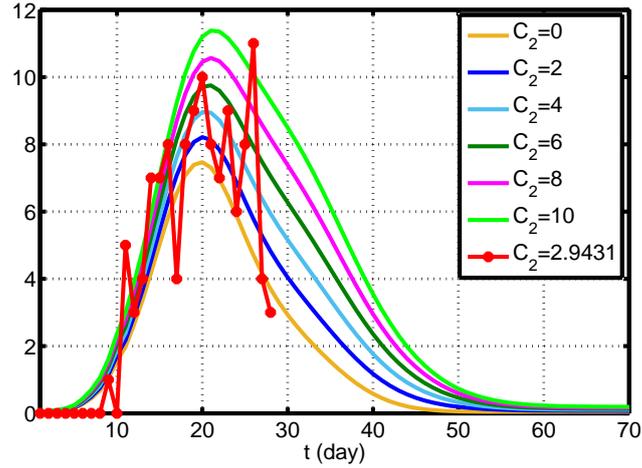


Fig. 11. The theoretical result of the total cumulative confirmed cases with Case II. after the incubation period is hard to estimate and its exact mechanism is not clear, so it is not considered in model.

DECLARATIONS

Ethics approval and consent to participate: Not applicable.

Availability of data and materials: Not applicable.

Consent for publication: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Availability of data and materials: The datasets used during the current study are available from the corresponding author on reasonable request.

Authors' contributions: Juan Zhang, Guiquan Sun, Mingtao Li and Zhen Jin conceived and designed the experiments; Juan Zhang, Wenyi Zhang and Lili Chang performed the experiments; Juping Zhang, Huarong Ren, Xin Pei, Xiaochun Cao and Fenfen Zhang analyzed the data; Zi-Ke Zhang, Guirong Liu and Meiping Yao contributed analysis tools; Juan Zhang , Zhongwei Jia , Ricardo J Soares Magalhães and Zhen Jin wrote the paper.

Acknowledgments

This work is supported by the National Natural Science Foundation of China under Grant (11601292, 61873154,11671241,11701348,11801398,11971278), key research and development project in Shanxi Province (202003D31011/GZ) , the National key research and development project of China (2019YFC1200500, 2019YFC1200501).

References

- [1] Li Q, M.Med., Guan X, Ph.D., Wu P, Ph.D., Wang X, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *The New England Journal of Medicine*. January 29, 2020.
- [2] Tang B, Wang X, Li Q, et al. Estimation of the transmission risk of 2019-nCoV and its implication for public health interventions. SSRN. 2020.
- [3] Zhao S, Ran J, Yang G, Lou Y, Gao D, et al. Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak. *bioRxiv*. Posted online Jan. 24, 2020. DOI: <http://dx.doi.org/10.1101/2020.01.23.916395>.
- [4] Shen M, Peng Z, Xiao Y, Zhang L. Modelling the epidemic trend of the 2019 novel coronavirus outbreak in China. *biorxiv*. DOI: <https://doi.org/10.1101/2020.01.23.916726>.
- [5] Health commission of Shanxi province. <http://wjw.shanxi.gov.cn/>.
- [6] National Bureau of Statistics of China. *China Statistical Yearbook*. 2019. <http://www.stats.gov.cn/tjsj/ndsj/2019/indexch.htm>.

Figures

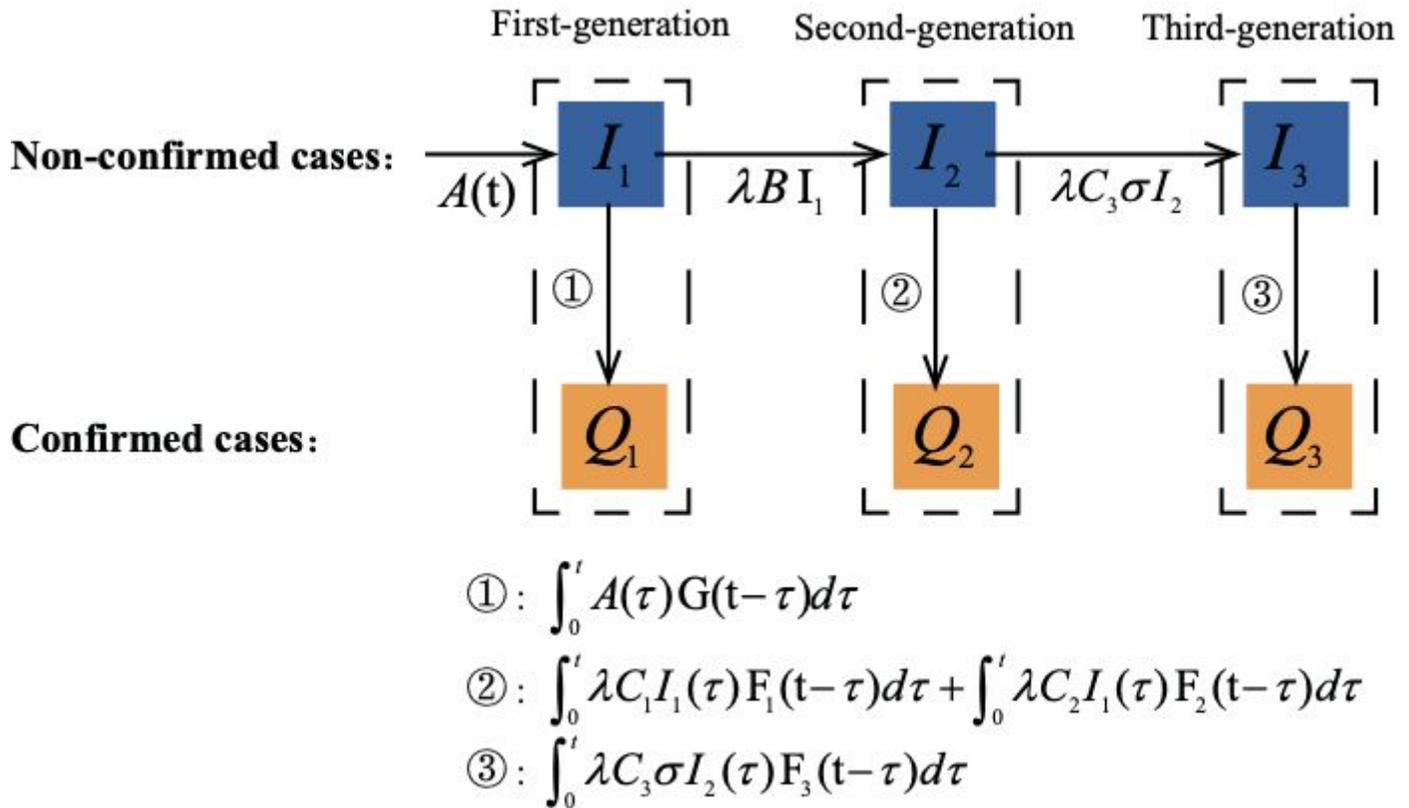


Figure 1

Transfer diagram on the dynamical transmission of coronavirus among subpopulations.

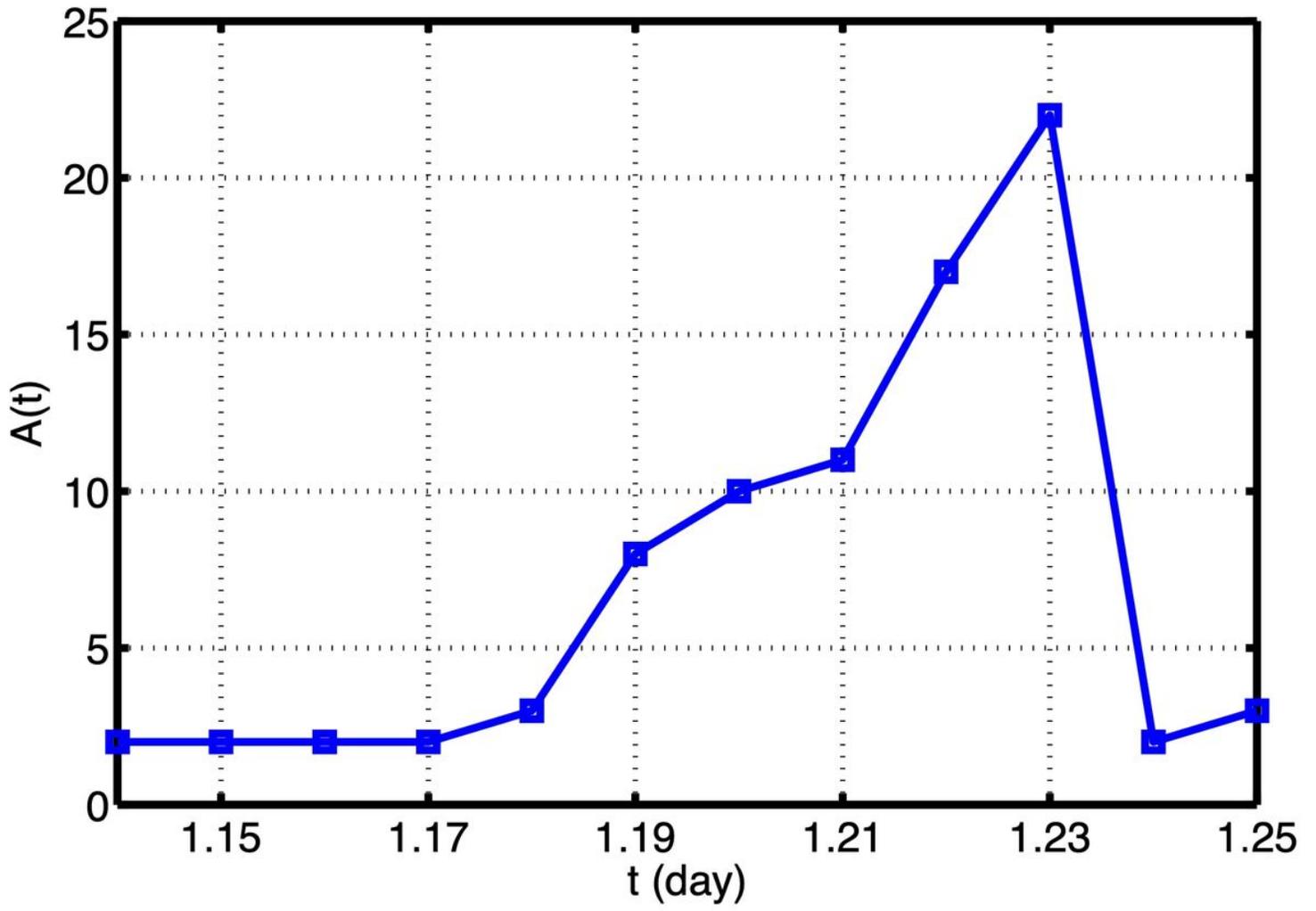


Figure 2

$A(t)$ with time.

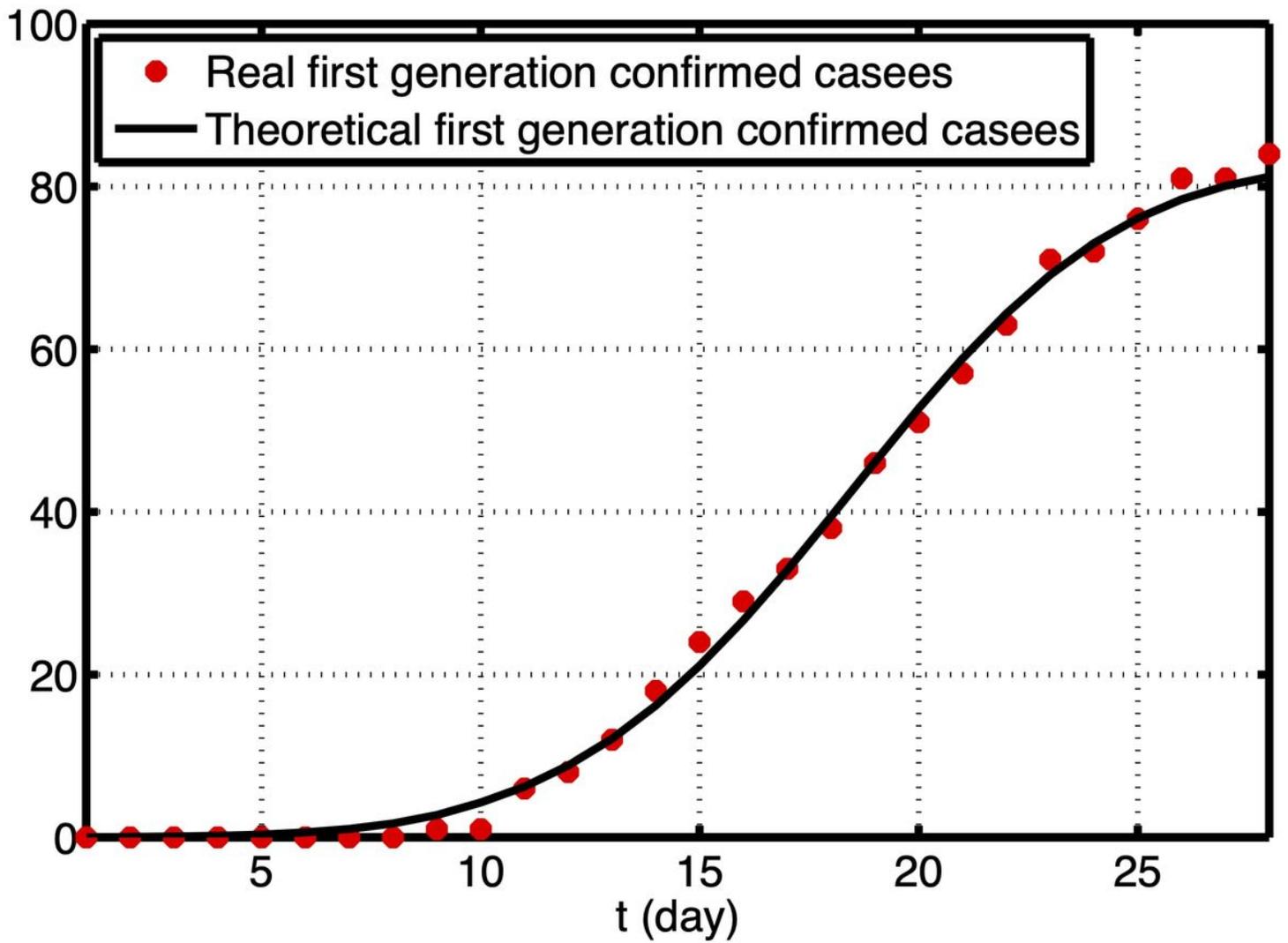


Figure 3

The fitting result of the first-generation cumulative confirmed cases. Where the red dots are actual reported cases from January 14 to February 10. Black solid line is the theoretical cases of $Q_1(t)$.

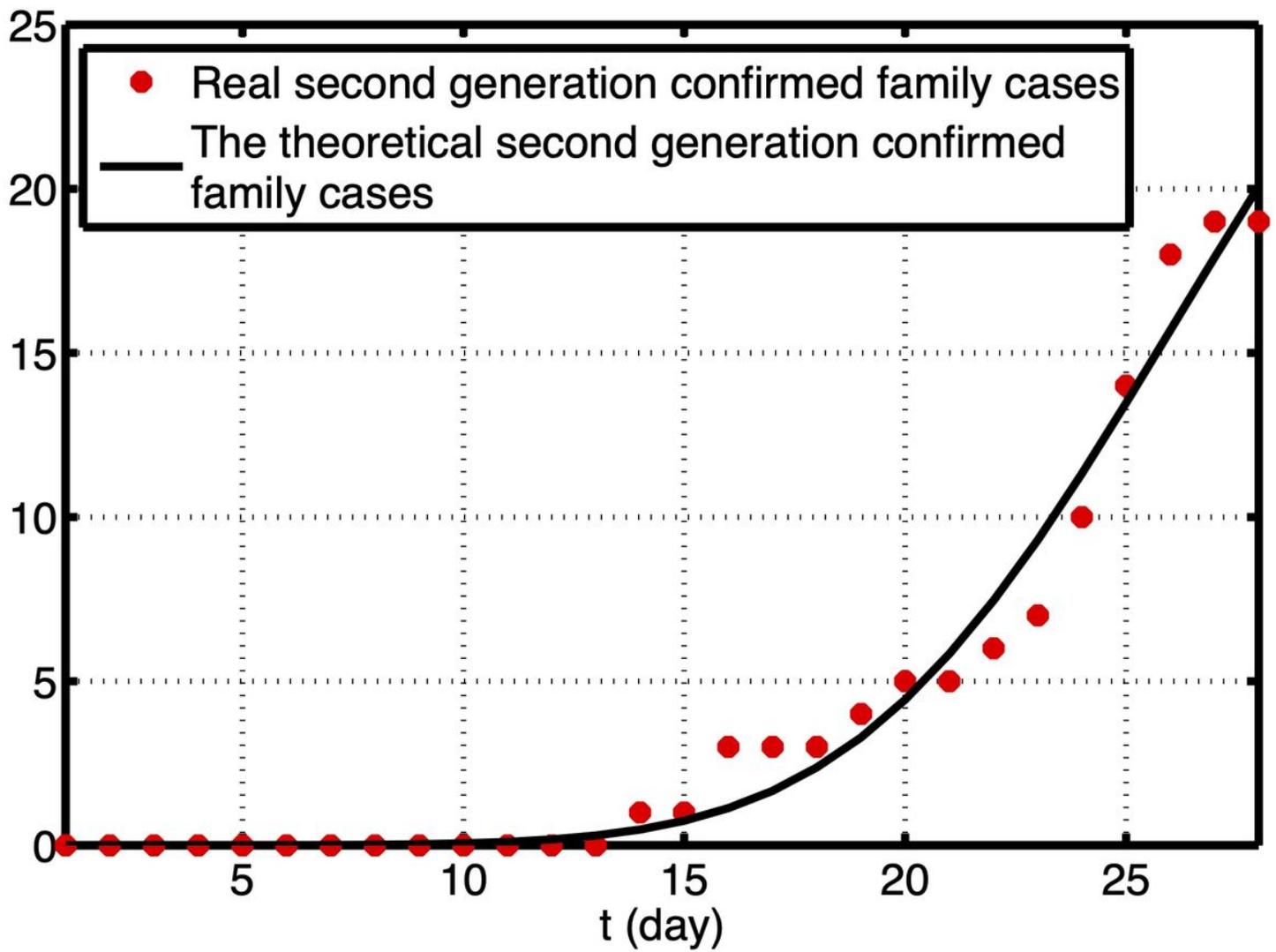


Figure 4

The fitting result of the second-generation cumulative confirmed family cases. Where the red dots are actual reported cases from January 14 to February 10. Black solid line is the theoretical cases of $\int_0^t \lambda C_1 I_1(\tau) F_1(t-\tau) d\tau$.

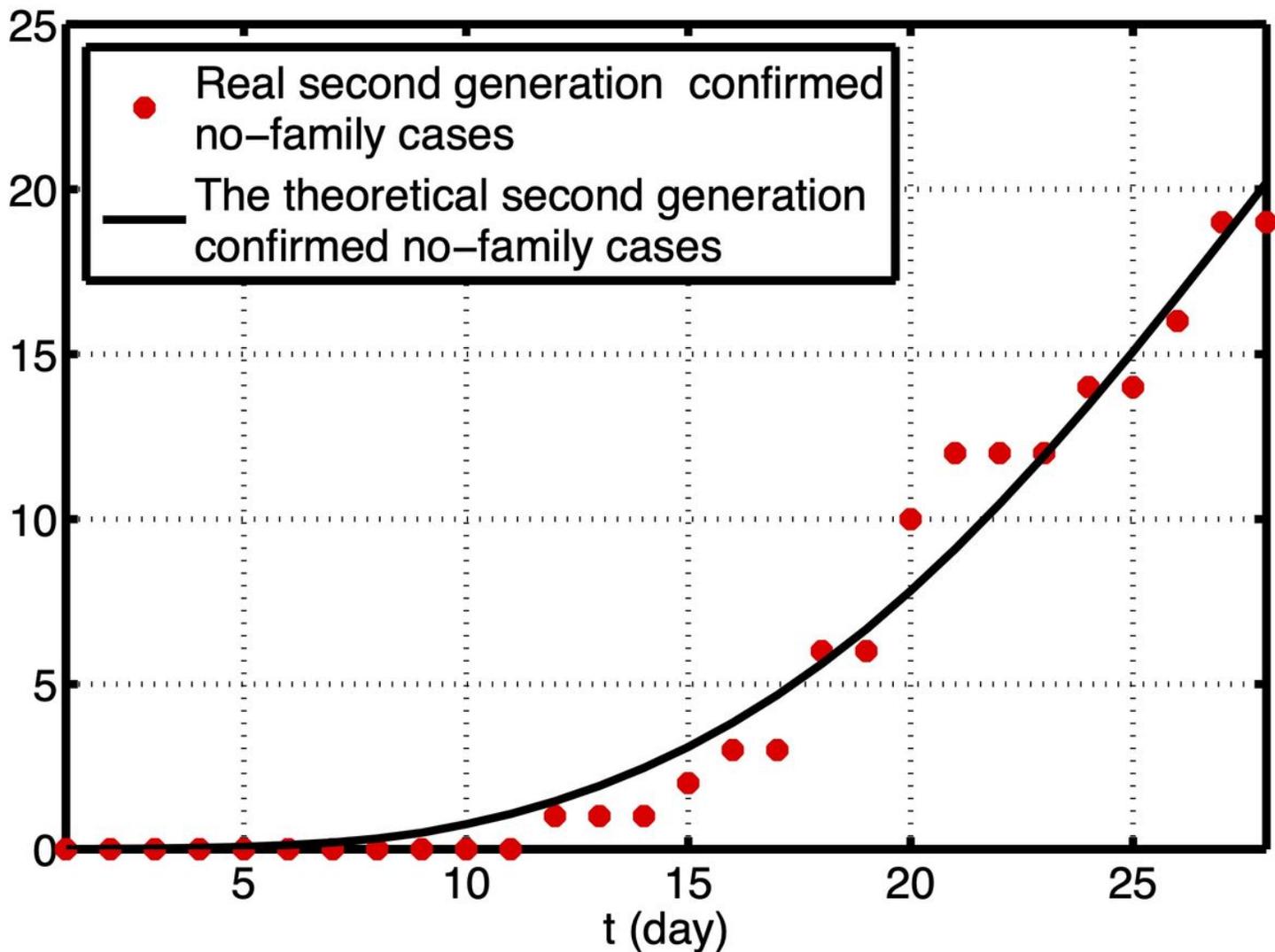


Figure 5

The fitting result of the second-generation cumulative confirmed non-family cases. Where the red dots are actual reported cases from January 14 to February 10. Black solid line is the theoretical cases of $\int_0^t \lambda C_2 I_1(\tau) F_2(t-\tau) d\tau$.

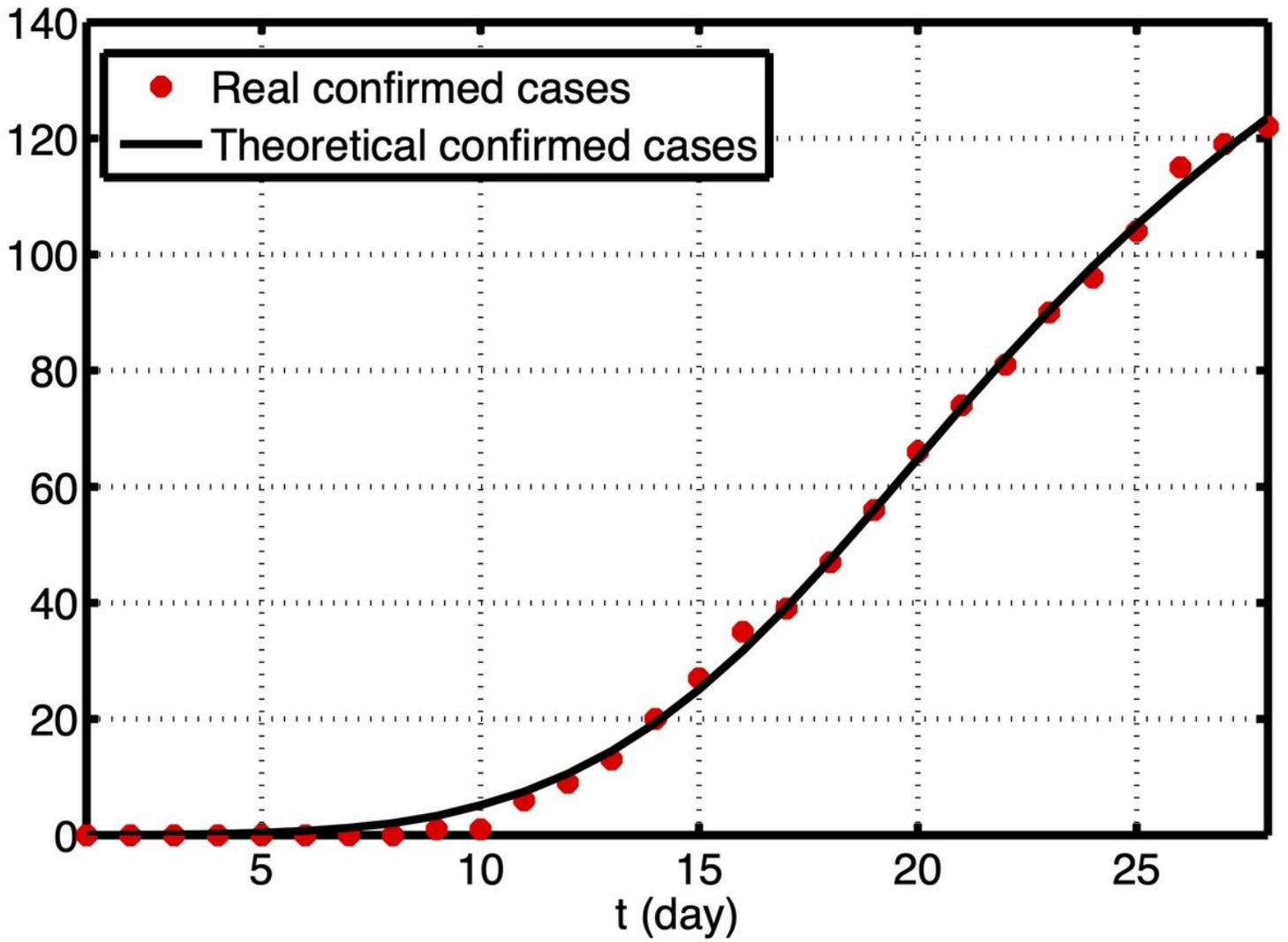


Figure 6

The fitting result of the total cumulative confirmed cases. Where the red dots are actual reported cases from January 14 to February 10. Black solid line is the theoretical cases of $Q(t)$.

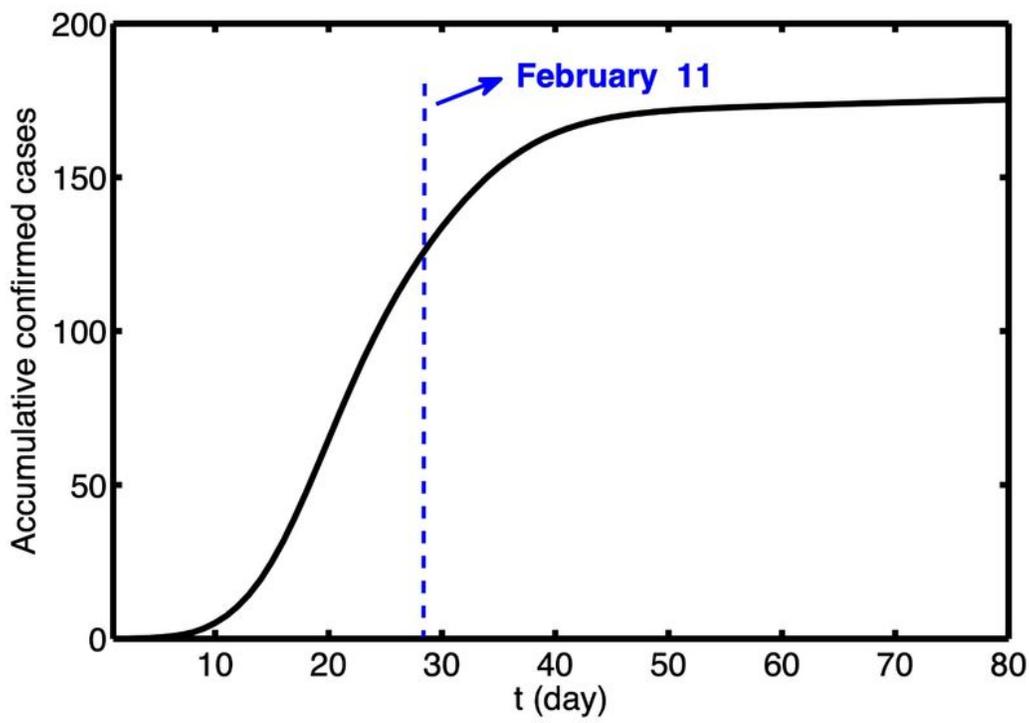
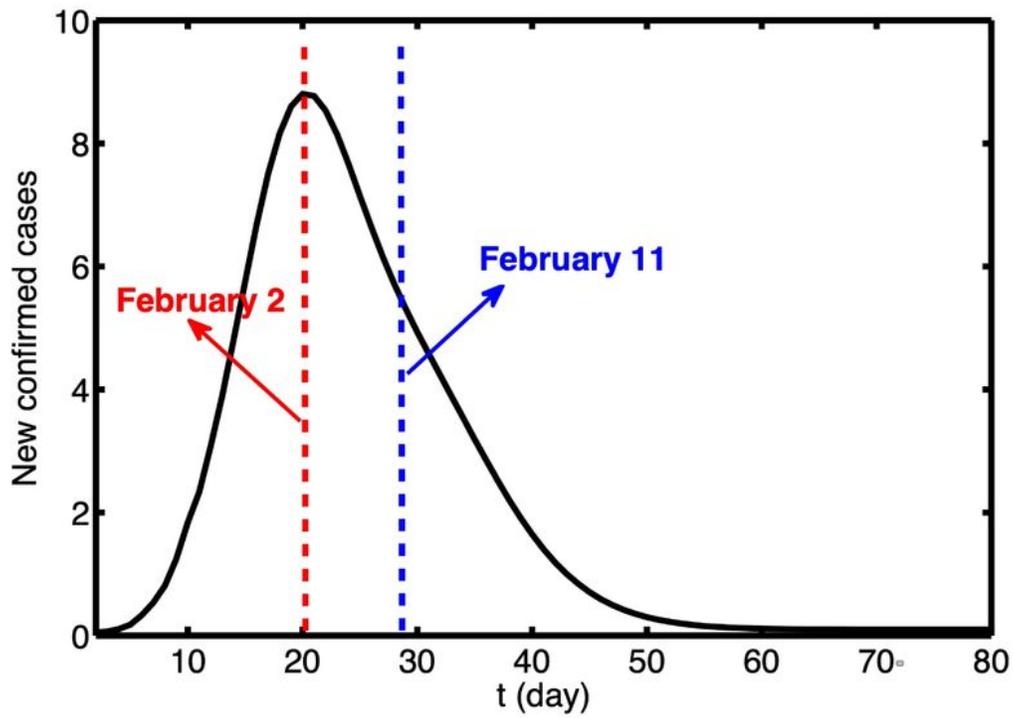


Figure 7

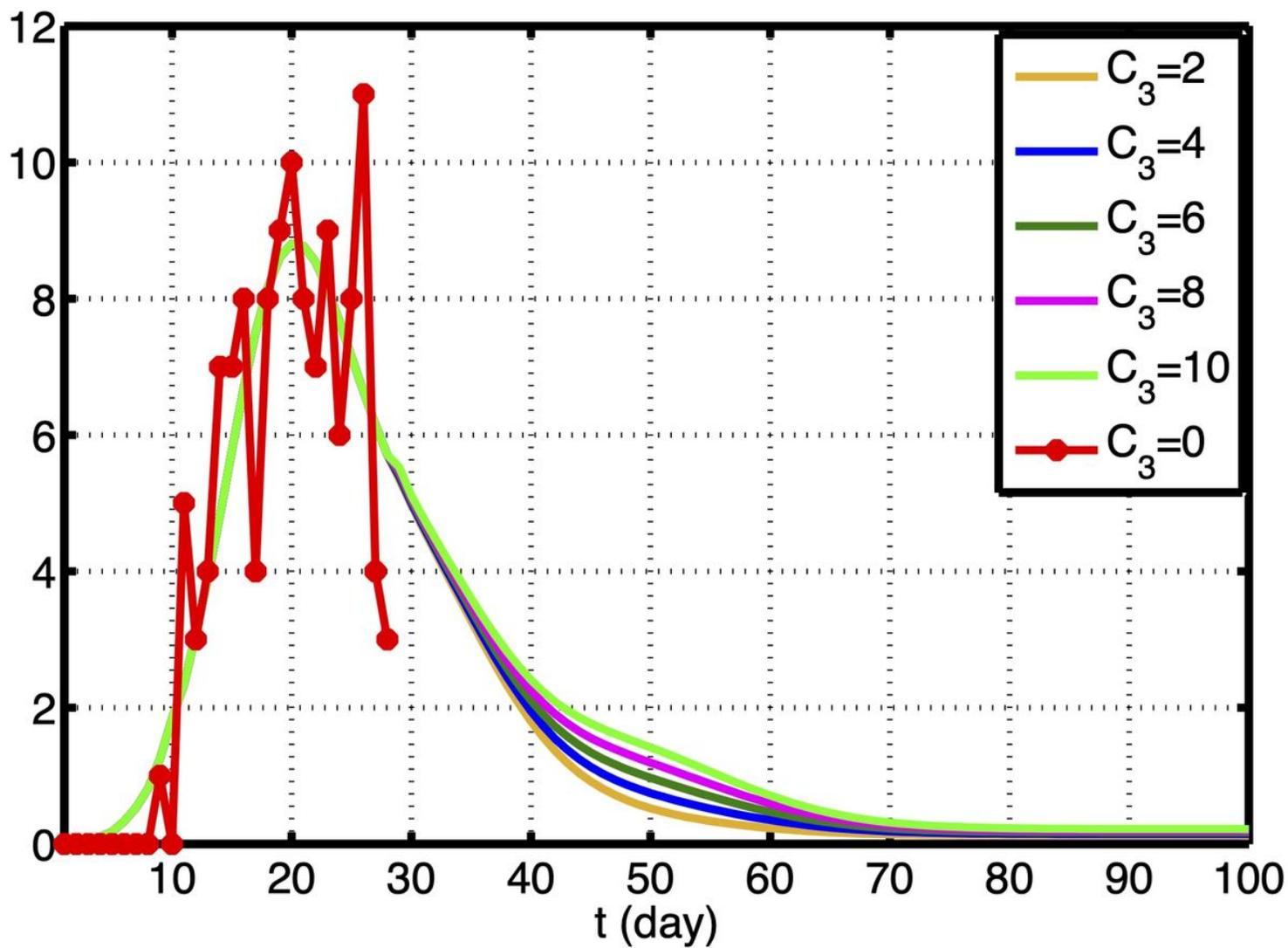


Figure 8

The prediction result of the total cumulative confirmed cases with the end of holiday. [0.9]

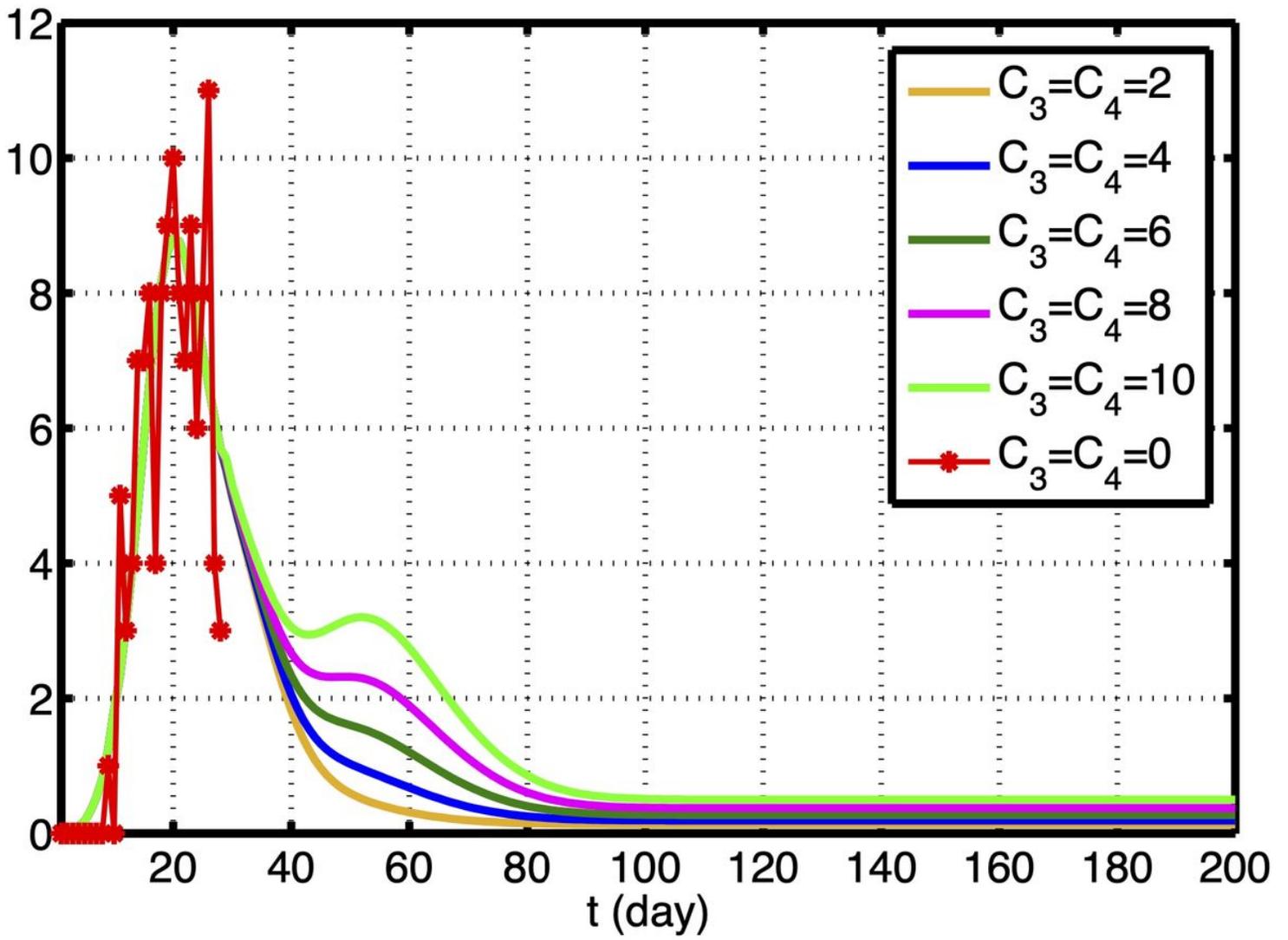


Figure 9

The theoretical result of the total cumulative confirmed cases with having the third-generation and fourth-generation infection.

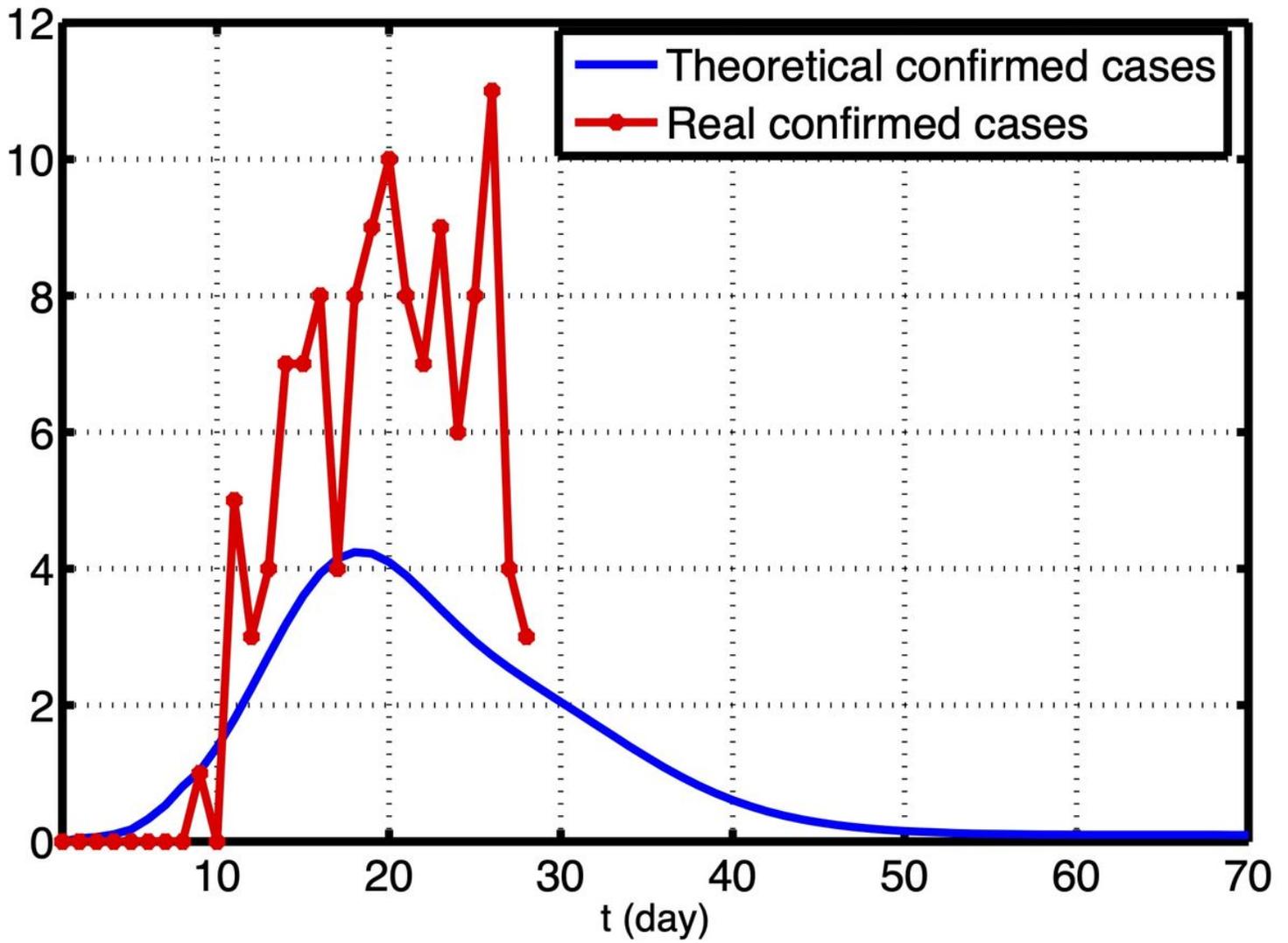


Figure 10

The theoretical result of the total cumulative confirmed cases with Case I.

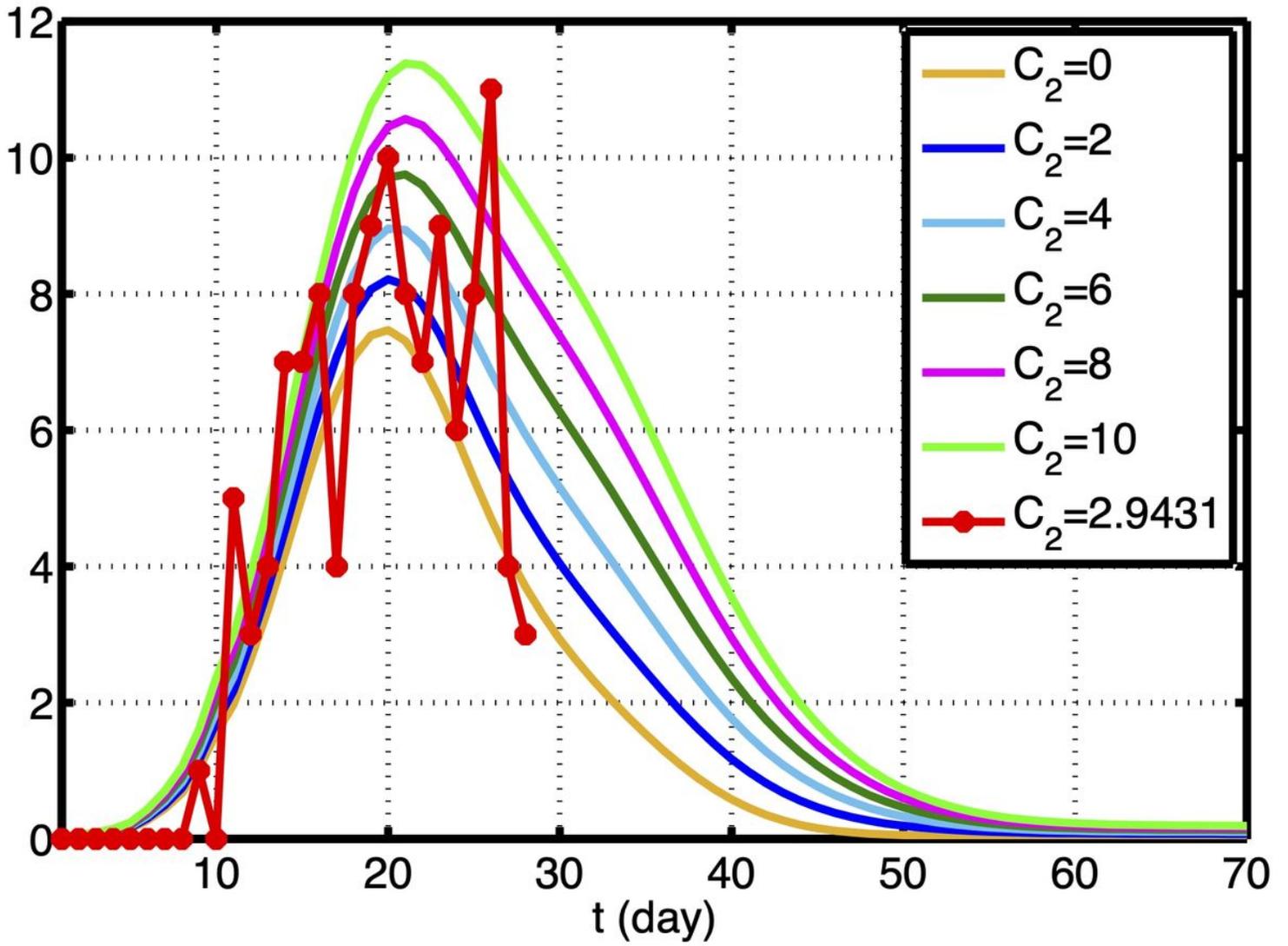


Figure 11

The theoretical result of the total cumulative confirmed cases with Case II.