Estimation of COVID-19 outbreak size in Harbin, China

Haitao Song (htsong@sxu.edu.cn)
Shanxi University https://orcid.org/0000-0003-1977-4308
Zhongwei Jia
Peking University
Zhen Jin
Shanxi University

Research Article

Keywords: COVID-19 outbreak, Mathematical model, Effective reproduction number, Harbin

Posted Date: June 1st, 2020

DOI: https://doi.org/10.21203/rs.3.rs-30516/v1

Version of Record: A version of this preprint was published at Nonlinear Dynamics on April 10th, 2021. See the published version at https://doi.org/10.1007/s11071-021-06406-2.

Estimation of COVID-19 outbreak size in Harbin, China

Haitao Song[1,2,†], Zhongwei Jia[3], Zhen Jin[1,2]

[1] Complex Systems Research Center, Shanxi University, Taiyuan 030006, China

[2] Shanxi Key Laboratory of Mathematical Techniques and Big Data Analysis on Disease Control and Prevention, Shanxi University, Taiyuan 030006, China

[3] National Institute of Drug Dependence, Peking University, Beijing 100191, China

[†] Correspondence to: Haitao Song (htsong@sxu.edu.cn).

Abstract

Background: Since the first level response to public health emergencies (FLRPHE) was launched on January 25, 2020 in Heilongjiang province, China, the outbreak of COVID-19 seems to be under control. However, an outbreak of COVID-19 caused by imported case developed in Harbin during April 2020. Here, we provide an estimate of the COVID-19 outbreak size in Harbin in April based on the number of found infected cases, assess the proportion of found and unfound infected cases in infected cases and evaluate the effective reproduction number which shows the transmission risk of COVID-19.

Methods: We used data from April 9 to April 30, 2020, on the number of found infected cases from the outbreak in Harbin to infer the number of infections in Harbin in the outbreak of COVID-19 and give the proportion of found and unfound infected cases in infected cases. Data on found infected cases were obtained from the the reports of Health Commission of Heilongjiang Province. A Susceptible-Unfound infected-Found infected-Removed model was used to fit the data on found infected cases of COVID-19 in Harbin using the least square method and simulate the transmission of COVID-19. The effective reproduction number was estimated.

Results: The COVID-19 outbreak size estimated in Harbin in April reaches 174, where 54% of infected cases were found and 46% of infected cases were not found out. Our findings suggest that the effective reproduction number decreased drastically in contrast with the value of 3.6 on April 9 after that the effective interventions were implemented by Heilongjiang province government. Finally, the effective reproduction number arrived the value of 0.04 which is immensely below the threshold value 1, which means that Heilongjiang province government got the outbreak of COVID-19 in Harbin under control.

Conclusions: The COVID-19 outbreak size in Harbin based on the assumptions that infected people with COVID-19 in the incubation period have the same infectivity with infectious people with COVID-19 could have been overestimated. As an increasing number of imported infected cases got into China and a growing number of asymptomatic infected people were

found, our study provides evidence that unfound infected cases would increase the risk of local outbreak of COVID-19 in China.

Keywords: COVID-19 outbreak, Mathematical model, Effective reproduction number, Harbin

Background

Since December 2019, the novel coronavirus (COVID-19) spread all provinces in China [1]. On March 2020, China government got the epidemic of COVID-19 under control by implementing the first level response to public health emergencies (FLRPHE) and strict control measures [2]. However, the COVID-19 spread rapidly all over the world, and 3,224,079 confirmed cases and 61,187 deaths were reported as of April 30, 2020.

The outbreak of COVID-19 in Harbin caused by an imported case and the first related case was diagnosed and two asymptomatic cases were found on April 9 [3–5]. As of April 30, 68 confirmed cases and 23 asymptomatic cases were reported from the COVID-19 outbreak in Harbin. On March 19, an American student (Han) carrying COVID-19 came back to Harbin from New York, and then infected her neighbor (Cao) during confinement period. Due to the infection within the family, Guo was infected with COVID-19, and then infected Chen by the dine together. Subsequently, Chen with cerebral apoplexy sought medical advice in the second hospital in Harbin and first affiliated hospital of Harbin medical university, which caused that many patients, health care providers, doctors and nurses were infected with COVID-19, and then led to the local outbreak of COVID-19 in Harbin. The important events related to the outbreak of COVID-19 in Harbin are shown in Table 1.

Table 1. The important events related to the outbreak of COVID-19 in Harbin.

Date	Event
Mar/19/2020	The American student (Han) arrived at Harbin from New York
Mar/27/2020	The dine together including Chen, Guo and Wang (Cao's mother) was held
Apr/2/2020	Chen arrived the second hospital in Harbin and was in hospital
Apr/6/2020	Chen left the second hospital in Harbin, and arrived at the first affiliated hospital of Harbin medical university and was in hospital
Apr/9/2020	One confirmed case and three asymptomatic cases were reported
Apr/10/2020	Chen was diagnosed with COVID-19
Apr/11/2020	Han was diagnosed with COVID-19
Apr/16/2020	Steering group for prevention and control of COVID-19 in Harbin was set up by Heilongjiang Provincial Party Committee
Apr/28/2020	Han was defined as the infectious source of the outbreak of COVID-19 in Harbin
Apr/28/2020	No new infected case was reported in Harbin

To estimate the transmission potential of COVID-19 in Harbin, the basic reproduction number and the effective reproduction number were considered. The basic reproduction number is defined as the expected number of secondary cases produced by a single infection in a completely susceptible population [6,7]. While the effective reproduction number is the mean number of secondary cases an infected person can cause in a population where there is some immunity or some intervention measures in place [8]. It is useful to compute the basic reproduction number to estimated the transmission capacity of COVID-19 when the outbreak occurs. However, as a result of interventions such as FLRPHE and isolation measures, the effective reproduction number changes in time, and it is necessary to investigate the effective reproduction number to combat COVID-19 [9].

To estimate the outbreak size of COVID-19 in Harbin, we established a mathematical model explaining the transmission dynamics of COVID-19. Using the SIFR (Susceptible-Unfound infected-Found infected-Removed) model, we estimated the effective reproduction number and the epidemic size of COVID-19 in Harbin, and investigated the proportion of found and unfound infected people in infected people.

Methods

Data source

In this study, parameters estimation were carried out using found infected cases from April 9 to April 30, 2020 in Harbin and the least square method by Matlab R2014a, and then the effective reproduction number and outbreak size of COVID-19 in Harbin were estimated.

Data on found infected cases of COVID-19 from April 9 to April 30, 2020 in Harbin were obtained from Health Commission of Heilongjiang Province [3]. The data set includes the cumulative number of found infected cases, new found infected cases and cured cases. These data used are from publicly available data sources.

Mathematical modelling

We estimated the effective reproduction number and outbreak size of COVID-19 based on the found infected cases in Harbin in April 2020.

We established the SIFR (Susceptible-Unfound infected-Found infected-Removed) model to describe the transmission dynamics of COVID-19 in Harbin. The report of the WHO-China Joint Mission on COVID-19 [10] shows that the infected people in the incubation period has infectivity. Susceptible people become infected people by contacting with unfound infected people. Some infected people are found and treated in hospital, and unfound infected patients with COVID-19 get the self-healing. Where unfound infected people includes infected people in the incubation period, unfound asymptomatic and symptomatic infected people. We denoted susceptible people by S, unfound infected people by I, found infected people by F, removed people by R, respectively. Here we assumed that infected people with COVID-19 in the incubation period have the same infectivity with infectious people with COVID-19, and found infected people are quarantined and could not infect healthy people. Therefore, we established the transmission dynamics of COVID-19 in Harbin which is given by

$$
\begin{cases}
\dfrac{\mathrm{d}S(t)}{\mathrm{d}t} = -\dfrac{\beta S(t)I(t)}{S(t)+I(t)+D(t)+R(t)}, \\
\dfrac{\mathrm{d}I(t)}{\mathrm{d}t} = \dfrac{\beta S(t)I(t)}{S(t)+I(t)+D(t)+R(t)} - \gamma_1 I(t) - pI(t), \\
\dfrac{\mathrm{d}F(t)}{\mathrm{d}t} = pI(t) - \gamma_2 F(t), \\
\dfrac{\mathrm{d}R(t)}{\mathrm{d}t} = \gamma_1 I(t) + \gamma_2 F(t),
\end{cases}
$$

where β is the transmission rate from susceptible to infected people, γ_1 is the recovery rate of unfound infected people, γ_2 denotes the removed rate of infected people and p is the found rate of infected people.

Using the next generation matrix theory [6, 7], the effective reproduction number was computed as

$$
R_e(t) = \frac{\beta S(t)}{(\gamma_1 + p)(S(t)+I(t)+D(t)+R(t))}.
$$

Final size of the COVID-19 outbreak in Harbin was computed as

$$
\int_0^t I(t)dt = \frac{1}{\gamma_1 + p}(S(0)+I(0)-S(t)-I(t)).
$$

Using the least square method, we fitted our model with the found infected cases from April 9 to April 30, 2020 in Harbin and obtained the related initial values and parameter values in Harbin which were shown in Table 2. Then we estimated the outbreak size of COVID-19 in Harbin in April 2020 using our model. Finally, we estimated the effective reproduction number $R_e(t)$ to understand and control the outbreak of COVID-19 in Harbin in April 2020.

Results

Firstly, we fitted our SIFR model with found infected cases using the least square method, and the Mean Absolute Error (MAE) is 0.0562 (Fig. 1). The related initial values and parameter values were shown in Table 2. Then we computed the basic reproduction number of 3.6 on April 9, 2020. Fig. 1 shows that the cumulative number of infected people reached 174, the cumulative number of found infected people was 94 and the cumulative number of unfound infected people was 80.

The effective reproduction number $R_e(t)$ for our model was shown in Fig. 2. On April 9, the effective reproduction number was 3.6 which was the maximum of $R_e(t)$. As time went on, $R_e(t)$ decreased quickly and was less than threshold value 1 after April 15. Subsequently, $R_e(t)$ arrived the minimum of 0.04 on April 30 which immensely below the threshold value 1. Where the outbreak of COVID-19 is under control when the effective reproduction number $R_e(t)$ is less than 1. This implies that the outbreak of COVID-19 in Harbin in April 2020 was under control when effective interventions were implemented.

Table 2. Related parameters and initial values in Harbin.

Parameter Values	Descriptions	Mean value	Source
β	The transmission rate of COVID-19 per day	0.8403	Estimated
p	The found probability of infected people per day	0.1128	Estimated
γ_1	The recovery rate of unfound infected people	0.1	[27]
γ_2	The removed rate of found infected people	0.1	[27]
Initial Values	**Descriptions**	**Mean value**	**Source**
$S(0)$	Initial number of susceptible people	163	Estimated
$I(0)$	Initial number of unfound infected people	11	Estimated
$F(0)$	Initial number of found infected people	4	Data
$R(0)$	Initial number of removed people	0	Data

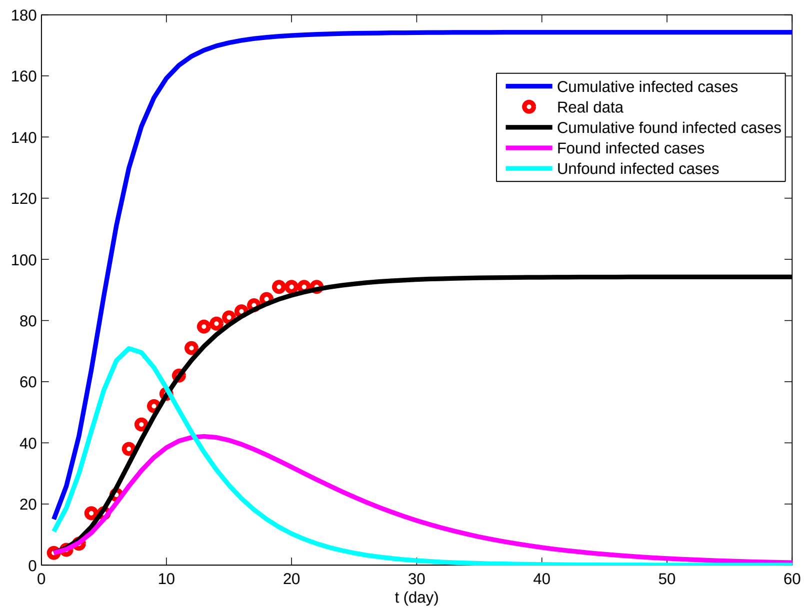

Figure 1: The simulated cumulative infected cases, simulated cumulative found infected cases, simulated found infected cases and unfound infected cases in Harbin are shown using our model. Parameter values and initial values are defined in Table 2.

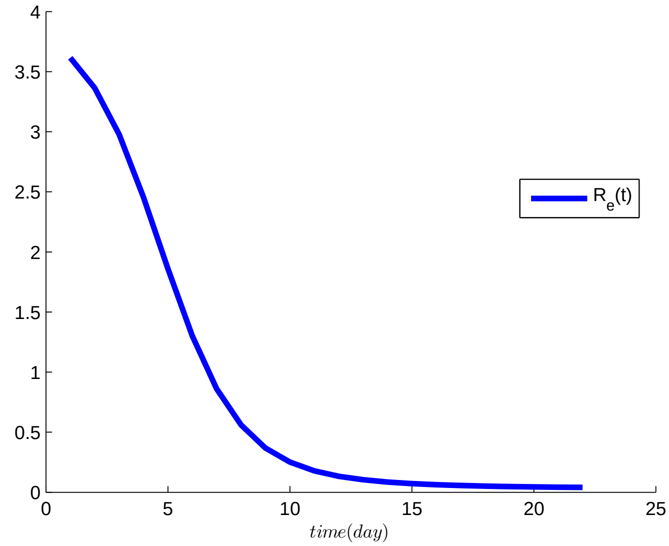

Figure 2: The estimated effective reproduction number of COVID-19 in Harbin in April 2020. Parameter values and initial values are defined in Table 2.

Discussion

This is the first study to estimate the transmission potential of COVID-19 outbreak of in Harbin in April 2020. The cumulative number of infected people finally reached 174, where 54% of infected people were found and 46% of infected people were not found out. Although all close contacts tracked were detected, some infected people were not found out. Indeed, the unfound infected people might be the infected people in the incubation period, unfound asymptomatic and symptomatic infected people. It is dangerous for public health to ignore unfound infected people. We must maintain vigilance against unfound infected people.

Our findings indicate that the effective reproduction number on April 9 got to 3.6 which is consistent with the estimated value in China [11–14]. However, when effective interventions were implemented by Heilongjiang province government, the effective reproduction number $R_e(t)$ drastically dropped and finally reached 0.04 which is greatly below the threshold value 1, which suggest that the outbreak of COVID-19 in Harbin in April 2020 was under control and no subsequent outbreak in Harbin.

The mathematical modelling used in the study is analogous to the transmission dynamics model of COVID-19 in [15–24]. The SIFR model help us to estimate the cumulative number of infected cases of COVID-19 in Harbin in April and the effective reproduction number using found infected cases. Nonetheless, there are several limitations. First, we assumed that infected people with COVID-19 in the incubation period have the same infectivity with infectious people with COVID-19, which caused that the outbreak size of COVID-19 in Harbin was overestimated. Second, the detail interventions were not incorporated into our model, which might lead to overestimate the outbreak size of COVID-19 in Harbin in April 2020. Third, very little is known about the effect of temperature and precipitation on the transmission of COVID-19.

Our estimation could be untrustworthy if temperature and precipitation have strong impact on the transmission of COVID-19.

Conclusions

Using the published data and our mathematical model, we estimated the COVID-19 outbreak size in Harbin, and provided the proportion of found and unfound infected cases in infected cases. The effective reproduction number was estimated based on the published data and our mathematical model. Although the transmission of COVID-19 has been under control in China, the epidemic situation of COVID-19 all over the world is serious. Now, an increasing number of imported infected cases got into China and a growing number of asymptomatic infected people were found, which might increase the risk of local outbreak of COVID-19 in China. Therefore, we should stay alert in case that unfound infected people might cause local outbreaks of COVID-19 in China such as the outbreak of COVID-19 in Harbin in April 2020 [25, 26].

Abbreviations

COVID-19: 2019 novel coronavirus; FLRPHE: the first level response to public health emergencies; SIFR: Susceptible-Unfound infected-Found infected-Removed; MAE: Mean Absolute Error; $R_e(t)$: Effective reproduction number; WHO: World Health Organization.

Authors' contributions

Haitao Song and Zhongwei Jia contributed to the study conception and design. Material preparation, data collection and analysis were performed by Haitao Song. The first draft of the manuscript was written by Haitao Song. Haitao Song, Zhongwei Jia and Zhen Jin commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding

This research was funded by the National Natural Science Foundation of China (11601291), and Program for the Outstanding Innovative Teams (OIT) of Higher Learning Institutions of Shanxi, and Shanxi Scholarship Council of China and Scientific and Technological Innovation Programs (STIP) of Higher Education Institutions in Shanxi.

Availability of data and materials

Not applicable.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

Not applicable.

Author details

[1] Complex Systems Research Center, Shanxi University, Taiyuan, Shanxi Province, Peoples Republic of China. [2] Shanxi Key Laboratory of Mathematical Techniques and Big Data Analysis on Disease Control and Prevention, Shanxi University, Taiyuan, Shanxi Province, Peoples Republic of China. [3] National Institute of Drug Dependence, Peking University, Beijing, Peoples Republic of China.

References

[1] Zhu N, Zhang D, Wang W, et al. A novel coronavirus from patients with pneumonia in China, 2019. New England Journal of Medicine, 2020.

[2] Chinadaily, "Tibet activates highest-level public health alert (January 30, 2020)"; https://www.chinadaily.com.cn/a/202001/29/WS5e318a36a3101282172739c1.html.

[3] Health Commission of Heilongjiang Province. http://wsjkw.hlj.gov.cn/index.php/Home/Zwgk/all/typeid/42 [2020-04-30].

[4] Chinese Center for Disease Control and Prevention. http://2019ncov.chinacdc.cn/2019-nCoV/.

[5] National Health Commission of the People's Republic of China. http://www.nhc.gov.cn/xcs/xxgzbd/gzbd_index.shtml [2020-04-30].

[6] Diekmann, O., Heesterbeek, J. A. P. & Metz, J. A. J. On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations, J. Math. Biol. 28, 365382 (1990).

[7] Van den Driessche, P. & Watmough, J. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Mathematical biosciences 180, 29-48 (2002).

[8] De Serres G, Gay N J, Farrington C P. Epidemiology of transmissible diseases after elimination. American Journal of Epidemiology, 2000, 151(11): 1039-1048.

[9] Cao Z, Zhang Q, Lu X, et al. Estimating the effective reproduction number of the 2019-nCoV in China. medRxiv, 2020.

[10] World Health Organization (WHO). Coronavirus disease (COVID-2019) outbreak, report of the WHO-China Joint Mission on COVID-19. https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf.

[11] Wu J T, Leung K, Leung G M. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. The Lancet, 2020, 395(10225): 689-697.

[12] Niehus R, De Salazar P M, Taylor A R, et al. Using observational data to quantify bias of traveller-derived COVID-19 prevalence estimates in Wuhan, China. The Lancet Infectious Diseases, 2020.

[13] Sen-zhong H, Zhihang P, Zhen J. Studies of the strategies for controlling the COVID-19 epidemic in China: Estimation of control efficacy and suggestions for policy makers. SCIENTIA SINICA Mathematica, 2020.

[14] Tang B, Wang X, Li Q, 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] Li R, Pei S, Chen B, et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). Science, 2020.

[16] Tian H, Liu Y, Li Y, et al. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. Science, 2020.

[17] Imai N, Cori A, Dorigatti I, et al. MRC Centre for Global Infectious Disease Analysis: Wuhan coronavirus reports 1C3. January, 2020. https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/news–wuhan-coronavirus (accessed Jan 25, 2020).

[18] Read J M, Bridgen J R E, Cummings D A T, et al. Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic predictions. MedRxiv, 2020.

[19] Zhao S, Lin Q, Ran J, 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. International journal of infectious diseases, 2020, 92: 214-217.

[20] Wu J T, Leung K, Bushman M, et al. Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. Nature Medicine, 2020: 1-5.

[21] Prem K, Liu Y, Russell T W, et al. The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. The Lancet Public Health, 2020.

[22] Tian J, Wu J, Bao Y, et al. Modeling analysis of COVID-19 based on morbidity data in Anhui, China. 2020.

[23] Wang K, Lu Z, Wang X, et al. Current trends and future prediction of novel coronavirus disease (COVID-19) epidemic in China: a dynamical modeling analysis. Mathematical Biosciences and Engineering, 2020, 17(4): 3052.

[24] Chen T M, Rui J, Wang Q P, et al. A mathematical model for simulating the phase-based transmissibility of a novel coronavirus. Infectious diseases of poverty, 2020, 9(1): 1-8.

[25] Song H, Liu F, Li F, et al. The Impact of Isolation on the Transmission of COVID-19 and Estimation of Potential Second Epidemic in China. Preprints 2020, 2020030456 (doi: 10.20944/preprints202003.0456.v1).

[26] Jia Z, Lu Z. Modelling COVID-19 transmission: from data to intervention. The Lancet Infectious Diseases, 2020.

[27] You C, Deng Y, Hu W, et al. Estimation of the time-varying reproduction number of COVID-19 outbreak in China. Available at SSRN 3539694, 2020.

Figures

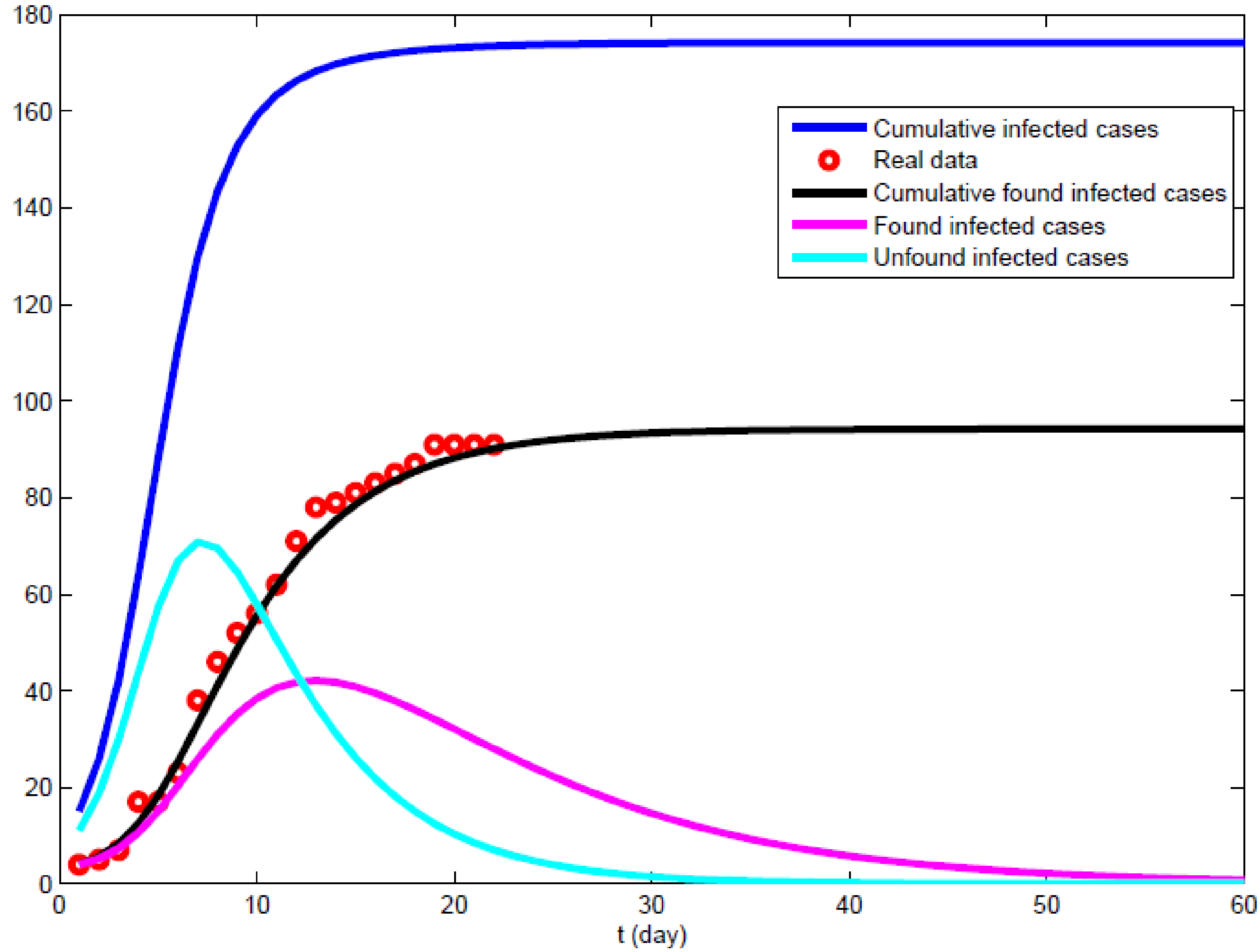

Figure 1

The simulated cumulative infected cases, simulated cumulative found infected cases, simulated found infected cases and unfound infected cases in Harbin are shown using our model. Parameter values and initial values are defined in Table 2.

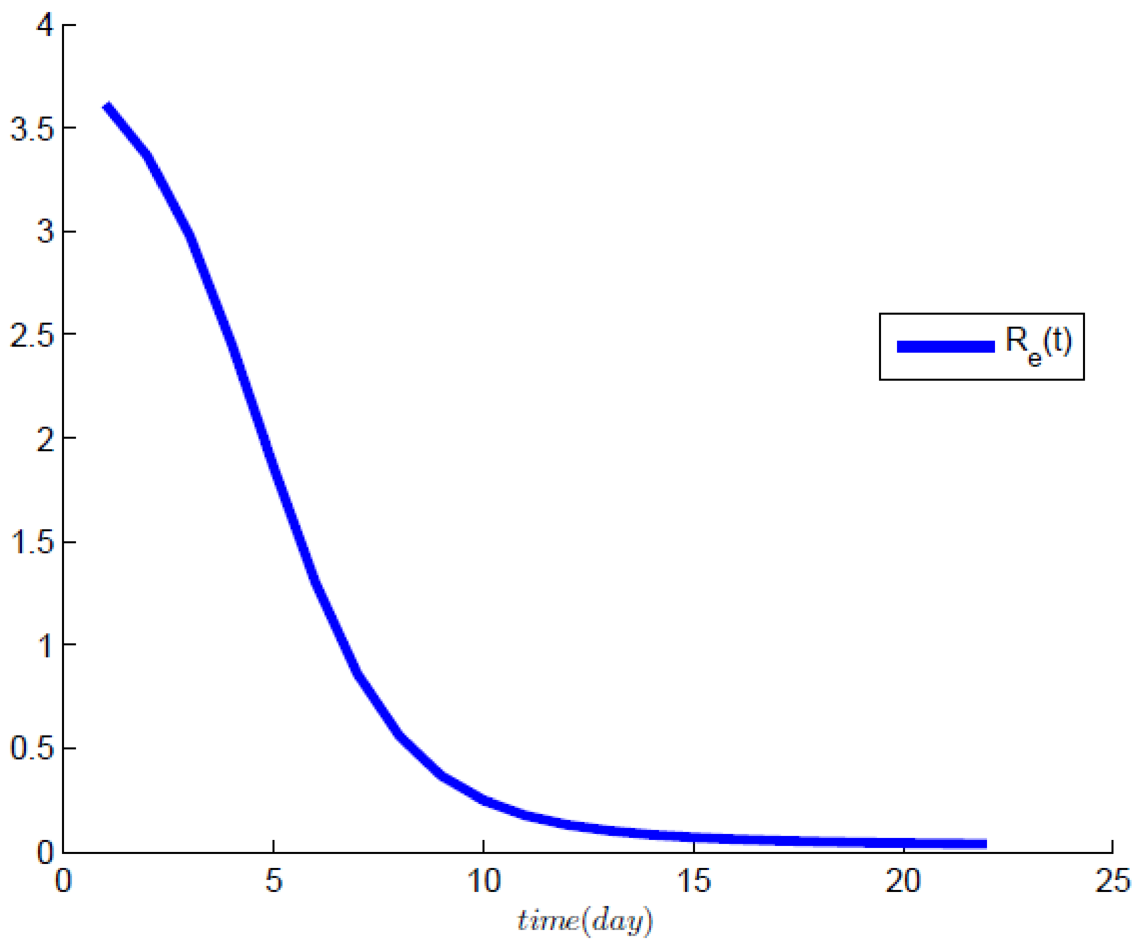

Figure 2

The estimated effective reproduction number of COVID-19 in Harbin in April 2020. Parameter values and initial values are defined in Table 2.