Mathematical Biosciences and Engineering

http://www.aimspress.com/journal/MBE

MBE, 17(4): 2842–2852.
DOI: 10.3934/mbe.2020158
Received: 22 February 2020
Accepted: 20 March 2020
Published: 25 March 2020

Research article

Modeling analysis of COVID-19 based on morbidity data in Anhui, China

Jingjing Tian[1,†], Jiabing Wu[2,†], Yunting Bao[1], Xiaoyu Weng[1], Lei Shi[1], Binbin Liu[1], Xinya Yu[1], Longxing Qi[1,*] and Zhirong Liu[2,*]

[1] School of Mathematical Sciences, Anhui University, Hefei 230601, China

[2] Anhui Provincial Center for Disease Control and Prevention, Hefei 230601, China

† These authors contributed equally to this work.

* **Correspondence:** Email: qilx@ahu.edu.cn (L. Qi); Tel: +0551-63861952;
lzr@ahcdc.com.cn (Z. Liu); Tel: +0551-63674900.

Abstract: Since the first case of coronavirus disease (COVID-19) in Wuhan Hubei, China, was reported in December 2019, COVID-19 has spread rapidly across the country and overseas. The first case in Anhui, a province of China, was reported on January 10, 2020. In the field of infectious diseases, modeling, evaluating and predicting the rate of disease transmission is very important for epidemic prevention and control. Different intervention measures have been implemented starting from different time nodes in the country and Anhui, the epidemic may be divided into three stages for January 10 to February 11, 2020, namely. We adopted interrupted time series method and develop an SEI/QR model to analyse the data. Our results displayed that the lockdown of Wuhan implemented on January 23, 2020 reduced the contact rate of epidemic transmission in Anhui province by 48.37%, and centralized quarantine management policy for close contacts in Anhui reduced the contact rate by an additional 36.97%. At the same time, the estimated basic reproduction number gradually decreased from the initial 2.9764 to 0.8667 and then to 0.5725. We conclude that the Wuhan lockdown and the centralized quarantine management policy in Anhui played a crucial role in the timely and effective mitigation of the epidemic in Anhui. One merit of this work is the adoption of morbidity data which may reflect the epidemic more accurately and promptly. Our estimated parameters are largely in line with the World Health Organization estimates and previous studies.

Keywords: COVID-19; morbidity; mitigation; basic reproduction number; mathematical model; interrupted time series analysis

1. Introduction

Coronavirus belongs to the *Coronaviridae family* in phylogeny. It was first identified as a pathogenic microorganism of human disease in the mid-1960s. The name comes from the fact that the ratchet protein on the surface of the virus envelope protrudes like a crown under an electron microscope. Before 2019, six known coronaviruses types have been discovered: α genus of 229E, β genus of OC43, HKU1, Middle East respiratory syndrome coronavirus (MERS-COV) and severe acute respiratory syndrome (SARS-COV) [1–3]. In December 2019, an unidentified pneumonia was found in Wuhan, Hubei province, China. The responsible virus was later confirmed as a new coronavirus [4]. The National Board of Health name the disease as a novel coronavirus pneumonia (NCP). The World Health Organization (WHO) temporarily named the virus as the novel coronavirus 2019 (2019-nCov), and the disease as coronavirus disease 2019 (COVID-19) on 11 February. At the same time, the International Committee for the classification of viruses named the virus as severe acute respiratory syndrome coronavirus-2 (SARS-COV2) [5]. The disease spread rapidly from Wuhan to all parts of the country and overseas. Till March 12, 2020 at 5:44 p.m, a total of 80,981 people had been confirmed infected in China, and the total death toll had reached 3173. New cases of COVID-19 have emerged in 109 countries overseas (including Japan, South Korea and so on), with 38,620 confirmed cases and 1465 deaths get rid of China from the WHO [6].

Anhui, a neighbouring province of Hubei, is vulnerable to the disease. Since the first confirmed case reported on January 10, 2020 in Anhui, a total of 990 cases have been confirmed and 6 deaths have been reported by March 12. Figure 1 shows the heat map of the total number of confirmed cases in China on February 6. The epidemic poses a challenge to the prevention and control of the epidemic in Anhui .

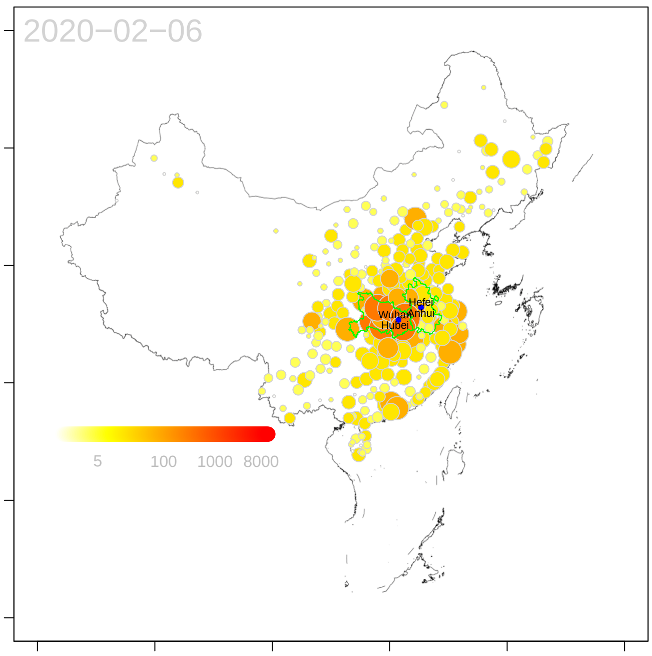

Figure 1. Heat map of the total number of confirmed cases in China on February 6.

In the face of the severe epidemic, the Chinese government promptly adopted a series of policies to mitigate the epidemic. The lockdown of the epicenter Wuhan has been implemented since January 23, 2020. On February 6, Anhui province started to adopt the centralized quarantine measures for close contacts of cases. These policies effectively prevented the spread of the disease in Anhui. This is the first work to use mathematical modeling techniques combined with observed data to evaluate the control effectiveness in Anhui. As the outbreak of COVID-19 now became a pandemic, the lesson learnt in Anhui, China is surely valuable for mitigation in other localities. Data-driven mathematical modelling plays an important role in epidemic mitigation, in preparedness for future epidemic and in the evaluation of control effectiveness. The basic reproduction numbers (number of secondary cases may be caused by a typical primary case in a wholly susceptible population, $\mathcal{R}_0$) is one of the key epidemiological parameters in quantifying the rate of disease transmission, the size of the outbreak and the needed control effort [7, 8]. Many data-driven modelling works in different regions have been done [9–11]. In this work, we focused on the epidemic in Anhui based on daily morbidity data which reflect the epidemic more accurately and promptly than laboratory confirmed data. Different from previous studies [8–15], we adopted relatively simple models with flexible transmission rates and model structures.

2. Data-driven modeling

2.1. Interrupted time series analysis

The interrupted time series analysis is a quasi-experimental research design. By collecting and analyzing relevant data of the outcome indicators measured at multiple time points before and after the implementation of an intervention, the change trend of the outcome indicators before and after the implementation of the intervention is compared to evaluate the impact of the intervention on the outcome indicators [16, 17]. We obtained the daily morbidity data from January 10 to February 11, 2020 from the Anhui Province Center of Disease Control and Prevention. During this period, two interventions occurred, Wuhan lockdown on January 23 (i.e., no imported cases from Wuhan to Anhui thereafter), and centralized quarantine management measures for close contacts implemented since February 6. We adopted an interrupted time series model to evaluate the effects of these two interventions.

We define the interrupted time series model as follows: $Y_t = \beta_0 + \beta_1 + time_t + \beta_2 * intervention1_t + \beta_3 * time\,after\,intervention1_t + \beta_4 + intervention2_t + \beta_5 * time\,after\,intervention2_t + e_t$, where Y_t is the daily morbidity index value, β_0 is the baseline value (also initial value at $t = 0$); β_1 is the baseline slope value (before any intervention); β_2 refers to the contribution of the first intervention on the baseline value; β_3 is the contribution of the first intervention on the baseline slope; That is, $\beta_1 + \beta_3$ is the trend after the implementation of the first intervention; similarly, $\beta_{4,5}$ are contribution of the second intervention on the baseline value and baseline slope; $\beta_{4,5}$ replaced $\beta_{1,2}$ after the implementation of the second intervention. Thus, $\beta_1 + \beta_5$ is the trend of daily morbidity after the second intervention.

As shown in Figure 2, the model prediction (the black curve) is divided into three sections, before January 23 (phase I), between January 23 and February 6 (phase II) and after February 6 (phase III). The slopes of in phases I-III are $\beta_1, \beta_1 + \beta_3, \beta_1 + \beta_5$, respectively. The instantaneous and lasting effects of the two interventions are clearly demonstrated.We fit the model with Eviews and find the r-squared $R^2 = 0.92$ and the Durbin-Watson statistic $DW = 2.32$, which indicates high fitting performance of the model. Parameter estimates are listed in the Table 1.

Table 1. Fitting parameter values of interrupted time series model.

Parameter	Estimate	p-value
β_0	7.076923	0.0778
β_1	2.923077	0.0000
β_2	38.29670	0.0000
β_3	-6.057143	0.0000
β_4	-4.742857	0.5012
β_5	-0.323077	0.8442

According to Table 1, the linear growth trend of the morbidity number during phase I was $\beta_1 = 2.9$ ($p-$value $< 1e-5$). The morbidity number had a sudden lift $\beta_2 = 38$ ($p-$value $< 1e-5$) on January 23, 2020. After that the slope of the morbidity number decreased by $\beta_3 = 6.057$ ($p-$value $< 1e-5$) . These results imply that right before the lockdown of Wuhan a large number of imported cases from Wuhan or other Hubei cities arrived into Anhui. The lockdown effect was substantial as the morbidity data showed an immediate down trend. However, the effect of the second intervention (after February 6) was noticeable but no significant. As shown in Figure 2, the daily number of cases continued to decrease after February 6. To sum up, it can be seen that the lockdown measures in Wuhan as well as the centralized quarantine management measures in Anhui had a great impact against the spread of COVID-19 in Anhui.

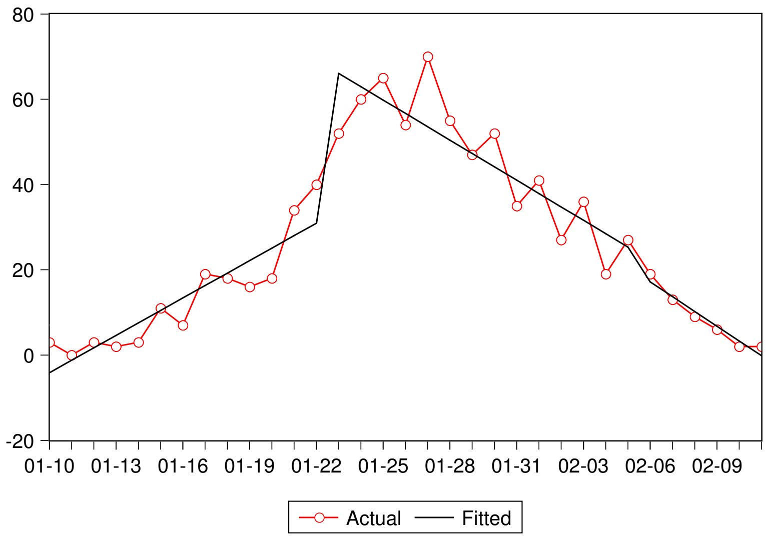

Figure 2. Trend of morbidity number of COVID-19 based on piecewise regression model.

2.2. Mechanistic modeling

According to the Anhui Provincial Center for Disease Control and Prevention, the onset of the first case in Anhui was on January 10, 2020 and the case was confirmed on January 22, 2020. There is a 12 days confirmation delay. According to the report of the National Health Commission, both pre-symptomatic patients and symptomatic patients may be contagious. Recent studies showed that the mean serial interval (time delay between the symptom onset of a primary case and a second case) is short, and could be shorter than the incubation period (time delay from infection to symptom onset). Thus pre-symptomatic transmission (one day or two prior symptom onset) is fully justified [9, 10].

Our interrupted time series analysis showed that both Wuhan lockdown and Anhui centralized quarantine management had an impact on the daily morbidity number in Anhui. During the initial phase of the COVID-19 outbreak, the rate of exposure could be high due to lack of awareness and protection. Subsequently, the relevant announcement and measures issued by the government attracted public's attention and enhanced level of self-protection, and strict quarantine measures were adopted. The rate of exposure gradually declined, and the form of epidemic transmission (main transmitter) may have changed accordingly. We consider these factors and formulate our mechanistic models for the three phases.

Phase I (prior to January 23, 2020): During this period public awareness of the severity of the disease was low and cases (especially those of mild symptom) might be treated as common cold. Both pre-symptomatic cases (especially right before symptom onset) and symptomatic cases were infectious [8]. Considering the infectivity before and after the symptom onset might be different, we divided the infected people into two categories: E and I in phase I, and allowed the infectivity of class E to be less than that of class I. However, the mobility of class I might be lower than that of class E, thus the infectivity level of E might be comparable to I. The population N can be divided into four categories: susceptible (S), exposed and pre-symptomatic population (E), symptomatic population (I), and recovered population (R). The total population N equals $S + E + I + R$. The transmission dynamic model reads:

$$\begin{cases} \dfrac{dS}{dt} = -\dfrac{\beta S(kE+I)}{N}, \\ \dfrac{dE}{dt} = \dfrac{\beta S(kE+I)}{N} - qE, \\ \dfrac{dI}{dt} = qE - \gamma I, \\ \dfrac{dR}{dt} = \gamma I. \end{cases} \tag{2.1}$$

where, β is the transmission rate of disease, q is the rate at which exposed individual showing symptom, k controls the infectiousness of exposed pre-symptomatic individuals relative to symptomatic individuals, γ is the cure rate at which symptomatic individuals moving to recovered class. We ignore COVID-19 induced death since there were no deaths from diseases in the early stage and few deaths in the later stage in Anhui.

Phase II (between January 23 and February 6, 2020): public began to realize the seriousness of the epidemic after Wuhan lockdown on January 23 and lockdown in other cities in Hubei shortly after. The country also issued the epidemic situation announcement, and began to take quarantine measures, symptomatic cases went to the hospital to see a doctor and would be quarantined before confirmation. Therefore, after January 23, only pre-symptomatic cases in E class can effectively contact susceptibles to spread the disease. And there was also a decrease in the rate of exposure after January 23. Therefore, on the basis of model (2.1), we incorporate the policy intervention factors, and change the exposure rate to $\alpha_1\beta$. We assume all symptomatic cases were quarantined in quarantined class (Q), and σ is the cure rate at which quarantined cases moving to recovered class. Then SEQR model (2.2) reads:

$$\begin{cases} \dfrac{dS}{dt} = -\dfrac{\alpha_1 \beta S E}{N}, \\ \dfrac{dE}{dt} = \dfrac{\alpha_1 \beta S E}{N} - qE, \\ \dfrac{dQ}{dt} = qE - \sigma Q, \\ \dfrac{dR}{dt} = \sigma Q. \end{cases} \tag{2.2}$$

Phase III (after February 6, 2020): The centralized quarantined management measures for close contacts was adopted on February 6 in Anhui and resulted in a decline in the chance of contact among the population. Therefore, after February 6, the rate of exposure rate was changed to $\alpha_2\beta$. The model is the same as (2.2), except for α_2 replacing α_1.

3. Parameter estimation and numerical simulation

We adopted MCMC method implemented in R package to estimate parameters in models (2.1 and 2.2). We fit $I(t)$ or $Q(t)$ to the observed morbidity data. Note that our morbidity are daily current symptomatic cases. We obtained $\beta = 0.4029$, $q = 0.41$ (per day), $k = 0.78$ and $\gamma = 0.182$ (per day) in the first stage. Since during the first phase, individuals in E and I were not quarantined, both of them may spread the disease. We obtained the basic reproduction number during the phase 1: $\mathcal{R}_0 = k\frac{\beta}{q} + \frac{\beta}{\gamma} = 2.9764$. We did not distinguish imported cases and local cases and treat all cases as local. A substantial proportion of cases in phase I would be imported cases from Wuhan or other cities in Hubei. If the proportion of imported cases over the time during the phase I was relatively constant, our method is still relatively reasonable since the mortality is still a proxy of the local epidemic. Given that most of the infections occurred among family members, we presume the ratio of imported versus local case were relatively stable during phase I, and gradually vanished later. The actual and predicted values of the morbidity number during phase I are compared in Figure 3.

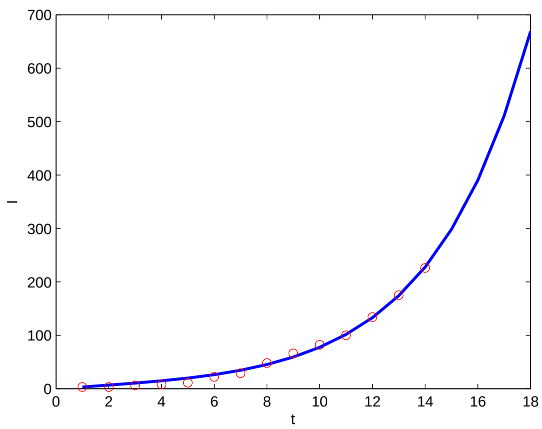

Figure 3. Morbidity number and actual value in phase I (actual value in red circle and predicted value in blue).

We used MCMC to fit model 2 to the morbidity number between January 23 and February 6, 2020. We obtained $\alpha_1\beta = 0.208$, $q = 0.24$, $\sigma = 0.04$. Using generation matrix method, the basic reproduction number was estimated $\mathcal{R}_0 = \frac{\alpha_1\beta}{q} = 0.8667$. This reduction in $\mathcal{R}_0$ was due to lockdown effects, holiday effects, and public being vigilant. The actual and predicted values of the morbidity number during phase II are compared Figure 4.

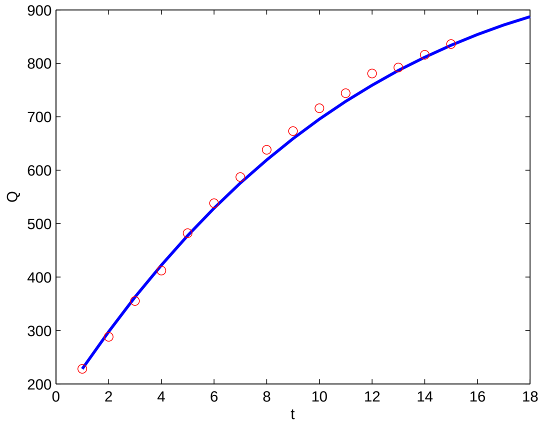

Figure 4. Morbidity number and actual value in phase II (actual value in red circle and predicted value in blue).

Similarly, we fitted model (2.2) (with α_2 replacing α_1) to the morbidity number after February 6, 2020. We obtained $\alpha_2\beta = 0.1311$, $q = 0.229$, $\sigma = 0.051$. The basic reproduction number is $\mathcal{R}_0 = \frac{\alpha_2\beta}{q} = 0.5725$. This reduction in $\mathcal{R}_0$ was due to further strengthened quarantine measure in Auhui.

The comparison between actual value and predicted value of daily morbidity in phase III is show in Figure 5. Under the policies of Wuhan lockdown and centralized quarantined management in Anhui, the morbidity number has been reduced and sharply decreased.

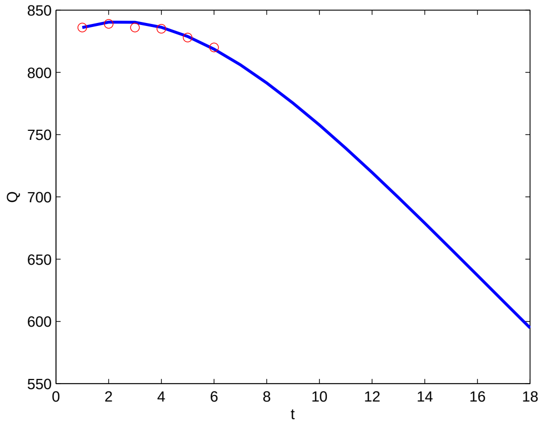

Figure 5. Morbidity number in Phase III (actual value is red circle and predicted value is blue).

The parameter descriptions and estimates are listed in Table 2.

Table 2. Parameter description and estimate in Anhui province 2019–2020.

Parameter	Description	Values	Methods
β	phase I transmission rate per capita per day	0.4029	MCMC
q	rate of symptom onset of pre-symptomatic exposed cases	0.229–0.41	Data and MCMC
$\alpha_1\beta$	phase II transmission rate per capita per day	0.208	MCMC
σ	The cure rate	0.04–0.051	Data and MCMC
$\alpha_2\beta$	phase III transmission rate per capita per day	0.1311	MCMC

Finally, the real data during the whole epidemic period compared to the predicted value of the three-stages model is shown in Figures 6 & 7.

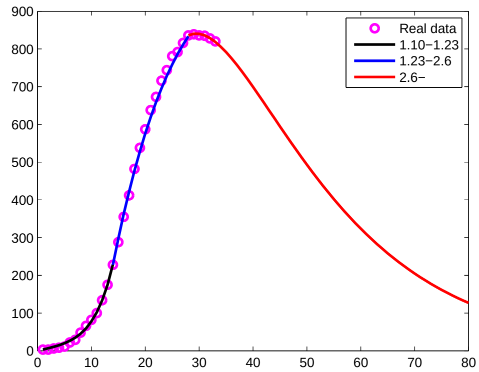

Figure 6. Comparison between the morbidity number per day and the observed data.

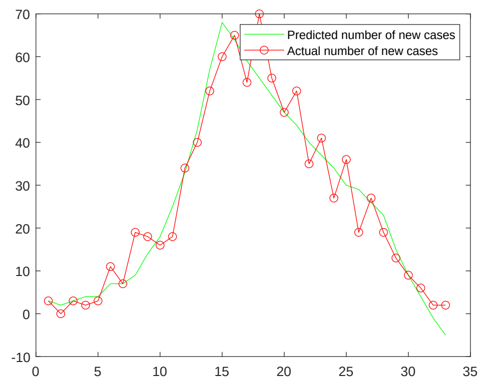

Figure 7. Comparison between the number of new cases per day in the three stages and the observed data.

In order to view the effect of relevant measures taken by the government on epidemic prevention and control more intuitively, we considered a hypothetical scenario that neither Wuhan lockdown nor centralized quarantine measures were in place and compared the morbidity number with and without government measures in Figure 8.

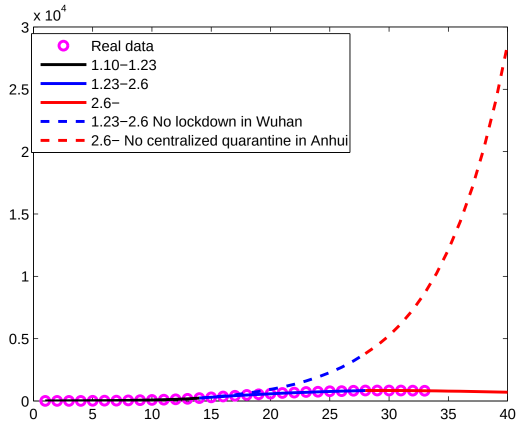

Figure 8. Comparison of the morbidity number per day in Anhui province with or without control measures.

Furthermore, we considered a second hypothetical scenario when the Wuhan lockdown was in place but no centralized quarantine measures in Anhui. The comparison of morbidity number of patients in Anhui is shown in Figure 9. It can be seen from the graph that the number of new cases of COVID-19 would continue for a long time, rather than decrease as seen now.

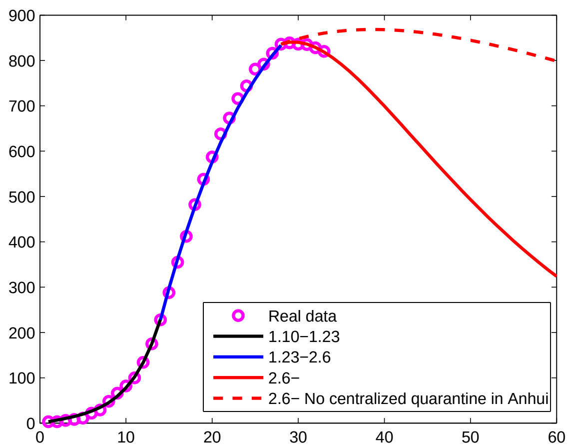

Figure 9. Comparison of the COVID-19 cases in Anhui province with Wuhan lockdown but no centralized quarantine.

To sum up, it can be seen that Wuhan lockdown and centralized quarantine measure in Anhui had played a timely and effective role in the prevention and control of the epidemic.

4. Conclusion and discussion

Based on the results of time series analysis, and the different measures taken by the government as time nodes, this paper establishes a transmission dynamics model which is divided into three phases. It can be seen from the analysis results of each stage that under the gradual implementation of various government measures, the basic reproduction number of the transmission of COVID-19 in Anhui decreased gradually, from 2.9764 to 0.8667 and then to 0.5725. These result of simulations indicate that the epidemic in Anhui has been controlled and is now in a critical period of transmission interruption. At the same time, this also further showed that when a major epidemic occurs, the timely implementation of national policies plays an important role in the control of the epidemic.

In addition, from the analysis of the data and model results, it can be concluded that the Wuhan lockdown in the second phase reduced the rate of exposure of transmission substantially in Anhui province by 48.37%. A centralized quarantine measure adopted in Anhui reduced the rate of exposure by 67.46% or 36.97%) compared with the rate of exposure rate in phase I or phase II, respectively. Therefore, both Wuhan lockdown and Anhui timely centralized quarantine measures have greatly reduced the speed of the transmission. The concentration quarantine measures in Anhui caused a substantial decrease in the rate of exposure, thus should be considered timely and effective. From the perspective of public health, this work further illustrated the impact of national and provincial policies in public health.

Acknowledgments

This research is supported by National Natural Science Foundation of China (11401002,11771001), Natural Science Fund for Colleges and Universities in Anhui Province (KJ2018A0029), Teaching Research Project of Anhui University (ZLTS2016065) and Anhui Provincial Department of Science and Technology, Anhui Provincial Health Commission Emergency Research Project "Epidemiological and clinical characteristics of new coronavirus pneumonia" (No. 202004a07020002).

Conflict of interest

The authors declare there is no conflicts of interest.

References

1. J. S. Kahn, K. Mcintosh, History and recent advances in coronavirus discovery, *Pediatr. Infect. Dis. J.*, **24** (2005), S223–S227.
2. Z. A. Memish, N. Nishra, K. J. Olival, S. F. Fagbo, V. Kapoor, J. H. Epstein, et al., Middle east respiratory syndrome coronavirus in bats, Saudi Arabia, *Emerg. Infect. Dis.*, **19** (2013), 1819–1823.
3. M. E. Killerby, H. M. Biggs, C. M. Midgley, S. I. Gerber, J. T. Watson, Middle east respiratory syndrome coronavirus transmission, *Emerg. Infect. Dis.*, **26** (2020), 191–198.
4. J. Cohen, D. Normile, New SARS-like virus in China triggers alarm, *Science*, **367** (2020), 234–235.

5. D. S. Hui, E. I. Azhar, T. A. Madani, F. Ntoumi, R. Kock, O. Dar, et al., The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health-The latest 2019 novel coronavirus outbreak in Wuhan, China, *Int. J. Infect. Dis.*, **91** (2020), 264–266.
6. World Health Organization (WHO), Coronavirus. Available from: https://www.who.int/health-topics/coronavirus.
7. K. O. Kwok, A. Tang, V. W. I. Wei, W. H. Park, E. K. Yeoh, S. Riley, Epidemic Models of Contact Tracing: Systematic Review of Transmission Studies of Severe Acute Respiratory Syndrome and Middle East Respiratory Syndrome, *Comput. Struct. Biotechnol. J.*, **17** (2019), 186–194.
8. C. Castillo-Chavez, C. W. Castillo-Garsowa, A. Yakubu, Mathematical Models of Isolation and Quarantine, *J. Am. Med. Assoc.*, **290** (2003), 2876–2877.
9. B. Tang, Estimation of the transmission risk of the 2019-ncov and its implication for public health interventions, *J. Clin. Med.*, **9** (2020), 1–13.
10. T. Chen, J. Rui, Q. Wang, Z. ZHao, J. Cui, L. Yin, A mathematical model for simulating the transmission of Wuhan novel Coronavirus, *bioRxiv*, (2020).
11. G. M. Leung, K. Leug, Crowdsourcing data to mitigate epidemics, *Lancent Dig. Health*, (2020).
12. M. Chinazzi, J. T. Davis, M. Ajelli, C. Gioannini, M. Litvinova, S. Merler, et al., The effect of travel restrictions on the spread of the 2019 novel cornavirus(2019-nCov) outbreak, *Science*, (2020), 1–12.
13. S. Zhao, S. S. Musa, Q. Lin, J. Ran, G. Yang, W. Wang, et al., Estimating the unreported number of novel coronavirus (2019-nCov) cases in China in the first half of January 2020: a data-driven modelling analysis of the early outbreak, *J. Clin. Med.*, **9** (2020), 388.
14. C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, et al., Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, *Lancet*, **395** (2020), 497–506.
15. P. Wu, X. Hao, E. H. Y. Lau, J. Y. Wong, K. S. Leung, J. T. Wu, et al., Teal–time tentative assessment of the epidemiological characteristics of novel coronavirus infections in Wuhan China, as at 22 January 2020, *Euro. Surveill.*, **25** (2020).
16. R. Kozma, N. K. Kasabov, J. S. Kim, A. Cohen, Integration of connectionist methods and chaotic time-series analysis for the prediction of process data, *Int. J. Intell. Syst.*, **13** (1998), 519–538.
17. A. K. Wagner, S. B. Soumerai, F. Zhang, D. Ross-Degnan, Segmented Regression Analysis of Interrupted Time Series Studies in Medication Use Research, *J. Clin. Pharm. Ther.*, **27** (2002), 299–309.