Suspected Close Contacts as the Pilot Indicator of the Growth Trend of Confirmed Population During the COVID-19 Pandemic: A Simulation Approach

Sisi Huang1#, Anding Zhu2#, Yan Wang3, Yancong Xu4, Lu Li5, Dexing Kong1,2,3

Abstract

Background: Regarding to the actual situation of the new coronavirus disease 2019 epidemic, social factors should be taken into account and the increasing growth trend of confirmed populations needs to be explained. A proper model needs to be established, not only to simulate the epidemic, but also to evaluate the future epidemic situation and find a pilot indicator for the outbreak.

Methods: The original susceptible-infectious-recover model is modified into the susceptible-infectious-quarantine-confirm-recover combined with social factors (SIDCRL) model, which combines the natural transmission with social factors such as external interventions and isolation. The numerical simulation method is used to imitate the change curve of the cumulative number of the confirmed cases and the number of cured patients. Furthermore, we investigate the relationship between the suspected close contacts (SCC) and the final outcome of the growth trend of confirmed cases with a simulation approach.

Results: This article selects four representative countries, that is, China, South Korea, Italy, and the United States, and gives separate numerical simulations. The simulation results of the model fit the actual situation of the epidemic development and reasonable predictions are made. In addition, it is analyzed that the increasing number of SCC contributes to the epidemic outbreak and the prediction of the United States based on the population of the SCC highlights the importance of external intervention and active prevention measures.

Conclusions: The simulation of the model verifies its reliability and stresses that observable variable SCC can be taken as a pilot indicator of the coronavirus disease 2019 pandemic.

Keywords: COVID-19; SIR model; social factors; numerical simulation; suspected close contacts; confirmed case; temporary hospital

Introduction

The ongoing pandemic of coronavirus disease 2019 (COVID-19), caused by the novel severe acute respiratory syndrome coronavirus-2, is spreading throughout the world. As of April 5, 2020, COVID-19 has recorded over 1.2 million confirmed cases globally and the number of newly confirmed cases in a single day exceeded 100,000 for the first time.¹ The situation of COVID-19 is severe in many countries. Evaluation of the epidemic trend and scale is most important for governments to fight against COVID-19. Facing the huge scale of the potential susceptible population, the Google search engine is one of the data sources used for evaluation.² However, the prediction of users’ search footprint is not accurate and timely. Mathematical models are more suitable tools to predict the growth trend and the inner relationship among observable variables.

As early as January 17, 2020, the World Health Organization Collaborating Centre for Infectious Disease Modeling and the Medical Research Council’s Centre for Global Infectious Disease Analysis of the Imperial College of England estimated that 1723 people in Wuhan were potentially infected with severe acute respiratory syndrome coronavirus-2 as of January 12, and warned that the outbreak in Wuhan may be much greater than the current number reports.³ Early in the Wuhan outbreak, there was no accurate clinical testing method to confirm the infected. Computed tomography images of suspected clinical cases were therefore regarded as a relatively reliable method for final diagnosis.⁴ Although Wang et al. reported that spiral computed tomography is a sensitive examination method to make an early diagnosis of COVID-19,⁵ the golden criteria for confirmation counts on nucleic acid detection.⁶ However, it was difficult to evaluate the actual scale of infections because of the lack of nucleic acid testing resources during the early outbreak.

Therefore, some mathematical modeling efforts were made to solve this problem. Chen and Yu developed a second derivative model to characterize the COVID-19 epidemic during the first two months.⁷ Roosa et al. developed a short-term forecasting

Editor: Stijn van der Veen

¹ School of Mathematical Sciences, Zhejiang University, Hangzhou 310027, China. ² School of Management Engineering, Zhejiang Gongshang University, Hangzhou 310018, China. ³ School of Mathematical Sciences, Chongqing Normal University, Chongqing 401331, China. ⁴ Department of Mathematics, Hangzhou Normal University, Hangzhou 311121, China. ⁵ College of Basic Medical Sciences, Zhejiang Chinese Medical University, Hangzhou 310053, China.

² Corresponding authors: Dexing Kong, School of Mathematical Sciences, Zhejiang University, Hangzhou 310027, P.R. China. E-mail: dkong@zju.edu.cn

³ SH and AZ contributed equally to this study.

Author contributions: Sisi Huang and Dexing Kong designed the study. Sisi Huang and Anding Zhu wrote the paper. Yan Wang collected the data. All the authors participated in the performance of the research and data analysis. All the authors approved the final version for publication.

Funding: No funding was received for this study.

Conflicts of interest: The authors declared no conflicts of interest.

Copyright © 2020 the Author(s). Published by Wolters Kluwer Health, Inc. Infectious Microbes & Diseases (2020) Vol:No

Received: 14 April 2020 / Accepted: 26 April 2020

http://dx.doi.org/10.1097/IMB.0000000000000266
model and tried to predict the early growth of the Wuhan epidemic outbreak from February 5 to February 24.\(^9\) Similarly, Kucharski et al. adopted the random walk model to simulate the spreading trajectory in Wuhan. By using the history data of confirmed cases, they estimated the median daily reproduction number and observed the decline of this key parameter from 2.35 to 1.05 as the travel restrictions were introduced.\(^9\) The susceptible-exposed-infectious-removed model is a classical model. Lin et al. adopted the susceptible-exposed-infectious-removed model to simulate the effect of individual reactions and governmental actions.\(^10\) In addition, Fanelli and Piazza employed the susceptible-exposed-infectious-removed model to simulate the growth trends in China, Italy, and France.\(^11\) On the other side, the COVID-19 epidemic is difficult to accurately predict because of the nonlinear nature of the epidemic interventions.\(^12\)

Agent hospitals (“Fangcang hospital”) are the temporal hospitals built to serve the suspected close contacts (SCCs) who need to be isolated and under clinical observation. From the very beginning of the Wuhan outbreak, beds in the agent hospitals had been increasing. Similarly, the British built Excel coronavirus hospitals and one of them named Nightingale has 4000 beds to be used as an ideal temporary site to treat patients. Therefore, the scale of SCC is an observables variable for the anti-coronavirus administrator. In the present study, we intend to develop a mathematical epidemic model to investigate the relationship between the observable SCC and the final outcome of the growth trend of confirmed cases. The simulation results based on the actual data of China, South Korea, Italy, and U.S. will reveal the effect of the SCC pilot indicator.

**Brief introduction to the SIDCRL model**

The susceptible-infectious-recover (SIR) model is the classical typical compartment model.\(^13,14\) The entire population is divided into three agents: the susceptible agent S, infected agent I, and recovered agent R. However, applying these classic models directly does not consider complex social factors such as external interventions, which makes it improper to describe and simulate the current situation of COVID-19. However, other models established during the epidemic do not capture or stress the exact reason for the rapid growth of the confirmed cases per day. In fact, Academician Nanshan Zhong also has emphasized the urgent need to clarify the dynamic characteristics of viral transmission in the early days,\(^8\) which can also be illustrated by modeling and simulation. Chen and Yu introduced the close-population agent in their second derivative model.\(^7\) Similarly, we introduce the SCC agent into the classical SIR model. In addition, because the Chinese government takes the tough regulation policy, we cannot ignore the impact of social factors, such as external intervention and isolation. We name the newly developed model as SIDCRL model, which takes into account the SCC agent (S), the infected but not isolated agent (I), the infected, not confirmed but isolated agent (D), the confirmed and isolated agent (C), the cured agent (R), and social factor (L) (see the Methods section).

The present study selected four countries, that is, China, South Korea, Italy, and the United States as the representative countries. China is currently considered as the first country that has controlled the epidemic as well as suffered the first outbreak. In particular, the cases confirmed in Hubei province accounts for most of the total number of the entire country. Therefore, we select Hubei as the representative of China and the simulation conducted in Hubei shows the epidemic situation in China. South Korea is a typical country in Asia that has also quickly controlled the epidemic. Situations in Italy and United States represent Europe and North America, respectively. In comparison, the regulation policies of these two countries are not as tough as the former Asian countries.

**Results**

**The reliability of the SIDCRL model**

The transitions among agents of the SIDCRL model is established based on the nature of COVID-19 as well as the external social factors (see the Methods section). The simulation in Hubei province fits the actual data well (Figure 1A). The effects of the social factors and the cured rate reflect the situation (see the

---

**Figure 1. Simulation of Hubei, China.** The actual data is as of March 9, 2020. The symbol corresponds to the variable in the model. C + R means the cumulative confirmed and Δ(S + R) indicates the newly confirmed per day. Both the solid line and the dashed line are the numerical simulation results. The red dots in subfigure (A) represent the confirmed case recorded by the actual data. The simulation reflects the reality and shows the validity of the model.
Methods section). In the early stage of the outbreak, the lack of timely interventions caused the saturation of medical resources, which then caused social aggregation of infections and cross infections. With the rapid follow-up of manpower and corresponding management measures, the number of the cured rose rapidly. This is reflected in the special function of the social factors and the increasing cured rate is illustrated by a sigmoid function. The curve of the newly confirmed per day is a major characteristic of the epidemic outbreak. Therefore, the figures (Figures 1B, 2B, 2D, and 3B) show that the rapid rise of the SCC is followed by the rapid increase of the number of newly confirmed per day with a certain lagging time.

Moreover, there is further evidence by observing the peak points. From the simulation result of Hubei, the number of SCC agent reached its peak of 7435 on January 17, and the simulated highest newly confirmed per day should be about 5195 people per day on February 16, with 30 days lagging (Figure 1B). Similarly, in the simulation of Korea, the maximum number of SCC is 16,240 on February 21 and the number of simulated newly confirmed cases reaches a maximum of 633 on March 2 with 10 days lagging, corresponding to that of the actual data of 851 on March 3 (Figure 2B). And in Italy, the maximum number of SCC is 7179 on March 9 and that of the simulated newly confirmed is 6529 on March 28 with 19 days lagging, corresponding to one peak of the actual data 6203 (the second largest) on March 26 (Figure 3B).

Interestingly, the number of SCC is not so high in Italy and the reason may lie in fast transferability between the state variables.

**Rapid rise of the SCC indicates the epidemic outbreak**

We investigated the relationship between SCC (S) and the newly confirmed per day Δ(C + R) (Figures 1B, 2B, 2D, and 3B). It should be noted that some agents from S will be infected and outflow into other compartments (see the Methods section). The curve of the newly confirmed per day is a major characteristic of the epidemic outbreak. Therefore, the figures (Figures 1B, 2B, 2D, and 3B) show that the rapid rise of the SCC is followed by the rapid increase of the number of newly confirmed per day with a certain lagging time.
The relation between SCC and the epidemic outbreak can be easily explained from our models. On the one hand, if SCC is large and then there is a greater infection base, which causes the rise of $SI$ in Eq. (1) (see the Methods section) so that the number of the newly confirmed is going up. On the other hand, the steady drop of SCC may herald the end of the epidemic as the daily number of newly confirmed cases also drops subsequently. Therefore, we may take SCC as a pilot indicator of the epidemic.

The prediction for the United States emphasizes the importance of controlling the SCCs

By simulating, we find that the U.S. has the greatest number of SCCs, which has mounted up to the order of $10^5$ since the beginning of May (Figure 3B). It implicates the severe challenge for the US to fight against and overcome the epidemic. For prediction we successively increase the number of SCC and thus get the correspondingly steeper growth trend of the cumulative confirmed (Figure 3A). According to our simulation as of April 1, in the worst scenario, the simulated cumulative confirmed cases could reach a maximum of 1,500,000 by the end of April. However, a minimum of 600,000 is also possible if the number of SCC is effectively controlled and reduced by home isolation and other active external interventions. Fortunately, what we already see from the simulation is that there is already a local peak of SCC on May 24. Therefore, suppose the lagging time is 20–30 days in the US and the local peak is exactly the global peak as we hope, then optimistically the peak point of the newly confirmed will be
around April 18 (±5 days). In conclusion, the simulation results again emphasize the importance of controlling and reducing the number of SCC.

**Discussion**

We establish a mathematical model called SIDCRL model with dynamic social factors illustrating epidemic situations in China, South Korea, Italy, and the United States. The model coincides with the actual situation and gives predictions. Also, it can be induced from the simulation that the increasing number of SCCs contributes to the epidemic outbreak.

**SIDCRL is a model with high scalability**

Numerous mathematical models have been proposed since the epidemic outbreak. But as it is indicated in some studies that the prediction using more complex models are not necessarily more reliable, we include all these social attributes in $L(t)$ instead of more subdivided parameters. As a result, the model can be portably applied in almost all areas around the world just by adjusting the system increase factors. In addition, the sigmoid function is used to describe the cured rate, which is obviously reasonable and necessary.

**Control measures should be taken to reduce the SCCs**

The situation in Hubei is representative of China’s epidemic development. During the period, China promoted a series of strong prevention and control measures to reduce the number of the SCCs. On January 23, Wuhan announced the “lockdown.” On January 23 and January 25, China began to build the Vulcan Hill and Raytheon Mountain hospital to solve the shortage of medical resources. By February 4, 11 agent hospitals were built, providing more than 10,000 beds. Patients were effectively quarantined and received treatment while the residents also tried to isolate themselves at home, which noticeably slowed down the epidemic in China since mid-February. China provides a valuable example for South Korea, Italy, the United States, and other epidemic centers. In particular, the simulation stresses the relation between the increasing number of SCC and the outbreak.

That is, during the epidemic, it is highly important to avoid contacting with people so as to reduce the risk to be infected.

**The number of the SCC can be viewed as a pilot indicator of the outbreak**

The definition and the detailed meaning of the SCCs have been described in the Methods. In fact, this is a measurable and recordable quantity. However, what we commonly know as the number of close contacts is just a cumulative number and ignores its coherent connection with the increasing daily or cumulative number of confirmed cases. Therefore, the present study calls on governments and the international community to pay attention to this important indicator and strictly record the number of SCC. Here, we want to put a simple way to record variable $S$ and control the number of the SCCs. Once an infected case is confirmed, list all persons who have had a significant exposure and add the number into SCC. Keep close supervision of contact follow-up. The agents can be removed from the system if they keep healthy for a certain isolation period of time and meanwhile subtract the corresponding number from SCC. In conclusion, the rigorous record of $S$ really kills two birds with one stone. First, the strict control of close contacts greatly reduces the spread of infections and viruses and at the same time takes its effectiveness in quarantine and isolation. Second, the curve trend of SCC can be analyzed to simulate the trend of the number of the newly confirmed per day. Actually, it could be a more practical and sensible method to predict the epidemic situation, which makes the number of SCC a pilot indicator of the growth trend of confirmed populations during the COVID-19 pandemic.

**Materials and methods**

The data of Hubei, China is collected from the National Health Commission of the People’s Republic of China and Hubei Provincial Health Commission. The data of South Korea, Italy, and the United States is collected from the real-time updates from Johns Hopkins University in the United States. All the simulations and analysis are conducted in MATLAB Release 2017b.

![State transition diagram of the SIDCRL model](image)
SIDCRL model

Differential equation system is used to characterize the COVID-19 epidemic. According to the current situation, the dynamic system considered the following states or agents: SCC agents $S(t)$, infected but not isolated agents $I(t)$, infected, not confirmed but isolated agents $D(t)$, confirmed and isolated agents $C(t)$, cured agents $R(t)$. The parameter $L$ is introduced as a social factor, and this modified SIR model is named as SIDCRL model. The dynamic system combines the social factors and the natural transition (shown in Figure 5).

Each day, there are $\beta SI$ SCC agents from compartment $S$ getting infected. But considering that there is a certain lagging time (hours to days) until the infected are confirmed, $(1 - a)\beta SI$ of these infected agents are not isolated and classified into compartment $I$, while $a\beta SI$ agents are isolated but not confirmed and classified into compartment $D$. As for the infected but not isolated compartment $I$, all agents are isolated every day, among which $(1 - g)al$ agents have not been confirmed and $gal$ agents are confirmed. Moreover, $bd$ agents in infected, not confirmed but isolated compartment $D$ are confirmed. At the same time, there are system increase factors $u_iL(i = 1, 2, 3)$ caused by social attribute entering into compartments of S.I.D.C, respectively. Finally, among the patients C who are diagnosed and quarantined, there are $\gamma C$ agents getting cured every day and being classified into compartment $R$.

Based on the above analysis, we can easily obtain the differential equations with the state variables $S, I, D, C, R$ and relevant parameters:

$$
\begin{align*}
\frac{dS}{dt} &= u_1L(t) - \beta SI - (1 - a)\beta SI + u_2L(t) - u_1, \\
\frac{dI}{dt} &= a\beta SI - bD + (1 - g)al + u_2L(t), \\
\frac{dD}{dt} &= g \times (1 - a)\beta SI, \\
\frac{dC}{dt} &= gal + bD - \gamma(t)C, \\
\frac{dR}{dt} &= \gamma(t)C.
\end{align*}
$$

(1)

where the actual meaning in details of each state variable is described as follows:

1. SCC $S(t)$: People who have had close contact with patients and are not immune to the virus will be listed as SCCs. This is a zero-based quantity and increases only when the newly confirmed case occurs. People such as the patient’s family or relatives and hospital staff are very likely to be recorded as SCCs. The dynamic of $S$ considers the factors of social aggregation and isolation, and is two-way connected with the entire population outside this system, which are generally described by $u_1L(t)$;

2. Infected, not isolated persons $I(t)$. These infected people are not isolated. They infect the SCCs and cause the virus to spread;

3. Infected, not confirmed but isolated persons $D(t)$. This group is still under medical observation and is an important source of the newly confirmed;

4. Confirmed and isolated persons $C(t)$. It is the difference between the cumulative number of the confirmed and that of the cured persons. By definition it includes deaths;

5. Cured persons $R(t)$. It is the cumulative number of people cured by medical treatment.

In equation (1), the actual meaning of each parameter is as follows:

1. $\beta$ is the infection rate reflecting natural transition.

| Parameter | Value | Parameter | Value |
|-----------|-------|-----------|-------|
| Infection rate $\beta$ | 0.0001 | Isolation rate $a$ | 0.5 |
| Scaler factor $u_1$ | 2885 | Isolation rate $a$ | 0.5 |
| Scaler factor $u_2$ | 1200 | Confirm rate $g$ | 0.3 |
| Scaler factor $u_3$ | 10 | Confirm rate $b$ | 0.3 |
| Peak point $h_1$ | 20 | Peak value $K$ | 0.055 |
| Rate of change $\sigma$ | 0.2 | Initial value $P_0$ | 0.02 |
| Rate of change $\epsilon$ | 0.2 | Initial point $h_2$ | 38 |

\[ L(t) = e^{-(t-h_1)/\sigma^2}, \]

(2)

where $h_1$ is the peak point and $\sigma$ controls the rate of change.

Assumptions and parameter estimation

The model does not consider the number of unconfirmed cases caused by insufficient kits. The confirmed patients are assumed to be isolated and cannot infect others. The isolation rate and confirm rate are assumed to be similar in Hubei, China as the rest of the infected countries. The training data set is selected from the most recent period of data, which is supposed to be less susceptible to distortion. The parameters are estimated by the idea of least-squares fitting and parameter values simulated in Hubei, China are listed in Table 1.

Numerical simulation

The differential equation system (1) is a high-dimensional nonlinear non-autonomous system so that its explicit solution is not easy to obtain. But the discrete-time simulation can be performed using numerical simulation. The equations of S.I.D.C. R can be viewed as the state transfer functions in differential form. Place the initial time and initial value of the simulation and the value of each state variable at the next moment can be determined by the state transition functions and the current variable value.

To be specific, the equation $S = u_1L(t) - \beta SI$ can be written as

$$
\frac{ds}{dt} = (u_1L(t) - \beta SI)dt.
$$

(4)

Notice $S(t + dt) \approx S(t) + ds$, by Eq. (4) we get

$$
S(t + dt) \approx S(t) + (u_1L(t) - \beta SI)dt.
$$

(5)
Acknowledgment

The authors thank Dr. Chenhui Qiu for his valuable discussions and suggestions.

References

[1] Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis 2020;20(5):533–534.

[2] Husnayain A, Fuad A, Su E. Applications of google search trends for risk communication in infectious disease management: a case study of COVID-19 outbreak in Taiwan. Int J Inf Dis 2020;95:221–223. DOI: 10.1016/j.ijid.2020.03.021.

[3] Natsuko Imai, Ilaria Dorigatti, Anne Cori, Christl Donnelly, Steven Riley, Neil M. Ferguson. Estimating the potential total number of novel Coronavirus cases in Wuhan City, China. Imperial College London (22-01-2020), doi: https://doi.org/10.25561/77150.

[4] Zhu Y, Liu Y-L, Li Z-P, et al. Clinical and CT imaging features of 2019 novel coronavirus disease (COVID-19). J Infect 2020.

[5] Wang K, Kang S, Tian R, Zhang X, Wang Y. Imaging manifestations and diagnostic value of chest CT of coronavirus disease 2019 (COVID-19) in the Xiaogan area. Clin Radiol 2020;75(5):341–347.

[6] Shen M, Zhou Y, Ye J, et al. Recent advances and perspectives of nucleic acid detection for coronavirus. J Pharm Anal 2020;10(2):97–101.

[7] Chen X, Yu B. First two months of the 2019 coronavirus disease (COVID-19) epidemic in China: real-time surveillance and evaluation with a second derivative model. Glob Health Res Policy 2020;5(1):1–9.

[8] Roosa K, Lee Y, Luo R, et al. Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020. Infect Dis Model 2020;5:256–263. DOI: 10.1016/j.idm.2020.02.002.

[9] Kucharski AJ, Russell TW, Diamond C, et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. Lancet Infect Dis 2020;20(5):553–558.

[10] Lin Q, Zhao S, Gao D, et al. A conceptual model for the coronavirus disease 2019 (COVID-19) outbreak in Wuhan, China with individual reaction and governmental action. Int J Infect Dis 2020;95:231–236. DOI: 10.1016/j.ijid.2020.02.058.

[11] Faneli D, Piazza F. Analysis and forecast of COVID-19 spreading in China, Italy and France. Chaos Solitons Fractals 2020;134:109761. DOI: 10.1016/j.chaos.2020.109761.

[12] Roda WC, Varughese MB, Han D, Li MY. Why is it difficult to accurately predict the COVID-19 epidemic? Infect Dis Model 2020;5:271–281. DOI: 10.1016/j.idm.2020.03.001.

[13] May RM, Anderson RM. Population biology of infectious diseases: Part I. Nature 1979;280:361–367. DOI: 10.1038/280361a0.

[14] Jones DS, Plank M, Sleeman BD. Differential equations and mathematical biology. Chapman & Hall/CRC Mathematical and Computational Biology. 2nd ed. CRC Press; 1983: 411–413.

[15] Guan WJ, Ni ZY, Hu Y, et al. Clinical characteristics of coronavirus disease 2019 in China. N Engl J Med 2020;382(18):1708–1720.

[16] Linton NM, Kobayashi T, Yang YC, et al. Incubation period and other epidemiological characteristics of 2019 novel coronavirus infections with right truncation: a statistical analysis of publicly available case data. J Clin Med 2020;9(2):E538.

[17] Lin J, Duan J, Tan T, Fu Z, Dai J. The isolation period should be longer: lesson from a child infected with SARS-CoV-2 in Chongqing, China. Pediatr Pulmonol 2020;55(6):E6–E9.

[18] Tang B, Xia F, Tang S, et al. The effectiveness of quarantine and isolation determine the trend of the COVID-19 epidemics in the final phase of the current outbreak in China. Int J Infect Dis 2020;95:288–293. DOI: 10.1016/j.ijid.2020.03.018.

[19] National Health Commission of the People’s Republic of China. Available from: http://www.nhc.gov.cn/. Accessed April 10, 2020.

[20] Hubei Provincial Health Commission. Available from: http://wjw.hubei.gov.cn/. Accessed April 10, 2020.