Epidemiological Modeling Analysis Reveals the Transmission Potential of COVID-19 Asymptomatic Patients: A Prospective Study of Epidemiological Transmission in America

Yuhang Zhao  
Shandong University

Zhonglei Ji  
Affiliated Hospital of Weifang Medical University

Ping Wang  
Shandong Academy of Chinese Medicine

Xiangwei Meng  
Shandong University

Zhixiang Ma  
Shandong University

Xiyuan Li  
Shandong University

Haihang Sun  
Shandong University

Wei Li (✉️ cindy@sdu.edu.cn)  
School of Control Science and Engineering, Shandong University

Research article

Keywords: COVID-19, Asymptomatic, Mathematical model, America

Posted Date: November 9th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-103012/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Epidemiological modeling analysis reveals the transmission potential of COVID-19 asymptomatic patients: a prospective study of epidemiological transmission in America

Yuhang Zhao¹, Zhonglei Ji², Ping Wang³,⁴, Zhixiang Ma¹, Xiangwei Meng¹, Xiyuan Li¹, Haihang Sun¹, Wei Li*¹

1. Biomedical Engineering Institute, School of Control Science and Engineering, Shandong University, Jinan, Shandong Province, 250061, China

2. Affiliated Hospital of Weifang Medical University, Weifang, Shandong Province, China

3. State Key Laboratory of Precision Measurement Technology and Instruments, Tianjin University, Tianjin, 300072, China

4. Shandong Academy of Chinese Medicine, Jinan, Shandong Province, 250014, China

*Corresponding Author:

Wei Li, Email: cindy@sdu.edu.cn
Abstract

Background: The asymptomatic of COVID-2019 are getting more and more attention from all walks of life. Now, America has become the epicenter of the pandemic, with more reported cases and deaths than other regions of the world. Studying the development asymptomatic populations may play a key role in managing the outbreak effectively.

Methods: We propose a new model to predict the course of the epidemic and simulate the transmission of the asymptomatic. The model considers seven stages of infection: susceptible (S), exposed (E), infected (I), asymptomatic (A), confirmed (C), recovered (R), dead (D), we named it as SEIACRD. We used a model to study the interaction between asymptomatic patients and viral transmission.

Result: Our model confirms that about 12 million people will be infected with the virus. Changes in mortality rates will be volatile, first falling, then rising. Not only the number of patients, but also the spread of the epidemic will be affected by the ability to detect asymptomatic persons. Contact with asymptomatic infected patients also significantly promoted the spread of the virus. But these methods had no significant effect on changes in patient mortality.

Conclusion: American Asymptomatic patients have a strong interaction with epidemic transmission. They are no less at risk of transmission than symptomatic patients. In terms of controlling the number of infections, efforts to improve detection capacity are more effective than simply suppressing the spread of the virus. Extensive testing and effective social distancing measures should be taken to protect more people from the virus.

Keywords

COVID-19, Asymptomatic, Mathematical model, America

Introduction

COVID-19 is one of the most dangerous epidemics in history. A group of researchers, using mathematical modeling to study the spread of infectious diseases in Italy\(^1\). They argue that many people who carry the virus have no symptoms, and that their role in spreading the virus is easily overlooked. The scholars analyzed the inflection data which is collected from Wuhan\(^2\). They draw the conclusion that even if the patient has no symptoms, they still have a high probability of passing the virus to the next person. Currently, the epidemic is getting worse in the United States.
Mathematical model plays an important role in the implementation of epidemic prevention measures\textsuperscript{3,4}. Common infectious disease models are divided into SI, SIR, SIRS, and SEIR\textsuperscript{5,6}. As the transmission of the SARS in 2003, the patients who are infected with novel coronavirus will secrete enough antibodies to protect themselves against a second infection\textsuperscript{7,8}. Some scholars pointed out that the incubation period of new pneumonia is usually about 5 to 20 days\textsuperscript{9}. Median incubation period was 3.0 days (range 0-24 days)\textsuperscript{10}. This feature will change the tendency of the current infected number\textsuperscript{11,12}. The COVID-19 data sources we used were collected by the COVID Tracking Project, an organization that records epidemic data in the United States. The data source has sufficient authority and influence to meet the requirements of our research.

**Methods**

According to the data source provided by the COVID Tracking Project in the United States, the patient test results were classified into positive, negative, pending, hospitalized, in ICU Cumulative, recovered, dead, and so on. Due to the burden on the medical system, asymptomatic populations have not yet been classified as a separate group to statistics. In our research, asymptomatic infected persons (A) and confirmed patients (C) were introduced into the mathematical model. The model considers seven stages of infection: susceptible (S), exposed (E), infected (I), asymptomatic (A), confirmed (C), recovered (R), dead (D).

The model simulation will be carried out under the following assumptions. The total population does not change during the period of the epidemic. We believe that there is a dynamic equilibrium between the number of deaths and births, if we don’t consider the influence of the COVID-19, during this phase. We can assume the changes in different populations as the differential equation:

\[
S'(t) + R'(t) + I'(t) + E'(t) + A'(t) + C'(t) + D'(t) = 0.
\]

So the total population is a constant \( N \), which is the population of the country. So, we can expression it as

\[
S(t) + E(t) + I(t) + A(t) + C(t) + R(t) + D(t) = N.
\]

We believe that, when anyone of the population can be classified as a part of \( S(t) \), \( R(t) \) and \( D(t) \), and the sum of \( E(t) \), \( I(t) \), \( A(t) \) and \( C(t) \) is equal to 0, there will be no new virus transmission route in the society and the epidemic will end. It can be expressed as a formula \( S(t) + R(t) + D(t) = N \) and \( E(t) + A(t) + I(t) + C(t) = 0 \).

The research mainly studied asymptomatic infection of the virus. Severe patients,
mild patients and other patients were included in the $I(t)$ population and were not distinguished any more. $C(t)$ is the people who has been attacked by the virus and detected by the test. This group of the infected person is the subset of the $I(t)$ and $A(t)$. They will get more attention from doctors and their families, so they have less chances to contact with others. It is really helpful to reduce the coefficient of the transmission and promote the opportunity. So, We need to separate them from the common inflected citizens. Designing the group of $C(t)$ separately is helpful to improve the accuracy of the model.

The SEIACRD dynamical system consists of seven ordinary deferential equations, describing the evolution of the population in each stage over time:

\[
\frac{\partial S}{\partial t} = -S(t)[\alpha E(t) + \beta A(t) + \mu(t)]
\]

(1)

\[
\frac{\partial E}{\partial t} = S(t)[\alpha E(t) + \beta A(t) + \mu(t)] - (r_1 + r_2)E(t)
\]

(2)

\[
\frac{\partial I}{\partial t} = r_1 \times E(t) - (r_3 + r_5 + r_7) \times I(t)
\]

(3)

\[
\frac{\partial A}{\partial t} = r_2 E(t) - (r_4 + r_6)A(t)
\]

(4)

\[
\frac{\partial C}{\partial t} = r_5 I(t) + r_6 A(t) - (r_8 + r_9)C(t)
\]

(5)

\[
\frac{\partial R}{\partial t} = r_7 I(t) + r_8 A(t) + r_9 C(t)
\]

(6)

\[
\frac{\partial D}{\partial t} = r_9 I(t) + r_8 C(t)
\]

(7)

In the discrete time series, the matrix is easy to be analyzed, so the seven differential equations are transformed into matrix equations:
Figure 1. A infectious disease mathematical model of SEIACRD based on improved SEIR model. S, susceptible; E, exposed; I, infected (symptomatic infected); A, asymptomatic (infected but without symptoms); C, confirmed patients (symptomatic or symptomatic infected, detected); R and D, recovered or dead from asymptomatic infection, which segment of the population will no longer be infected or infect with others.

The mathematical model is verified by the computer. The iterative method is used to calculate the matrix equation of the discrete-time series. The variations of different populations in the model over time are obtained and its approximate numbers are calculated. We compared these values with the data collected by Johns Hopkins University. The virus was poorly understood, leading to missed diagnoses and misdiagnoses in the initial stages of transmission. Therefore, February 27 was selected as the time node for the start of the simulation. We divided the Johns Hopkins data into a training set and a validation set. We selected data from February 27 to July 25 as the basis of the training set, there are 150 days. And the data from July 26 to August 31 were chosen as the validation set, about 37 days. A total of 187 days were used. The first 150 days were used to for parameter adjustment. The last 37 days are only used to check the accuracy.

We first ran the mathematical model for 250 days. The curve is the model result and the scatter line is the real data. Through comparison, we can find that the model has high accuracy, and can be in good agreement with the actual changes in the first 150 days. At the same time, we found some changes in the current infection trend on the 90th day of the record by the COVID-19 Tracking Project. After searching the literature, we believe that the social activities which are held in the United States have increased the possibility of human contact and changed the direction of the curve. The blue curve in the second figure represents the cumulative number of the positive over 150 days, which is the sum of E, I, A, C, R, and D in the model. The 150th day is July 25th. As we can see from the graph, there are about two million people infected. The calculated results are in good agreement with the actual situation.
The 37 days from July 25 to August 31 were the validation phase of the model's prediction. We keep the operating parameters of the model unchanged and compare the actual data with the predicted data. We found that the model and the data were consistent. The model was still successful in showing trends in the actual number of infections. This indicates that our model has good predictive ability.

**Figure 2.** Comparison of the model and actual data. The curve is the model result and the scatter line is the real data. (a) It is the result of the model calculation. (b) The result was compared with real data from February 27 to July 25. (c) The prediction accuracy of the model was verified with the data from July 25 to August 31.

**Result**

In our model, we have simulated the whole process of the epidemic from its development to its end. We changed the number of simulated days to 1000. We can see that the actual number of cases changes over the course of 90-100 days, as the number of cases gradually flattens out into a new period of growth. We have also revised the epidemic parameters accordingly. According to the model, the epidemic will still continue for a long time. And the number of people infected is expected to rise in the future. Regardless of influence and transmission with other countries, About 6 percent of the U.S. population could be affected by the virus. About 95.02 percent of them will be recovered. But about five percent of those who carry the virus are at risk of dying. As is shown in figure 2, by calculating each parameter in the model separately, the time deduction method of each part can be obtained. We can research the data more clearly when the X-component in the picture is lengthened. The number of deaths is a monotonically increasing function. In the 112\textsuperscript{th} day, the actual number was 2.25 million infected and 110,000 dead because of the virus. The number of deaths was about 5.3\% of the total. The number of asymptomatic infected persons was about 150,000 or 10\% of the current number of infected persons. But in the 150\textsuperscript{th} day, The actual number of infected people has risen to 4.16 million. And the death group has reached 138549 about 3.3\% of the total. If extended to the end of the epidemic, about 474 thousand infected people may lose their lives shown in the curve of death. It is about 3.9\% of the total number of people infected. The comparison shows that, the death rate will gradually decrease. Even if the number of infections and deaths will increase rapidly, but the overall mortality rate is lower than the current. According to our calculations, the number of infections will reach a peak in the 260\textsuperscript{th} day. Meanwhile, it will reach an inflection point in about 260 days. And then, COVID-19 will fade away gradually until it finally disappears.

**Figure 3.** Simulate the long-term evolution of COVID-19 outbreaks. (a) Important parameters of the epidemic evolution process. (b) Comparison of infection, cure, and death in models and data,
extrapolated to 1000 days.

The highest or lowest parameters of the transmission, which has occurred since the outbreak, were used to predict the spread of the epidemic in the future (Figure 4). If the future virus transmission is the same as the period with the most severe propagation parameters, the result of the run is shown in the two figures above. The transmission rate for asymptomatic patients increased by approximately 0.5% from the current value. It is represented that the asymptomatic are more likely to transmit the virus because they do not get sufficient attention, which can influence the overall spread of the epidemic. The following two graphs show the overall spread of the epidemic that the parameters are the same as the lowest value that has already occurred. It represents the rate of contact with asymptomatic infected people has dropped by one percent. It can be seen from the figure that although the parameters do not change much, the results of the two are quite different after the iterative operation. The group of less realized were eventually infected about 5.4 percent of American citizens. The expected mortality rate is around 3.8 percent. The result was a 1.8% increase over the final number of infections in the original model. The other group which focused on the infectiousness of asymptomatic patients, about 2.7% percent of folks are captured by the virus, which is 0.9% less than the previous group. The mortality rate is about 3.7 percent of those infected by the virus.

**Figure 4.** The different transmission probability of asymptomatic patients affects the future development of the epidemic. (a), (b) The future of the epidemic will be the same as the most dangerous situation at this stage. (c), (d) The future of the epidemic will be the same as the most favourable situation at this stage.

We also simulated the relationship between the detection ability of different asymptomatic patients and the spread of the epidemic (Figure 5). The above two figures simulate the spread of the epidemic if the detection capacity is the same as the current minimum detection capacity, that is, the ability to detect the virus cannot change as the number of patients increases. It is simulated that the asymptomatic can not be detected quickly because the ability to detect the virus can not follow the increased number of patients. Its ability to detect asymptomatic patients is reduced by 5%. The following two graphs simulate the route of the whole epidemic as people's detection ability and attention increased by 5%. After the iterative calculation, there are also large differences between the two conditions. The group with the lower detection rate not only shows a modest increase in the number of viruses carries, but also the duration of the outbreak will add about 100 days longer than normal. It has nearly 23 million more patients than the original model. The epidemic has a greater impact on society. The other group which increased the detection ability will end the outbreak early. And it will reduce about 300,000 patients compared to the original model. Therefore, it can be concluded that large-scale screening of asymptomatic
patients is conducive to the end of the epidemic and return to normal society earlier. In addition, no significant change in death probability was found during the whole simulation process. We note that in terms of the number of cases, the rate of CoVID-19 detection plays a bigger role than the rate of transmission, and the impact is more significant in terms of the duration of the epidemic. As shown in the figure, the low detection rate is more than 200 days earlier than the high detection rate. This shows that people have the capacity to end the epidemic as soon as possible if appropriate responses are taken.

**Figure 5.** The relationship between the detection ability of different asymptomatic patients and the spread of the epidemic. (a), (b) The lower level of testing can complicate efforts to contain the outbreak. (c), (d) Better detection capabilities will reduce the pressure from the outbreak.

**Discussion**

The research on the transmission rate of asymptomatic patients should receive sufficient attention from the government, institutions, and researchers. China, Japan and Italy have carried out extensive screening of asymptomatic patients. The disaster has been effectively controlled. As the management experience of these countries, research and treatment of asymptomatic patients may hold the key to solving the global epidemic. Our model shows that asymptomatic infected persons have a strong interaction with epidemic transmission under different social restriction levels. We draw the same conclusion with a group of research who extended the SEIR model to study the epidemic in Mexico. Once the patients show symptoms, both the patients themselves and the society will be vigilant. It allows people to consciously reduce the connection chances with the infection in symptomatic patients. Asymptomatic persons are not easy to be found because of a certain degree of concealment. Unless the nucleic acid tests are verified, these patients will continue to hide in the population. According to some Japanese experts’ analyses about the classification of patients on the diamond princess, asymptomatic patients accounted for 18% of the total. It is much smaller than we would expect for the asymptomatic population in the United States. We believe that the difference in sample size is responsible for this phenomenon. Our sample base is much larger than the diamond Princess. With such a large and hard-to-control population, the virus has a much greater chance of spreading in the United States than in the Diamond Princess. We also found that, compared to a smoother curve in most countries, the trend in the United States started to get worse in about 100 days. In other countries, treating COVID-19 is considered extremely important. But American politicians and the public seem never to pay enough attention to this disease. People do not take adequate protective measures when they get together. The exposure of ordinary people to the virus has greatly increased. Previous studies have shown that this increase in exposure
is very dangerous\textsuperscript{11}. If this situation continues without any protection action, the impact of the epidemic will be even more severe\textsuperscript{27}. The successful development and widespread use of a vaccine could change the spread of the epidemic. This is not taken into account in the model design. Social isolation measures, atmospheric factors\textsuperscript{28} and hospital affordability are also key factors affecting the spread of the virus\textsuperscript{29}. These factors are not taken into account in the model, which may result in a certain degree of deviation in the results. Governments need to adopt more effective statistical and testing measures\textsuperscript{30}. The world should work together\textsuperscript{31}. It is an effective method to minimize the risk of exposure to the virus. Effective self-protection by each individual is key to preventing asymptomatic carriers from spreading the virus.

**Conclusion**

In summary, our results suggest a strong interaction between asymptomatic infected persons and the spread of the epidemic in the United States. Asymptomatic infections are more difficult to manage because they carry the virus and do not exhibit clinical symptoms. If very strong detection strategies are implemented, the length of the total transmission time of the virus will be reduced, resulting in a significant reduction in the total number of expected infections. Effective measures to prevent asymptomatic infected persons from infecting others, such as appropriate social distancing measures, will also significantly reduce the risk of infection in normal people. But these two measures may not reduce the mortality rate in the affected population. We believe the spread of the virus in the United States will continue for some time. Extensive testing and effective social distancing measures should be taken to protect more people from the virus.

**Abbreviations**

COVID-19: Coronavirus Disease-2019; SEIACRD: a mathematical model which includes susceptible, exposed, asymptomatic, confirmed, recovered, dead; SI: a mathematical model which includes susceptible and infected; SIR: a mathematical model which includes susceptible, infected, and recovered; SIRS: a mathematical model which includes susceptible, infected, recovered and re-susceptible; SEIR: a mathematical model which includes susceptible, exposed, infected, recovered; SARS: Severe acute respiratory syndrome; WHO: World Health Organization.

**Acknowledgements**

We would like to thank Biomedical Engineering Institute, School of Control
Science and Engineering, Shandong University for supporting our study. We would also like to thank the Shandong Provincial Department of Science and Technology for its financial support.

Authors’ contributions

Yuhang Zhao proposed the model, performed the mathematical derivations, and wrote the paper. Zhonglei Ji and Wang Ping reviewed and revised the manuscript. Zhixiang Ma and Xiangwei Meng contributed to the development of data analysis methods. Xiyuan Li, Haihang Sun provided first-hand insight into the disease evolution. Wei Li guided the study and supervised the process of the research. All authors approved the manuscript.

Funding

This work was supported by the Major scientific and technological innovation projects of Shandong Province under Grant 2019JZZY020904, and by the Shandong Key Research and Development Project under Grant 2018GSF118136. We also acknowledge support from the National Natural Science Foundation of China (21728701).

Availability of data and materials

The data set we used is gathered from the COVID Tracking Project, open access repository: https://covidtracking.com/data/download. The data set analysed can be available from the corresponding author on reasonable request.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.
Author details

1 Biomedical Engineering Institute, School of Control Science and Engineering, Shandong University, Jinan, Shandong Province, 250061, China. 2 Affiliated Hospital of Weifang Medical University, Weifang, Shandong Province, China. 3 Shandong Institute of advanced technology Chinese academy of sciences, Jinan, Shandong Province, 250000, China. 4 State Key Laboratory of Precision Measurement Technology and Instruments, Tianjin University, Tianjin, 300072, China. 5 Shandong Academy of Chinese Medicine, Jinan, Shandong Province, 250014, China.

References


5 Yu X, Yang R. COVID-19 transmission through asymptomatic carriers is a challenge to containment. Influenza and Other Respiratory Viruses 2020.


A infectious disease mathematical model of SEIACRD based on improved SEIR model. S, susceptible; E, exposed; I, infected (symptomatic infected); A, asymptomatic (infected but without symptoms); C, confirmed patients (symptomatic or symptomatic infected, detected); R and D, recovered or dead from asymptomatic infection, which segment of the population will no longer be infected or infect with others.
Figure 2

Comparison of the model and actual data. The curve is the model result and the scatter line is the real data. (a) It is the result of the model calculation. (b) The result was compared with real data from February 27 to July 25. (c) The prediction accuracy of the model was verified with the data from July 25 to August 31.
Figure 3

Simulate the long-term evolution of COVID-19 outbreaks. (a) Important parameters of the epidemic evolution process. (b) Comparison of infection, cure, and death in models and data, extrapolated to 1000 days.
Figure 4

The different transmission probability of asymptomatic patients affects the future development of the epidemic. (a), (b) The future of the epidemic will be the same as the most dangerous situation at this stage. (c), (d) The future of the epidemic will be the same as the most favourable situation at this stage.
Figure 5

The relationship between the detection ability of different asymptomatic patients and the spread of the epidemic. (a), (b) The lower level of testing can complicate efforts to contain the outbreak. (c), (d) Better detection capabilities will reduce the pressure from the outbreak.