

Estimating the cure rate and case fatality rate of the ongoing epidemic COVID-19

Ying Diao^{1, 2#} Xiaoyun Liu^{3#} Tao Wang^{1#} Xiaofei Zeng⁵ Chen Dong⁶ Yuanming Zhang⁷
Changlong Zhou⁸ Xuan She¹ Dingfu Liu^{4*} Zhongli Hu^{1*}

1. College of Life Science, Wuhan University, Wuhan 430072, P.R. China;
2. College of Landscape Architecture and Life Sciences, Chongqing University of Arts and Sciences, Chongqing, 402160, P.R. China.
3. Wuhan Windoor Information Technology Co. Ltd., Wuhan, 430040, P.R. China
4. Hubei Academy of Agricultural Sciences, Wuhan, 430064, P.R. China
5. School of Medicine, Southern University of Science and Technology, Shenzhen, 518055, PR China
6. College of Biological Engineering, Henan University of Technology, Zhengzhou 450001, P.R. China;
7. College of Plant Science and Technology, Huazhong Agricultural University, Wuhan 430070, China
8. Yongchuan Hospital, Chongqing Medical University, Chongqing, 402160, P.R.China

#Contributed equally.

* Joint corresponding authors.

Correspondence to: Prof Dingfu Liu, Hubei Academy of Agricultural Sciences, Wuhan, 430064, P.R. China, wonderlau@126.com or Prof Zhongli Hu, College of Life Science, Wuhan University, Wuhan 430072, P.R. China, huzhongli@whu.edu.cn

Abstract

The epidemic caused by the novel coronavirus COVID-19 in Wuhan at the end of 2019 has become an urgent public event of worldwide concern. However, due to the changing data of the epidemic, there is no scientific estimate of the cure rate and case fatality rate of the epidemic. This study proposes a method to estimate the cure rate and case fatality rate of COVID-19. The ratio of cumulative discharges on a given day to the sum of cumulative discharges on a given day and cumulative deaths before j days is used to estimate the cure rate. Moreover, the case fatality ratio can also be estimated. After simulation calculations, j is statistically appropriate when it is 8-10, and it is also clinically appropriate. When j is 9, based on the available data, it is inferred that the cure rate of this epidemic is about 93% and the case fatality rate is about 7%. This method of estimating the cure rate can be used to evaluate the effectiveness of treatment in different medical schemes and different regions, and has great value and significance for decision-making in the epidemic.

Introduction

In December 2019, an epidemic caused by the new coronavirus COVID-19 occurred in Wuhan, China. The number of infected people and the spread of the virus have continued to increase, putting pressure not only on China but also the public health security worldwide (1). The rapid research on COVID-19 virology, epidemiology, and clinical medicine has played a positive role in the prevention and treatment of the disease (2-11). Considering the similarity of coronaviruses, the public habitually compares COVID-19 with SARS, especially regarding the data on cure rate (p) (CR) or case fatality rate (q) (CFR). In 2003, SARS emerged. The cumulative number of SARS diagnoses in Mainland China was 5,327, and the cumulative death toll was 349, with a CFR of 6.5 %; meanwhile, its cumulative number globally was 8,096, and the cumulative death toll was 744, with a CFR of 9.6 % (12). However, “there is currently no clear formula for the CR of patients with new pneumonia,” said by Jiao Yahui, the deputy director of the National Health and Medical Commission's Medical Affairs and Hospital Administration Bureau (13). The popular method of calculating the ratio of cumulative deaths to cumulative diagnosed cases to determine the CFR is accurate to confirm that an epidemic has ended, such as in the case of SARS. However, this estimation method has major flaws when applied in COVID-19, resulting in wrong judgments made by decision-makers regarding its future situation. For example, up to 24:00 on February 16, 2020,

the National Health and Medical Commission reported a total of 70,548 confirmed cases, 10,844 discharged patients, and 1,770 deaths. The CR was 15.37 %, and the CFR was 2.51 %. Furthermore, the hospital treatment rate was as high as 82.12 %. Clearly, the current CR and CFR is not the true CR and CFR caused by NOVID-19, nor it can accurately represent the current treatment trend of such a large number of hospitalized patients; if the cumulative discharge and the cumulative death are directly calculated, the CR and CFR would be 85.97 % and 14.03 %, respectively, which also cannot provide enough useful information. Therefore, we designed a method to calculate the CR of the new coronary pneumonia and then estimated the current CR and CFR to provide reference for the diagnosis, treatment, and control decision of this epidemic disease.

Method

Data sources

Data were acquired from various provinces, municipalities, and the country from January 20 to February 14. These data included the number of confirmed diagnoses, deaths, and discharges according to the National Health Commission of Wuhan, the National Health Commission of Hubei Province, and the COVID-19 Working Group of the State Council. However, few missing data from January 20 to January 23 were noted. The supplementary principle of the missing data is as follows: when the intermediate data are missing, the average of the two data (before and after) fills the missing data; when the data of Wuhan and the national data are missing, the data of Hubei Province are used as replacement (see Supplement Table S1).

Defined and corrected estimations of CR and CFR

"Cure" is the end of treatment and the end of death. If the epidemic has ended and all the patients have been treated, the cumulative number of discharged patients is $\sum x_i$, and the cumulative number of deceased patients is $\sum y_i$. When the cumulative number of diagnosed patients is $\sum (x_i + y_i)$, then the CR (\hat{p}) and CFR (\hat{q}) are estimated as follows correspondingly:

$$\hat{p} = \frac{\sum x_i}{\sum x_i + \sum y_i} \quad \text{and} \quad \hat{q} = 1 - \hat{p} = \frac{\sum y_i}{\sum x_i + \sum y_i} \quad (1)$$

Considering the occurrence of the epidemic situation of the novel pneumonia, some of patients have been cured and discharged, while some are still sick, some died, and some were

still under hospital treatment. The insight into whether the patients in the hospital will be cured or die remains unclear. Therefore, speculating the future trend while the epidemic is still ongoing by using the abovementioned formula to calculate the CR or the CFR directly is inappropriate.

According to the results of clinical studies, the novel pneumonia is a self-limiting disease with mild early symptoms, which may worsen after a week, and after the most dangerous period of time, the patient may recover gradually. According to the treatment plan, the patient must wait for all the symptoms to disappear, with two negative results of COVID-19 nucleic acid test, and be isolated for several days before discharge. In other words, the daily reported deceased patients were not the same batch as discharged patients. Hence, confirmers of the same period should consider that a time difference (j day) exists between the cured and death cases.

Therefore, the estimated CR (\hat{p}) should be the ratio of the cumulative number of discharges on the i thday ($\sum x_i$) to the sum of $\sum x_i$ and $\sum y_{i-j}$, expressed as follows:

$$\hat{p} = \frac{\sum x_i}{\sum x_i + \sum y_{i-j}} \quad (2)$$

where y_{i-j} is the number of cumulative deaths on the $(i-j)$ th day and j is the parameter that needs to be estimated. The estimated CFR (\hat{q}) should be the ratio of $\sum y_{i-j}$ to $\sum(x_i + y_{i-j})$, expressed as follows:

$$\hat{q} = 1 - \hat{p} = \frac{\sum y_{i-j}}{\sum x_i + \sum y_{i-j}} \quad (3)$$

Determination of the empirical value j

The empirical value j needs to be estimated before the estimation of the CR and CFR. In statistics, if we do not consider the discovery of specific drugs and medical means in the near future, the best estimates for each rate should be very close to its true value. In other words, these daily estimates are stable, and their variation is minimal. Thus, in this study, the empirical value j was determined according to the minimum variance (or coefficient of variation) among the daily estimates.

Results

j value

Because of the change in diagnostic criteria on February 12, resulting in a surge in

cases, only pre-February 11 data were selected. Taking the national data as an example, starting from February 11 and backtracking, in Equation (2), $j = 1, 2, 3, \dots, 15$ was calculated, and the CR was calculated daily. When $j = 10$, the value of CR was the most stable among the days, and the coefficient of variation was the smallest.

According to the published data, in addition to calculating the national CR, the same calculations were performed according to the division data of Wuhan, Hubei, Hubei except Wuhan, and the country except Hubei (Fig.1). The coefficient of variation of the estimated cure rate of wuhan, national, Hubei except Wuhan, Hubei and country except Hubei reached the minimum value at $j = 8, 10, 10, 10, 12$, respectively, indicating that this is the smallest difference in the estimated value of each day.

CR and CFR

According to the published data, $j = 8, 9, 10$ were selected. As shown in Fig. 2, the day-to-day CR from February 16 was calculated according to the division of the country, Wuhan, Hubei, Hubei except Wuhan, and the country except Hubei (see Table S2-4). The average CR is shown in Table 1.

According to Table 1, when $j=9$, if no specific drugs and better treatments emerge later and the pathogenicity of NOVID-19 is not significantly different, the CR of the new coronary pneumonia is approximately 93 %, while its CFR is approximately 7 %, which is relatively close to the CFR of SARS (6.6 %) in Mainland China. In Wuhan, the CR of COVID-19 is approximately 90 %, while its CFR is approximately 10 % or more.

Considering the large number of cases and the low CR in Wuhan, the CR in Hubei Province is lower than the national average, with only approximately 90%. The results of Hubei except Wuhan are similar to the national average. The CR in Hubei is the highest in the country, possibly related to factors such as the large number of imported cases, the different age structure of imported cases, relatively few patients, and better treatment conditions in provinces outside Hubei.

Discussion

***j* value**

COVID-19 happens all of a sudden, with no known cure. Hence, no treatment experience has been recorded, and no special drugs are currently known. It is mainly managed by symptomatic support treatment, all relying on enhanced immunity.

Assuming that the COVID-19 epidemic event satisfies the statistical randomness, the virus has no variation, and the daily CR and CFR should be basically stable, with approximately the same medical means and measures in a certain period of time and the estimated value of $j = 9$ able to meet this condition.

According to clinical experience, the time point of exacerbation was more concentrated on the 9th to 12th day of the disease course. After the exacerbation of the patient's condition, a rapid progress process takes place, and death easily occurs (14). In this article, $j = 9$ was consistent with the development of the disease course. Furthermore, according to the data released by the National Health and Medical Commission's press conference on February 4, the average length of hospitalization for discharged patients in country except Hubei is more than 9 days, whereas that in Hubei Province is 20 days (13). In this article, $j = 9$ was consistent with the actual treatment interval of the disease.

The CR estimate positively correlated with the j value size. The larger the j value, the greater the CR estimates; the smaller the j value, the smaller the CR estimates. In relation to the estimated value of CR, CFR will vary with the value of j . Therefore, to obtain accurate estimates of CR and CFR, we need to choose a j value, which is also worthy of further study.

However, using this strategy and method is certainly more reasonable than directly “calculating the CR and CFR for the COVID-19 in the country with the ratio of cumulative discharge ($\sum x$) and cumulative death ($\sum y$) to cumulative confirmed cases.”

Moreover, Equation (2) is not only appropriate and applicable during the epidemic; it is also an accurate estimate. The result estimated by Equation (2) is the same as that provided by Equation (1) at the end of the epidemic.

CR and CFR

This study corrects the current problems in calculating the CR of COVID-19. The CR nationwide is estimated to be 93%, and the CR of Wuhan is approximately 87%, which is considerably higher than the current official data. However, the official CR nationwide on February 16 was only 15.37%. The higher CR we estimated not only helps to boost morale in

the country against epidemic disease and alleviate the anxiety of the masses, but also provides a more scientific basis for decision-making on epidemic prevention and control.

Meanwhile, the estimated national CFR of COVID-19 obtained in this study is approximately 7%, which is close to SARS. Considering that the epidemic source is Wuhan, the estimated CFR of Wuhan exceeds 10%. According to the clinical analysis data of Wuhan Jin Yin-tan Hospital, the CFR of the first 41 patients was 15% (9, 15), and the CFR of the first 99 cases was 11% (16). Meanwhile, in the clinical characteristic analysis of the Central South Hospital of Wuhan University, the CFR was 4.3 % (17). We further analyzed the data and found that the number of discharges from Tongji Hospital accounted for 68 % of the study cases, compared with 34.1 %. However, a large number of undischarged cases remain uncertain. Therefore, the statistics of CFR in the Central South Hospital have a marked deviation compared with that in the Jin Yin-tan hospital. The rapid increase in the number of patients, inadequate treatment conditions, delayed diagnosis and delayed treatment, and inadequate knowledge of the virus in the early days in Wuhan affected the treatment effect.

If it is predicted according to the estimated national CFR, under the current medical conditions, the number of national deaths caused by COVID-19 is approximately 4,938 (= $70,548 \times 7\%$), wherein most of them are in Wuhan. In essence, this approach is using the estimated CFR of the first 9 days (February 7) to predict the possible deaths of currently diagnosed patients. The epidemic is continuing, and the number of confirmed patients is still increasing. If no major medical change transpires, the predicted value of future deaths will be extremely large. If the prediction is based on the estimated CFR in Wuhan, the number of deaths caused by COVID-19 is even more alarming. Moreover, all undiagnosed deaths caused by COVID-19 are not included.

This article predicts this high CFR and death toll based on the current medical level and existing data. This result reminds us that we cannot be blindly optimistic. Governments at all levels and people from all walks of life must not take it lightly and must grasp the final “window period.” The patients should be separated and receivable and should receive care and treatment in the best way possible. Identifying high-risk personnel in the treated population, strengthening treatment, and minimizing the CFR should be taken into account. In particular, we need to continue to actively provide medical resources to Wuhan and Hubei Province to control the rate of severe illness and reduce the number of deaths. As the pathogenicity of the coronavirus weakens, with the enhancement of medical technology,

improvement of medical conditions, and rich experience in treatment, patients with COVID-19 will be treated more effectively. Consequently, the CR will increase, and the CFR will decrease.

Conflict of Interest Statement

The authors declare no conflict of interest.

References

1. WHO. Statement on the second meeting of the International Health Regulations Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV) (2020). [https://www.who.int/news-room/detail/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-\(2005\)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-\(2019-ncov\)](https://www.who.int/news-room/detail/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov))
2. D. S. Hui, E. I Azhar, T. A. Madani, F. Ntoumi, R. Kock, O. Dar, G. Ippolito, et al., The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health -The latest 2019 novel coronavirus outbreak in Wuhan, China. *International Journal of Infectious Diseases*. **91**, 264-266 (2020).
3. D. T. Andrea, Outbreak of a novel coronavirus. *Nature Reviews Microbiology* (2020) <https://doi.org/10.1038/s41579-020-0332-0>
4. P. Zhou, X. L. Yang, Z. L. Shi, et.al.A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* (2020) <https://doi.org/10.1038/s41586-020-2012-7>.
5. C. Ewen, C. David. Why snakes probably aren't spreading the new China virus. *Nature* (2020), doi: [10.1038/d41586-020-00180-8](https://doi.org/10.1038/d41586-020-00180-8) <https://www.nature.com/articles/d41586-020-00180-8>
6. T. Y. Qiu, T. T. Mao, Y. Wang, M. D. Zhou, Z. W. Cao. Identification of potential cross-protective epitope between 2019-nCoV and SARS virus. *Journal of Genetics and Genomics* (2020), <https://doi.org/10.1016/j.jgg.2020.01.003>
7. R. J. Lu, X. Zhao, J. Li, P. H Niu, B. Yang, H. L. Wu, W. L. Wang, H. Song, B. Y. Huang, N. Zhu, Y. H. Bi, X. J. Ma, F. X. Zhan, L. Wang, T. Hu, H. Zhou, Z. H. Hu, W. M. Zhou, et al., Genomic characterisation and epidemiology of 2019 novel coronavirus implications for virus origins and receptor binding. *The Lancet* (2020), [https://doi.org/10.1016/S0140-6736\(20\)30251-8](https://doi.org/10.1016/S0140-6736(20)30251-8).

8. N. S. Chen, M. Zhou, X. Dong, J. M. Qu, L. Zhang. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *The Lancet* (2020), [https://doi.org/10.1016/S0140-6736\(20\)30211-7](https://doi.org/10.1016/S0140-6736(20)30211-7)
9. C. L. Huang, Y. M. Wang, X. W. Li, L. L. Ren, B. Cao. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet* (2020).
[https://doi.org/10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5)
10. X. Xu, P. Chen, J. Wang, J. Feng, H. Zhou, X. Li, W. Zhong, P. Hao, Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission. *Sci China Life Sci.* **63** (2020). <https://doi.org/10.1007/s11427-020-1637-5>
11. N. John, China's response to a novel coronavirus stands in stark contrast to the 2002 SARS outbreak response. *Nature Medicine* (2020) <https://doi.org/10.1038/s41591-020-0771-1>
12. WHO. Summary of probable SARS cases with onset of illness from 1 November 2002 to 31 July 2003. https://www.who.int/csr/sars/country/table2004_04_21/en/
13. National Health Commission of the People's Republic of China. Press briefing (Feb 04, 2020). <http://www.nhc.gov.cn/xwzb/webcontroller.do?titleSeq=11220&gecstype=1> (accessed)
14. National Health Commission of the People's Republic of China. Press briefing (Feb 05, 2020) <http://www.nhc.gov.cn/xwzb/webcontroller.do?titleSeq=11221&gecstype=1>
15. Z. Xu, S. Li, S. Tian, H. Li, L. Q. Kong. Full spectrum of COVID-19 severity still being depicted. *The Lancet* (2020). [https://doi.org/10.1016/S0140-6736\(20\)30308-1](https://doi.org/10.1016/S0140-6736(20)30308-1)
16. N. Chen, M. Zhou, X. Dong et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *The Lancet* (2020).
17. D. W. Wang, B. Hu, C. Hu, F. F. Zhu, X. Liu, J. Zhang, B. B. Wang, H. Xiang, Z. S. Cheng, Y. Xiong, et al., Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus–Infected Pneumonia in Wuhan, China. *JAMA* (2020).
doi:10.1001/jama.2020.1585

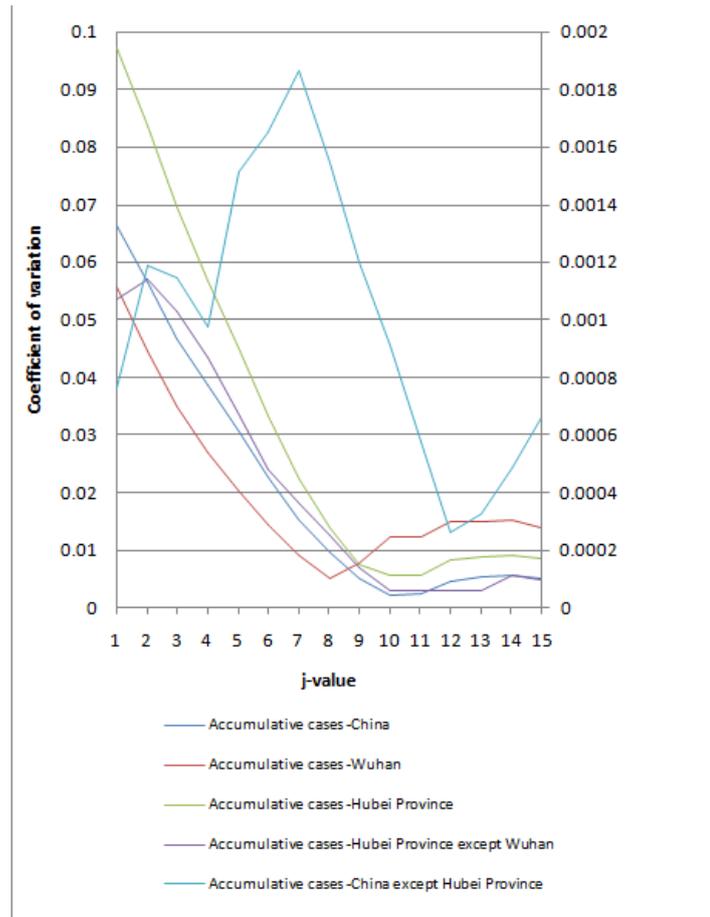


Figure 1 Coefficient of variation of estimated cure rate (p) at different values of j

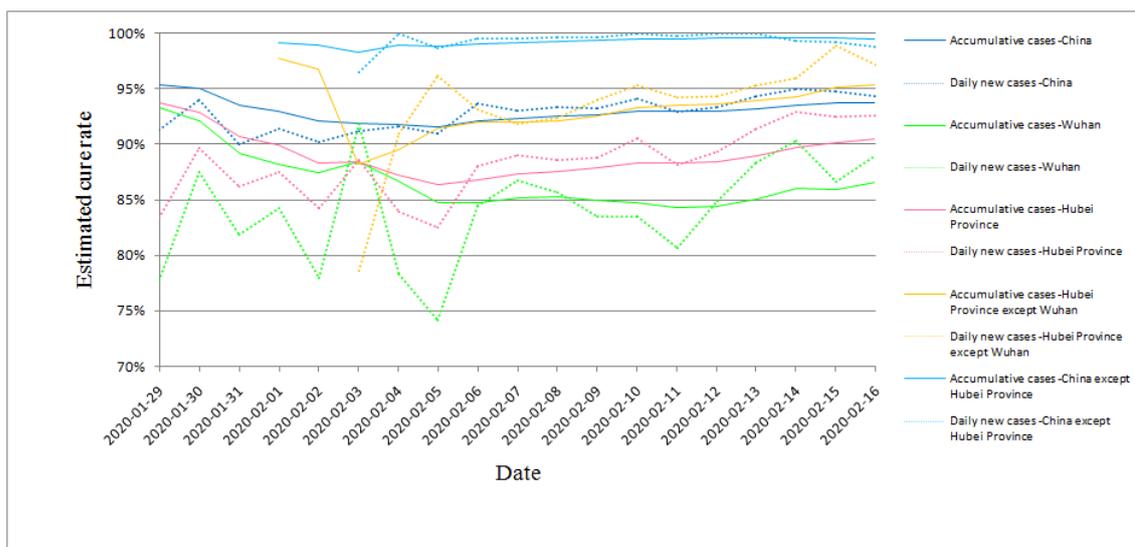


Figure 2 Estimated cure rate at $j = 9$

Table 1 Average of estimated cure rates

j-value	Nationwide		Wuhan		Hubei Province		Hubei Province except Wuhan		Nationwide except Hubei Province	
	Accumulative	Daily new	Accumulative	Daily new	Accumulative	Daily new	Accumulative	Daily new	Accumulative	Daily new
8	91.26%	91.30%	84.23%	81.97%	86.64%	86.20%	90.81%	92.05%	98.98%	99.25%
9	92.98%	92.77%	86.67%	84.06%	89.01%	88.29%	92.62%	93.42%	99.22%	99.32%
10	94.39%	94.09%	88.83%	86.77%	91.04%	90.40%	94.45%	94.75%	99.40%	99.55%