A new method to calculate the instant case fatality rate of COVID-19 in Wuhan and Hubei of China

Lei Cao
Xi’an Jiaotong University

Ting-ting Huang
Xi’an Jiaotong University

Jun-xia Zhang
Xi’an Jiaotong University

Qi Qin
Xi’an Jiaotong University

Si-yu Liu
Xi’an Jiaotong University

Hui-min Xue
Xi’an Jiaotong University

Ya-xin Gong
Xi’an Jiaotong University

Chang-hua Ning
Xi’an Jiaotong University

Xiao-tong Shen
Xi’an Jiaotong University

Jia-xin Yang
Xi’an Jiaotong University

Yan-ni Mi
Xi’an Jiaotong University

Xue Xiao
Xi’an Jiaotong University

Yong-xiao Cao (✉️ yxy@xjtu.edu.cn)
Xi’an Jiaotong University

Research article

Keywords: COVID-19, case fatality rate, cure rate, Wuhan, Hubei

DOI: https://doi.org/10.21203/rs.3.rs-38907/v1
Abstract

The worst-hit area of coronavirus disease 2019 (COVID-19) in China was Wuhan City and its affiliated Hubei Province, where the outbreak has been well controlled. The case fatality rate (CFR) is the most direct indicator to evaluate the hazards of an infectious disease. However, most reported CFR on COVID-19 represent a large deviation from reality. We aimed to establish a more accurate way to estimate the CFR of COVID-19 in Wuhan and Hubei and compare it to the reality. The daily case notification data of COVID-19 from December 8, 2019, to May 1, 2020, in Wuhan and Hubei were collected from the bulletin of the Chinese authorities. The instant CFR of COVID-19 was calculated from the numbers of deaths and the number of cured cases, the two numbers occurred on the same estimated diagnosis dates. The instant CFR of COVID-19 was 1.3%-9.4% in Wuhan and 1.2%-7.4% in Hubei from January 1 to May 1, 2020. It has stabilized at 7.69% in Wuhan and 6.62% in Hubei since early April. The cure rate was between 90.1% and 98.8% and finally stabilized at 92.3% in Wuhan and stabilized at 93.5% in Hubei. The mortality rates were 34.5/100 000 in Wuhan and 7.61/100 000 in Hubei. In conclusion, this approach reveals a way to accurately calculate the CFR, which may provide a basis for the prevention and control of infectious diseases.

Introduction

The novel coronavirus pneumonia (COVID-19) initially appeared in Wuhan, China. As of June 29, 2020, the number of confirmed COVID-19 cases in China has exceeded 85,204 of which 4,648 have died. Wuhan City and its affiliated Hubei Province were the most affected and account for at least 90% of the cases in China. At present, the outbreak in China has been controlled basically. However, the cases of COVID-19 have rapidly increased all over the world and World Health Organization announced it was a pandemic on March 11, 2020. Many countries such as the United States, Spain, Italy, France, Germany and England are most affected. As of June 29, 2020, the total number of patients has risen sharply to over 10.1 million worldwide, and over 499,000 patients died. The global pandemic is getting worse.

Assessing the hazards of infectious diseases is a vital concern in epidemiology. The case fatality rate (CFR), representing the percentage of deaths from a disease among the total number of infected patients during the entire outbreak, is the most direct path that reflects the severity of the disease. The accurate CFR of an infectious disease can be obtained after an outbreak is over on the basis of the numbers of confirmed cases and deaths. Given the severity of the COVID-19 outbreak situation throughout the world, it is important to estimate and predict the CFR of COVID-19. Hard outcomes such as the CFR have a crucial part in forming strategies at national and international levels from a public health perspective. It is imperative that health-care leaders and policy makers are guided by estimates of mortality and case fatality for the setting of public security strategies, the allocation of health resources and the adjustment of medical treatments. This information might also be of great help in the government's decision making and the public's understanding.
However, the mortality or “CFR” of COVID-19 published recently in some articles and announced by authorities are neither mortality nor CFR\(^9,10\). Wang D \textit{et al.} reported that 6 of 138 confirmed cases were dead and 47 of them were discharged, resulting in a CFR of 4.3\(^\text{11}\). Zhang YP analyzed the characteristics of 72314 COVID-19 cases reported in China’ Infectious Disease Information System. There were a total of 1,023 deaths out of 44,672 confirmed cases with a CFR of 2.3\(^\text{12}\). The same problem of the above articles was that the hospitalized patients were ignored. The hospitalized patients would be either cured or dead in future and should not be included in the denominator for calculation of a CFR. In addition, daily confirmed cases would terminate with death or cured a few days later, but on the same reporting day, the deaths were always the confirmed cases some time earlier. In this case, it could increase the error of the estimation of CFR. China is the first country to suffer and recover from COVID-19 outbreak. The accurate estimation of CFR in China may exert a huge impact on global epidemic prevention.

To avoid the above mentioned errors, we established a method to estimate the instant CFR of COVID-19. We collected the daily case notification data of COVID-19 announced by the Chinese officials and estimated death time and cured time. Finally, the instant CFR of COVID-19 was estimated according to the numbers of deaths and cured cases on the same estimated diagnosis dates.

**Methods**

**Research design**

The confirmed COVID-19 patients would eventually have two endpoints, death or cured. Every single patient would undergo a death time or cure time, and then finally reached the end. In the study, the death time or cure time were defined as a period of time from diagnosis date to death date or the time from diagnosis date to cure date, respectively. Firstly, the death time and cure time of COVID-19 patients were calculated by a new method described below in data analysis. Then the estimated date of diagnosis of the declared deaths and cured cases were determined by a death time or a cure time advanced from the deaths or cured cases, respectively. The instant CFR and cure rate of COVID-19 were obtained based on the numbers of deaths and cured cases, which occurred on the same estimated diagnosis dates.

**Data sources**

The daily case notification data of COVID-19 were collected from websites of the National Health Commissions of the People's Republic of China (http://www.nhc.gov.cn/ xcs/xxgzb/dq/index.shtml), the Health Commission of Hubei Province (http://wjw. hubei.gov.cn/fbjd/dtyw), and the Wuhan Municipal Health Commission (http://wjw.wh. gov.cn/front/web/list3rd/yes/802).

**Data collection**

We collected the daily case notification data of COVID-19 in Wuhan City and its affiliated Hubei Province from December 8, 2019, to May 1, 2020\(^\text{13}\). The data included the numbers of daily confirmed
cases/deaths/cured cases and the numbers of daily cumulative confirmed cases/deaths/cured cases. Two researchers collected the data independently every day and then checked and corrected the data. We collected data from the official website, which was considered exempt from approval.

Data analysis

According to the daily numbers of confirmed cases, deaths, and cured cases, the daily variation histograms of confirmed cases, deaths, and cured cases were plotted. Their trend curves of the histograms were fitted, and the dates corresponding to the highest point of the trend curves were found. We speculated that the time difference (period) between the two peak dates of the daily deaths and the daily confirmed cases in trend curves was considered as the mean death time (lag time from diagnosis to death) of the deceased patients. As the same, the time difference between the two peak dates of the daily cured cases and the daily confirmed cases in trend curves was considered as the mean cure time (lag time from diagnosis to recovery) of the cured patients. Depending on the above mean death time and cure time, the estimated initial diagnosis dates of the declared deaths or cured cases of COVID-19 were determined.

The estimated initial diagnosis date of dead patients should be \( n \) days prior to the death date. That is, the estimated diagnosis date (EDD) of the deaths is the date \( n \) days before the declared death date. The estimated initial diagnosis date of deaths is obtained from the date reported on the daily case notification data (DCND) of deaths minus the mean death time (DT). Similarly, the estimated diagnostic date (EDD) of cured cases was obtained from the date reported on the daily case notification data of cured cases minus the mean cure time (CT) (Fig. 1).

\[
\text{EDD of deaths} = \text{date on DCND} - \text{DT}
\]

\[
\text{EDD of cured cases} = \text{date on DCND} - \text{CT}
\]

The daily CFR (FR) is calculated from the number of cumulative deaths (NCD) divided by a sum of the number of cumulative deaths and the number of cumulative cured cases (NCCC) occurred on the same estimated diagnosis date.

\[
\text{FR} = \frac{\text{NCD}}{(\text{NCD} + \text{NCCC})} \times 100\%
\]

The daily cure rate (CR) is calculated from the number of NCCCs divided by the sum of the number of NCDs and NCCCs occurred on the same estimated diagnosis date.

\[
\text{CR} = \frac{\text{NCCC}}{(\text{NCD} + \text{NCCC})} \times 100\%
\]

The mortality rate was calculated based on the number of deaths over the population in Wuhan and Hubei.

Statistical analysis
GraphPad Prism 8 and Excel 2016 were used to record, calculate and analyze the data, and to draw figures of different patterns. The unit of time for the data collection of COVID-19 epidemic was a day. Fit spline/LOWESS was used to analyze and fit the trend curves of the histograms for daily deaths and daily confirmed cases.

Results

Daily new cases

The curves of daily new COVID-19 cases reported in Wuhan City and Hubei Province are shown in Figure 2. The confirmed cases initially occurred on December 8, 2019. The number of daily newly confirmed cases started to increase in early January. The number markedly increased in late January 2020, peaked in mid-February, and then decreased gradually. The death toll began to increase from mid-January to late January, peaked in early February 2020 and then decreased gradually. The COVID-19 patients were initially cured on January 10, 2020. The number of daily new cured patients began to increase in mid-January, obviously increased in early and mid-February, maintained a high level in late February 2020, and then decreased in the first half of March. The numbers of daily new confirmed cases, daily new deaths and daily new cured cases in Wuhan City were higher than those in Hubei Province. The characteristics in Hubei Province and Wuhan City were similar.

Daily cumulative cases

The daily variation curves of cumulative cases of COVID-19, including cumulative confirmed cases, cumulative deaths and cumulative cured cases in Wuhan City and Hubei Province, are shown in Figure 3. The numbers of cumulative confirmed cases began to increase in early January. The increase is sharply accelerated in late January and stabilized until early March. The cumulative death toll started to show an increasing trend in mid-January. The number was remarkably increased in early February, and the numbers increase slowed down in late February. The numbers of cumulative cured cases showed an increasing tendency in mid-January. The increase was visibly accelerated in February and remained stable until April. On April 16, the data of confirmed cases, deaths and cured cases of COVID-19 in Wuhan were adjusted by Chinese authorities. The number of confirmed cases increased from 50,008 to 50,333. The number of deaths increased from 2,579 to 3,869 and the number of cure cases decreased from 47,283 to 46,335. The cumulative cases before April 16 were altered based on the increase or decrease proportion.

Death time and cure time

After treatment, the COVID-19 patients died or were cured. The time from diagnosis to death is referred to as the death time. The time from diagnosis to discharge is the cure time. Figure 4 shows daily variation histograms of confirmed cases, deaths and cured cases in Wuhan and Hubei. The trend curves of the histograms were fitted, and the dates that corresponded to the highest point of the trend curves were found. The dates of highest point in the trend curves of daily confirmed cases, the deaths, and the cured
cases were February 9, February 15, and February 29, respectively, in Wuhan City; and February 8, February 14, and February 28, respectively, in Hubei Province. The death time and the cure time were 6 days and 20 days in Wuhan City and Hubei Province, respectively.

**Estimation of diagnosis dates of the deaths and cured cases**

The estimated diagnosis dates of deaths and cured cases were determined based on mean death or cure times, respectively. We assumed that the dead patients died on day 6 after diagnosis was confirmed in Wuhan and Hubei. The estimated diagnosis date of the dead patients should be on 6th day prior to death date. The cured patients discharged on day 20 after diagnosis. The estimated diagnosis date of the cured cases should be 20 days before the announced cure date in Wuhan and Hubei, respectively. Figure 5 shows the curves of the numbers of cumulative deaths and the number of cumulative cured cases of COVID-19 on the estimated diagnosis dates. The number of confirmed cases was equal to the sum of the number of cumulative deaths and the number of cumulative cured cases on the same estimated diagnosis date. Since the number of COVID-19 deaths on the same day was significantly less than the number of cured cases, the curve of estimated confirmed COVID-19 cases which was the sum of the cumulative cured cases and the cumulative death toll, was close to the curve of cumulative cured cases. The shapes and trends of the curves in Wuhan City are similar to but lower than those in Hubei Province.

**Instant CFRs**

The instant CFR of COVID-19 in Wuhan City was less than 10% (1.3-9.4% with an average of 7.5% from January 1 to April 9, 2020. It was low in the first half of January, gradually increased and exceeded 9.0% on the end of January, then gradually stabilized at around 8.0% in late February. It was no more than 8% in March and has been stable at 7.69% since early April (Figure 6A). The instant CFR of COVID-19 in Hubei Province showed the same trend as that in Wuhan City, but the CFRs were lower than that in Wuhan City. The instant CFR of COVID-19 in Hubei Province was less than 7.4% (1.2-7.4%) with an average of 6.2% from January 1 to April 10, 2020. It was lower than 4% in the first half of January, then increased to about 7.0% in late January, and has been stable at 6.62% since early April (Figure 6B). The curve tendency showed that the CFRs were stable after April, 2020.

**Instant cure rates**

The cure rate and CFR are opposing figures, since cure and death are a pair of competing events. The instant cure rate of COVID-19 in Wuhan City was between 90.1% and 98.8%, and has been stable at 92.3% since early April. While the cure rate in Hubei Province was between 92.6% and 98.7%. Overall, the cure rates didn't fluctuate much, and finally stabilized at 92.3% in Wuhan City and at 93.5% in Hubei Province on April 9 (Figure 6).

**The mortality rate**
The numbers of deaths in Wuhan and Hubei was 3,869 and 4,512 on April 30, 2020. The population of Wuhan City and Hubei Province are 11.21 and 59.27 million. So, the mortality rates were 34.5/100,000 in Wuhan City and 7.61/100,000 in Hubei Province.

Discussion

In view of the misleading of reported CFR calculation, we have, for the first time, established a method to calculate the instant CFR of COVID-19. This method resolved the problems of COVID-19 patients who are still hospitalized and the time lag between the diagnosis and outcome. By using this new calculation, we estimated the CFRs of COVID-19 in Wuhan City and its Hubei Province, where the epidemic was most severe inside China. The results showed that the instant CFR of COVID-19 was 1.3%-9.4% in Wuhan and 1.2%-7.4% in Hubei from January 1 to May 1, 2020. The CFR has stabilized at 7.69% in Wuhan and 6.62% in Hubei since early April, 2020.

The accurate CFR of an infectious disease can be obtained after the end of the outbreak. At present, the outbreak of COVID-19 is ongoing throughout the world. The numbers of confirmed cases, deaths and cured cases are constantly changing. Therefore, estimating the CFR is obviously challenging. However, it is critical to balance the socioeconomic burden of infection control interventions against their potential benefit for mankind. The current literature and the officially published CFRs have a large deviation.

Huang CL et al reported the clinical features of 41 patients infected with SARS-CoV-2 in Wuhan, China. As of January 22, 2020, 6 of 41 patients died, and 28 had been discharged (CFR 15%). Wang D et al reported the clinical characteristics of 138 hospitalized patients with COVID-2019 in Wuhan, China. Six patients died, and 47 were discharged. The remaining 85 patients were still hospitalized. The CFR was reported as 4.3%. Zhang YP extracted a total of 72,314 cases with COVID-19 reported China's Infectious Disease Information System through February 11, 2020. A total of 1,023 deaths occurred among 44,672 confirmed cases, for an overall CFR of 2.3%. However, he ignored the fact that there were 38,909 hospitalized cases in confirmed cases. At present, on official websites, the CFRs have been calculated by dividing the number of known deaths by the number of confirmed cases. These data reflect neither the CFR nor mortality and might be off by orders of magnitude. The confirmed cases included deaths, cured patients, and hospitalized patients. Some of the hospitalized patients may die in the future, and their data would not contribute to the calculation of case fatality. In addition, the diagnosis of viral infection will precede death or recovery by days to weeks. The number of deaths should be compared to the previously confirmed case numbers-accounting for this delay increasing the estimated CFR.

We assumed that the CFR of the whole outbreak process could be regarded as a collection of many successive instant CFRs. The instant CFR is directly related to various factors at a time, and it is easier to analyze various factors and take possible actions to influence the disease progress by knowing this CFR in advance. This instant CFR will gradually approach the CFR as time goes on until the outbreak ends gradually. The extreme instant CFR can be used as an approximate value of the CFR within a certain
period of time. This approach provides a way to accurately calculate the CFR without involving the data of hospitalized patients.

The instant unit of the CFR was set as a day and the daily CFRs were calculated. The daily CFR was equal to the daily cumulative number of deaths being divided by the sum of the number of cumulative deaths and the number of cumulative cured patients. The number of daily confirmed cases is equal to the number of deaths plus the number of cured cases. The confirmed COVID-19 patients had to go through a certain period of time (the death time or the cure time), resulting in the final outcomes: death or cure. The date of death minus the death time is the diagnosis date of death. Similarly, the cure date minus the cure time is the diagnosis date of cured patients. The daily CFR and cure rates were then calculated using the cumulative number of deaths and cured cases at the same diagnosis date. This rate varies from day to day with disease-related factors.

The outbreak of COVID-19 first appeared in Wuhan, where the situation was most serious with the highest CFR of China. The CFR varied over time. In the first half of January, the CFR was less than 10% because the outbreak was not very serious and prevalent. It was easy for patients to see a doctor. However, until the beginning of February, the number of patients largely increased, and it was difficult for patients to receive medical care treatment. The admission and diagnosis of COVID-19 were delayed, the hospitals were crowded out, and the medical condition deteriorated. As a result, the CFR rose sharply to 19.5%. In February, a large number of domestic medical resources supported Wuhan, and the medical conditions gradually improved. The treatment concept and methods were updated, the level of medical care was improved, and the treatment procedure became increasingly standardized. On the other hand, the virulence of SARS-CoV-2 may decline as the virus passes through generations. The proportion of severe cases decreased, and the CFR decreased to nearly 5.26% on April 1. We speculated that the CFR will be close to the final CFR. The trend in Hubei Province was similar to that in Wuhan, but the CFR was low. The CFRs in the first half of January 2020 were less than 10%. From the second half of January to the beginning of February, the rate rose to 10% ~16.6% and then decreased to 4.82% on April 1. We used this method to estimate the CFR (5.7%) of COVID-19 in China; this estimated CFR is close to its actual value (5.5%), supporting that this method is reliable. Using this method, we calculated the COVID-19 CFR of China and it turned out to be higher than the reported CFR of many articles and authorities, meaning that the CFR of COVID-19 in Wuhan City and Hubei Province were also underestimated.

In comparison with the outbreak of severe acute respiratory syndrome (SARS) in 2003, of 5,327 probable cases with SARS, 343 cases died in mainland China, giving an overall CFR of 6.4%\(^\text{16}\). In some cities or provinces, such as Beijing, Tianjin, Guangdong, Shanxi and Inner Mongolia, the CFR of SARS were 7.3%, 12.0%, 5.4%, 9.9% and 8.1%, respectively. In the present study, the CFRs were 7.69% and 6.62% in Wuhan City and Hubei Province, respectively. Therefore, the CFRs of COVID-19 in Wuhan, and Hubei are almost at the same level as that of SARS in mainland China.

This study has some limitations. The major influencing factors were the determination of mean death time and mean cure time. Indeed, not every patient followed this mean death time or cure time. The result
may deviate from the actual situation. In addition, our study was based on the official data. We believe that and mortality rate would be high and the CFR would be low if there’s some missing cases, false negative diagnosis of nucleic acid, or some asymptomatic infections were included. We suggest that the government disclose cases of infectious diseases to share resources for scientific study. If the accurate cure time and death time is obtained, it may help to correctly calculate the CFR and provide a basis for the prevention and control of infectious diseases.

Declarations

Availability of data and materials

The datasets during and/or analyzed during the current study available from the corresponding author on reasonable request.

Conflict of interest

The authors declare that there are no conflicts of interest.

Source of funding

The study was supported by the National Natural Science Foundation of China (grant no. 81670001)

Contributions

YXC and LC were responsible for the conceptual design and led the team. TTH and JXZ performed the modeling. QQ, SYL, HMX, YXG, CHN, XTS, and JXY were responsible for data collection and management. LC and YXC drafted the manuscript. YNM and XX revised the manuscript. All authors read the manuscript and contributed to editing.

References

1. Lu HZ, Stratton CW, Tang YW. Outbreak of pneumonia of unknown etiology in Wuhan, China: The mystery and the miracle. Journal of Medical Virology. Apr 2020;92(4):401–402.
2. Wang C, Horby PW, Hayden FG, Gao GF. A novel coronavirus outbreak of global health concern (vol 395, pg 470, 2020). Lancet. Feb 15 2020;395(10223):496–496.
3. Chan JFW, Kok KH, Zhu Z, et al. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. Emerging Microbes & Infections. Jan 1 2020;9(1):221–236.
4. N Z, D Z, W W, al e. A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med.
5. Ebrahim SH, Ahmed QA, Gozzer E, Schlagenhauf P, Memish ZA. Covid-19 and community mitigation strategies in a pandemic. BMJ. Mar 17 2020;368:m1066.
6. WHO. Coronavirus disease 2019 (COVID-19) Situation Report-40. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200229-sitrep-40-covid-19.pdf?sfvrsn=849d0665_2.

7. Chen Z, Akazawa K, Nakamura T. Estimating the CFR using a constant cure-death hazard ratio. *Lifetime Data Anal.* Sep 2009;15(3):316–329.

8. Rajgor DD, Lee MH, Archuleta S, Bagdasarian N, Quek SC. The many estimates of the COVID-19 CFR. *Lancet Infect Dis.* Mar 27 2020.

9. Mahase E Coronavirus covid-19 has killed more people than SARS and MERS combined, despite lower CFR. *BMJ.* Feb 18 2020;368:m641.

10. Battegay M, Kuehl R, Tschudin-Sutter S, Hirsch HH, Widmer AF, Neher RA. 2019-novel Coronavirus (2019-nCoV): estimating the CFR - a word of caution. *Swiss Med Wkly.* Jan 27 2020;150:w20203.

11. Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet.* Feb 15 2020;395(10223):497–506.

12. Zhang Y. The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China. *Chin J Epidemiol.* 2020;41(2):145–151.

13. Chen H, Wang T, Zhang Y. Comparison for four case fatality ratio estimation methods on the epidemic situations of infectious diseases. *Sci Technol Eng.* 2017;17(24):152–156.

14. Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *Lancet.* Feb 29 2020;395(10225):689–697.

15. Wang D, Hu B, Hu C, et al. Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. *JAMA.* Feb 7 2020.

16. Jia N, Feng D, Fang LQ, et al. Case fatality of SARS in mainland China and associated risk factors. *Trop Med Int Health.* Nov 2009;14 Suppl 1:21–27.

Figures
Figure 1

The relationship between estimated diagnosis date (EDD) and dates of deaths or cured cases on daily case notification data (DCND).
Figure 2

The daily variation of new confirmed cases, new cured cases and new deaths of COVID-19 in Wuhan City (A) and Hubei Province (B).
Figure 3

The daily variation of cumulative confirmed cases, cumulative cured cases and cumulative deaths of COVID-19 in Wuhan City and Hubei Province.

Figure 4

The numbers of daily new confirmed (A, B), deaths (C, D) and cured cases (E, F) of COVID-19 in Wuhan City (A, C, E) and Hubei Province (B, D, F). The red arrow indicated the highest point of the trend curve in the horizontal axis, which corresponded a date in the vertical axis.
Figure 5

The number of cumulative deaths, the number of cumulative cured cases and the number of deaths + cure cases of COVID-19 occurred on the same estimated diagnosis dates in Wuhan City (A) and Hubei Province (B) based on daily case notification data.
Figure 6

The instant fatality rates and cure rates of COVID-19 in Wuhan City (A) and Hubei Province (B) based on daily case notification data.