A Novel Approach to Calculate the Case Fatality Rate of COVID-19 in European countries

Qi Qin
Xi’an Jiaotong University

Yong-xiao Cao
Xi’an Jiaotong University

Lei Cao (leicao@mail.xjtu.edu.cn)
Xi’an Jiaotong University

Research article

Keywords: COVID-19, case fatality rates, Italy, Spain, France, Germany

Posted Date: July 17th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-42495/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

As a result of the global epidemic of coronavirus disease 2019 (COVID-19), many European countries and regions have been strongly affected. The case fatality rate (CFR) is the most important index to evaluate the hazards associated with an epidemic situation. We aimed to present a method to calculate the instant CFR and to evaluate and compare the instant CFR of COVID-19 in the four European countries that have been most impacted: Italy, Spain, France and Germany. The daily COVID-19 case data from January 30, 2020, to July 9, 2020 in Italy, Spain, France and Germany were collected from WHO reports. Death time was calculated as the difference between the peak dates of the number of daily confirmed cases and the number of deaths in each of the four countries. The estimated dates of diagnosis of the declared deaths were a death time prior to the dates of death. The instant fatality rate of COVID-19 was calculated as the ratio of the number of cumulative deaths to the number of cumulative confirmed cases; these deaths and confirmed cases occurred on the same estimated dates of diagnosis. As of July 9, 2020, the COVID-19 death time was 6, 4, 6 and 12 days in Italy, Spain, France and Germany, respectively. The instant CFR of COVID-19 was 14.4%~27.6%, 2.2%~14.7%, 8.2%~25.0% and 2.0%~10.5% in Italy, Spain, France and Germany, respectively. The average CFR of COVID-19 was highest in France (16.7%) and lowest in Germany (5.0%). Since late April 2020, the CFR has stabilized at approximately 15%, 20% and 5% in Italy, France and Germany, respectively. Since early June, 2020, the CFR in Spain has stabilized at approximately 11%. We have established a more accurate way to calculate the CFR that may provide a basis for the prevention and control of infectious diseases.

Introduction

Presently the outbreak of coronavirus disease 19 (COVID-19) in China has been well controlled. Unfortunately, the global epidemic has spread to 212 countries and territories. The number of confirmed COVID-19 cases is sharply increasing every day worldwide and the real-time reports can be tracked on websites provided by international forums such as Johns Hopkins University. At the end of January 2020, the WHO has declared that the COVID-19 became a Public Health Emergency of International Concern (PHEIC). As of July 9, 2020, over 12.2 million COVID-19 cases and over 550,000 deaths were confirmed worldwide. Many European countries also had early epidemics after China, but the current situation of some countries is becoming even worse than it was in China. From February 2020, Italy became the first country to face COVID-19. Other European countries, such as Spain, the United Kingdom, France, Germany, Russia and the Netherlands, were also caught in the epidemic.

It is important to assess the hazards of an infectious disease since it is a vital concern in epidemiology. The case fatality rate (CFR) is one of most important indexes to reflect the hazard of the infectious disease. It is of great significance to estimate and predict the fatality rate of COVID-19 in different countries during the epidemic. It can help us analyze the pros and cons of different countries' responses to the epidemic. It can also provide an important reference to decide on efficient strategies to face the huge public health challenge posed by the epidemic.
Since the COVID-19 first appeared in December 2019, many scientists and researchers have studied this disease from different scientific perspectives at different levels, some of which have been related to the CFR. However, we recently found that many studies calculated the CFR directly as the ratio of the number of cumulative deaths to the number of cumulative confirmed cases, without any correction. Giangreco and Stafford reported the CFR of COVID-19 in Italy and Germany, respectively, where the CFR was obtained by dividing the number of cumulative deaths by the number of cumulative confirmed cases. However, this crude calculation of the CFR could be misleading. When the number of total deaths is divided by the number of total confirmed cases, the calculation ignores that the date of the deaths and the date of initial diagnosis were not on the same day. There was a period of time for the patients underwent the whole COVID-19 disease course, including developing symptoms, being diagnosed and reported, and the final clinical outcome (death or recovery). There was a time lag between the diagnosis and outcome. The date of death should be a few days later than the date of diagnosis. As the denominator, the total number of confirmed cases contains some cases that would die a few days later, and these deaths should have added to the total number of deaths as the dividend. During a growing epidemic, the final clinical outcome of most of the reported confirmed cases is unknown. Simply dividing the cumulative reported number of deaths by the cumulative number of reported confirmed cases will therefore underestimate the CFR, especially in the early stage of an epidemic.

To avoid the abovementioned biases, we established a method to estimate the instant CFR of COVID-19 at different stages. In the present study, the method was expected to resolve the problems of COVID-19 patients with unknown outcomes and the time lag between the diagnosis and outcome. The analysis and comparison of the COVID-19 CFRs in different European countries during the epidemic has a significant role in providing a basis for the prevention and control of COVID-19 in Europe.

Methods

Data sources and collection

The daily data of newly confirmed cases/deaths/cured cases and the daily data of cumulative confirmed cases/deaths/cured cases of COVID-19 from Jan 30, 2020 to July 9, 2020 in Italy, Spain, France and Germany, were respectively collected from websites of the WHO (https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports), Johns Hopkins University (https://coronavirus.jhu.edu/map.html) and worldometer (https://www.worldometers.info/coronavirus) websites. 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.

Research design

The confirmed COVID-19 patients would eventually have one of two outcomes, death or recovery. After diagnosis, the patients would first experience a death time or a recovery time and then reach the end of
disease, death or recovery. In the present study, death time was defined as the period of time from the diagnosis date to the death date. Similarly, recovery time was defined as the period of time from the diagnosis date to the recovery date. The death time was calculated as the difference between the peak dates of the number of deaths and the number of confirmed cases in the four different countries. The estimated dates of diagnosis of the declared deaths cases were one death time earlier than the dates of death. The instant fatality rate for COVID-19 was calculated as the ratio of the number of cumulative deaths to the number of cumulative confirmed cases, and these deaths and confirmed cases occurred at the same estimated dates of diagnosis.

**Data analysis**

The daily variation histograms of confirmed cases and deaths in the four different countries were plotted. Their trend curves of these histograms were fitted, and the dates corresponding to the highest point of the trend curves were found. The difference in the peak dates between the number of deaths and the number of confirmed cases on the trend curves was considered as the mean lag time from diagnosis to death, namely, death time. On the basis of the above lag times (death time), the initial diagnosis dates of the declared COVID-19 death from daily reports were calculated.

The estimated diagnosis date of declared deaths were \( n \) days prior to the death date. That is, the estimated diagnosis date (EDD) of the deaths was the date \( n \) days before the reported death date. The estimated diagnosis date (EDD) of deaths was obtained from the death date reported on the daily notification data (DND) minus the death time (DT). (Fig. 1).

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

The daily CFR of the four countries was calculated from the number of cumulative deaths (NCD) divided by the number of cumulative confirmed cases (CCC).

\[
\text{CFR} = \frac{\text{NCD}}{\text{CCC}} \times 100\%
\]

Since recovery and death are a pair of competing events, the recovery rate and CFR are opposing figures. The daily recovery rates (RR) were calculated from the following equation:

\[
\text{RR} = 1-\text{CFR}
\]

The instant incidence rates were calculated on the basis of the number of cumulative confirmed cases in the population, and the instant mortality rates were calculated on the basis of the number of cumulative deaths in the population in each of the four different countries, respectively.

**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 daily newly confirmed cases, deaths and recovery cases in Italy, Spain, France and Germany are plotted in Fig. 2. In Italy, the first COVID-19 patient was reported on January 30, 2020. The number of new confirmed cases increased rapidly from March 16 to a peak of 6,557 on March 21. The first two deaths occurred on February 22. Then, the death toll rose sharply in mid-to-late of March, with a peak of number of deaths (969 deaths) on March 27. Italy's first recovery patient appeared on February 22, then the daily new recovery cases increased gradually, but the peak had not yet appeared. In Spain, the first COVID-19 patient was reported on January 31, 2020. The number of new confirmed cases increased rapidly in mid-to-late of March and reached a peak at the end of March. The first deaths occurred on March 4, and the peak of the number of deaths was at the beginning of April. Spain's first recovery patient appeared on March 5, and the number of recovery cases was still in increase. The COVID-19 epidemic started on January 24, 2020 in France. The number of new confirmed cases increased quickly in mid-to-late of March, and the peak was at the beginning of April. The first deaths occurred on March 2, with a peak number of deaths in early April. France's first cured patient appeared on March 10, with a peak number of cured cases in mid-April. The report of COVID-19 cases in Germany began on January 27, 2020. The number of new confirmed cases increased rapidly in mid-to-late of March, and the number reached a peak at the end of March and early April. The first deaths occurred on March 9, and the peak number of deaths was reached in the mid-April. Germany's first recovery patient was reported on March 10, with a peak number of cured patients in mid-April. On April 8, the number of recovery cases reached 37,604 in Germany, which is much higher than that of Italy, Spain and France.

Incidence rates and mortality

The incidence rates and mortality rate of COVID-19 in Italy, Spain, France and Germany are shown in Fig. 3. From late January to early March, the incidence rates among these countries grew slowly, and then they accelerated from early March. Italy had the highest incidence rate among the four countries before late March, and Spain overtook Italy in late March to have the highest incidence rate of 523/1,000,000. The incidence curves in France and Germany were similar, with incidence rates of 228/1,000,000 and 225/1,000,000, respectively. The mortality rate increased slowly from late January to early March, and then began to accelerate. The mortality rate curves in Italy and Spain are similar to the growth track of the incidence curve. In late March, Spain became the country with the highest mortality rate of 64/1,000,000. France's mortality has been much higher than that of Germany, which had the slowest growth rate and the lowest mortality rate of 10.6/1,000,000. The mortality rates of Italy and France were 57/1,000,000 and 44/1,000,000, respectively.
The lag time

There was a lag time from diagnosis date to death date, namely, death time. Figure 4 shows daily variation histograms of confirmed cases, and deaths. 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 the highest point in the trend curves of daily confirmed cases, and the deaths were March 24, and March 30, 2020, respectively, in Italy; 31 March and April 4, 2020, respectively, in Spain; April 3 and April 9, 2020, respectively, in France; and April 3 and April 15, 2020, respectively, in Germany. The lag times from diagnosis to death (death time) were 6 days, 4 days, 6 days, and 12 days in Italy, Spain, France, and Germany, respectively.

Estimation of diagnosis dates of the deaths

The confirmed COVID-19 patients had one of two outcomes: death or recovery. From the above estimation, we assumed that the average death time of COVID-19 in Italy was 6 days, that is, the initial diagnosis date of the declared deceased patients should be an average of 6 days prior to the death date. In the same way, the diagnosis dates of declared deaths were estimated to be 4 days, 6 days and 12 days earlier prior to the death date in Spain, France and Germany, respectively. Figure 5 shows the curves of the numbers of cumulative deaths and the number of cumulative confirmed cases of COVID-19 on the same estimated diagnosis dates. The deaths were not diagnosed on the same date as confirmed cases. Instead, the deaths were initially diagnosed as confirmed cases a period of time (death time) ago. Therefore, the number of daily cumulative confirmed cases should be the cases diagnosed on the same date when the deaths were diagnosed as confirmed cases.

Instant CFRs

Figure 6 shows that instant CFRs and recovery rates of COVID-19 in the four countries. In Italy, the instant CFR ranged from 14.4% to 27.6%, and the average CFR was 16.0%. In late April, 2020, the CFR stabilized at approximately 15.0%. In Spain, the highest instant fatality rate was 14.7%, the lowest fatality rate was 2.2%, and the average CFR was 11.5%. In early June, 2020, the CFR stabilized at approximately 11.0%. In France, the highest CFR was 25.0%, the lowest CFR was 8.2%, and the average fatality rate was 16.7%. In late April, 2020, the CFR stabilized at approximately 20.0%. In Germany, the highest instant fatality rate was 10.5%, the lowest CFR was 2.0%, and the average CFR was 5.0%. As of late April, 2020, the CFR stabilized at approximately 5.0%.

Instant recovery rates

The trends in the recovery rate and fatality rate are opposite because they are a pair of competitive events. In Italy, the highest instant recovery rate was 85.6%, the lowest recovery rate was 72.4%, and the average recovery rate was 84.0%. In late April, 2020, the recovery rate has stabilized at approximately 85.0%. In Spain, the highest instant recovery rate was 97.8%, the lowest cure rate was 85.3%, and the
average recovery rate was 88.5%. In early June, 2020, the recovery rate stabilized at approximately 89.0%. In France, the highest instant cure rate was 91.8%, the lowest cure rate was 75.0%, and the average recovery rate was 83.3%. In late April, 2020, the recovery rate stabilized at approximately 80.0%. In Germany, the highest instant recovery rate was 98.0%, the lowest recovery rate was 89.5%, and the average recovery rate was 95.0%. In late April, 2020, the recovery rate stabilized at approximately 95.0% (Fig. 6).

**Discussion**

In view of the misunderstanding of the CFR calculation in many published articles, 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 with uncertain outcomes and the time lag between the diagnosis and outcome. By using this new calculation, we determined and compared the CFRs of COVID-19 in Italy, Spain, France and Germany, which are the European countries where the epidemic was most severe. The results showed that the average CFR of COVID-19 was highest in France and lowest in Germany.

At present, the global epidemic of COVID-19 is still ongoing. Assessing the hazards of infectious diseases is a vital concern in epidemiology\textsuperscript{13}. Generally, an accurate CFR of an infectious disease can be obtained after the epidemic is over on the basis of the number of total deaths and the number of total confirmed cases. However, the number of confirmed cases, deaths and cured cases are constantly changing\textsuperscript{14} during the epidemic.

Recently, some scientists reported the “CFR” of COVID-19 in their publications. However, we found that this crude CFR\textsuperscript{9,10} was mostly obtained by dividing the number of cumulative deaths by the number of cumulative confirmed cases, which would underestimate the real CFR of COVID-19. For example, Zhang extracted a total of 72,314 COVID-19 cases reported in China's Infectious Disease Information System\textsuperscript{15}. There were a total of 1,023 deaths out of 44,672 confirmed cases for an overall CFR of 2.3%. While this may seem reasonable on the surface, the fact that there were still 38,909 hospitalized cases among the confirmed cases was ignored. The outcomes of these hospitalized patients remained uncertain. Some patients died a few days later. Therefore, this number of future deaths should be added to the dividend when calculating the CFR. In addition, Sun et al. reported that the CFR was 2.1% (361/17,205)\textsuperscript{16} in China, which also misled the public as it was estimated by using the same calculation as Zhang et al. These results may represent a large deviation from reality. Interestingly, Yang et al. used a simple linear regression model to estimate the CFR of COVID-19 in mainland China\textsuperscript{17}; the reported CFR was 2.10% (95% CI: 2.05–2.14%). However, at present, the outbreak has been basically controlled in China. Therefore, the CFR can be obtained by dividing the total deaths (4,644) by the total confirmed cases (84,451); the current estimated CFR in China is approximately 5.5%, which is obviously higher than expected in all above papers. Therefore, estimating the CFR is challenging\textsuperscript{18}.

We assumed the CFR during the epidemic process to be a collection of many successive instant CFRRs. We set the instant unit of the CFR as one day and calculated the daily CFR. The daily CFR was equal to
the number of daily cumulative deaths being divided by the number of daily cumulative confirmed cases. The difference from the above calculation is that the two numbers in the equation (deaths and confirmed cases) occurred on the same estimated diagnosis date. The deaths (dividends) were not diagnosed as confirmed cases (divisors) on the calculation date. Instead, they were diagnosed as confirmed cases a period of time (death time) ago. Therefore, in the CFR equation, when the dividend was the number of daily cumulative deaths, the divisor should be the number of daily cumulative confirmed cases diagnosed on the same date when the deaths (in the dividend) were diagnosed as confirmed cases. We used this method to estimate the CFR (5.7%) of COVID-19 in China; the estimated CFR is close to its actual value (5.5%), supporting that this method is reliable.

COVID-19 is a topic of general concern to all humankind. Due to the differences in attention, medical facilities, public consciousness of self-protective measures, and economics, there are some differences in the epidemic situation in different countries. The number of confirmed cases/deaths/recovery cases is changing every day, so it is difficult to use a simple number or average number to show the mortality rate and CFR. Hence, we used the instant CFR approach to estimate the impact of the epidemic and changes in the hazards associated with COVID-19 in different European countries.

According to the different levels of the CFR in different places, we are able to determine the status of the COVID-19 situation in different countries. The CFR could be an important reference and indicator for real-time measurement of epidemic inside a country and even for comparison among different countries. Germany has shown the lowest CFR of COVID-19, demonstrating that Germany has provided the best medical facilities for COVID-19 among these four European countries. France has the highest average CFR (16.7%) and it has surpassed that in Italy (16.0%) after June, indicating that the French needs to take more stringent measures to control COVID-19. Meanwhile, France has maintained the highest instant CFR after stabilization (20.0%). Since late April, 2020, the instant CFR has stabilized at approximately 15% in Italy, and 20% in France, respectively, which means that the French and Italian people and government should pay more attention to the epidemic right now.

During the outbreak of severe acute respiratory syndrome (SARS) in 2003, a total of 8,422 SARS cases were reported in 32 countries and regions. The global death toll from SARS was 919, with a CFR of nearly 11% for SARS worldwide. However, the current total number of COVID-19 cases and deaths for now is far more than the respective reported numbers for SARS. The SARS epidemic was most severe in China (Beijing, Hong Kong) and some countries in Southeast Asia, while SARS epidemic in Europe was relatively mild, and SARS had no widespread impact in European countries. The number of SARS cases in Italy, Spain, France and Germany was 4, 1, 7, 9, respectively, out of which there was only 1 death (from France). As a comparison, as for June 18, 2020, the instant (average) CFR of COVID-19 was 14.4%~27.6% (16.0%), 2.2%~14.7% (11.0%), 8.2%~25.0% (16.7%) and 2.0%~10.5 (5.0%) in Italy, Spain, France and Germany, respectively. Obviously, the current status of the COVID-19 epidemic in Europe is much more serious than it was for SARS in 2003. We have also determined that the incidence rates and mortality rates of COVID-19 for the four European countries are significantly higher than those for SARS in 2003. It seems that even though COVID-19 was first seen in Wuhan, China, the situation of China is currently well
controlled. However, the epidemic has had widespread transmission worldwide. Why is the epidemic of COVID-19 worse than that of SARS in Europe? The transmissibility of the SARS virus is not strong, so it has been limited to only some places. Since SARS had a strong lethality, it still showed a high CFR in the limited endemic areas. On the other hand, COVID-19 is highly contagious, and it has a stronger ability to spread among people, so the whole world has been affected.

The number of confirmed cases and the number of deaths in Europe are still rising. Our study developed a new method of estimating the CFR of COVID-19. Then, we compared the CFRs of COVID-19 in four European countries and found that, as for July 9, 2020, the average CFR of COVID-19 was highest in France and lowest in Germany. Since late April 2020, the CFR has stabilized at approximately 15%, 20% and 5% in Italy, France and Germany, respectively. Since early June, 2020, the CFR in Spain has stabilized at approximately 11%. This more accurate way is established to calculate the CFR may provide a basis for the prevention and control of infectious diseases.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

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

Competing interests

The authors declare that they have no competing interests.

Funding

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

Authors' contributions
YXC and LC were responsible for the conceptual design and led the team. QQ performed the modeling, responsible for data collection and management. QQ, LC and YXC drafted the manuscript, revised the manuscript. All authors read the manuscript and contributed to editing.

Acknowledgements

Not applicable.

References


**Figures**

![Figure 1](image)

**Figure 1**

The relationship of estimated diagnosis date (EDD) and dates of deaths.
Figure 2

The daily variation of newly confirmed COVID-19 cases, deaths and recovery cases in Italy (A), Spain (B), France (C), and Germany (D).

Figure 3

The incidence rates (A) and mortality rates (B) of COVID-19 in Italy, Spain, France, and Germany.
Figure 4

Daily variation histograms and fitted trend curves of confirmed cases (A, C, F, G) and deaths (B, D, F, H) in Italy (A, B), Spain (C, D), France (E, F), and Germany (G, H).
Figure 5

The numbers of cumulative deaths and the number of cumulative confirmed cases of COVID-19 on the estimated diagnosis dates in Italy (A), Spain (B), France (C), and Germany (D).
Figure 6

Instant fatality rates and recovery rates of COVID-19 in Italy (A), Spain (B), France (C), and Germany (D).