

The computation of case fatality rate for novel coronavirus (COVID-19) based on Bayes theorem

An observational study

Chi-Sheng Chang, MBA^{a,b}, Yu-Tsen Yeh, MS^c, Tsair-Wei Chien, MBA^d, Jui-Chung John Lin, DC^e, Bor-Wen Cheng, PhD^b, Shu-Chun Kuo, MD^{f,g,*}

Abstract

Background: When a new disease such starts to spread, the commonly asked questions are how deadly is it? and how many people are likely to die of this outbreak? The World Health Organization (WHO) announced in a press conference on January 29, 2020 that the death rate of COVID-19 was 2% on the case fatality rate (CFR). It was underestimated assuming no lag days from symptom onset to deaths while many CFR formulas have been proposed, the estimation on Bayes theorem is worthy of interpretation. Hence, it is hypothesized that the over-loaded burdens of treating patients and capacities to contain the outbreak (LSBHRS) may increase the CFR.

Methods: We downloaded COVID-19 outbreak numbers from January 21 to February 14, 2020, in countries/regions on a daily basis from Github that contains information on confirmed cases in >30 Chinese locations and other countries/regions. The pros and cons were compared among the 5 formula of CFR, including [A] deaths/confirmed; [B] deaths/(deaths + recovered); [C] deaths/(cases x days ago); [D] Bayes estimation based on [A] and the outbreak (LSBHRS) in each country/region; and [E] Bayes estimation based on [C] deaths/(cases x days ago). The coefficients of variance (CV=the ratio of the standard deviation to the mean) were applied to measure the relative variability for each CFR. A dashboard was developed for daily display of the CFR across each region.

Results: The Bayes based on (A)[D] has the lowest CV (=0.10) followed by the deaths/confirmed (=0.11) [A], deaths/(deaths + recoveries) (=0.42) [B], Bayes based on (C) (=0.49) [E], and deaths/(cases x days ago) (=0.59) [C]. All final CFRs will be equal using the formula (from, A to E). A dashboard was developed for the daily reporting of the CFR. The CFR (3.7%) greater than the prior CFR of 2.2% was evident in LSBHRS, increasing the CFR. A dashboard was created to present the CFRs on COVID-19.

Conclusion: We suggest examining both trends of the Bayes based on both deaths/(cases 7 days ago) and deaths/confirmed cases as a reference to the final CFR. An app developed for displaying the provisional CFR with the 2 CFR trends can improve the underestimated CFR reported by WHO and media.

Editor: Oliver Schildgen.

Declarations: Ethics approval and consent to participate: not applicable.

All data were downloaded from Google Sheet.

Consent to publish: Not applicable.

There are no sources of funding to be declared.

The authors have no conflicts of interest to disclose.

Supplemental Digital Content is available for this article.

Availability of data and materials: All data used in this study is available in Supplemental Digital Content files, <http://links.lww.com/MD/E101>.

The datasets generated during and/or analyzed during the current study are publicly available

^a Center for Quality Management, Chi Mei Medical Center, Liouying, ^b Department of Industrial Engineering and Management, National Yunlin University of Science and Technology, Yunlin, ^c Medical School, St. George's University of London, London, United Kingdom, ^d Research Department, Chi-Mei Medical Center, Tainan, ^e USA Sports Medicine, Sherman oaks, California, ^f Department of Optometry, Chung Hwa University of Medical Technology, Jen-Teh, ^g Department of Ophthalmology, Chi-Mei Medical Center, Yong Kang, Tainan City, Taiwan.

* Correspondence: Shu-Chun Kuo, Chi-Mei Medical Center, 901 Chung Hwa Road, Yung Kung Dist., Tainan 710, Taiwan (e-mail: ophkuo@gmail.com).

Copyright © 2020 the Author(s). Published by Wolters Kluwer Health, Inc.

This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial License 4.0 (CCBY-NC), where it is permissible to download, share, remix, transform, and buildup the work provided it is properly cited. The work cannot be used commercially without permission from the journal.

How to cite this article: Chang CS, Yeh YT, Chien TW, Lin JCJ, Cheng BW, Kuo SC. The computation of case fatality rate for novel coronavirus (COVID-19) based on Bayes theorem: an observational study. *Medicine* 2020;99:21(e19925).

Received: 17 February 2020 / Received in final form: 19 March 2020 / Accepted: 19 March 2020

<http://dx.doi.org/10.1097/MD.00000000000019925>

Abbreviations: CFR = case fatality rate, COVID = Novel Coronavirus, CV = coefficient of variance, LSBHRS = the level of staff burnouts and healthcare resource scarcities, MERS = the Middle East respiratory syndrome, SARS = the severe acute respiratory syndrome, WHO = World Health Organization.

Keywords: Bayes theorem, case fatality rate, coefficients of variance, COVID-19, dashboard, World Health Organization

Key Points

- Using Bayes theorem to verify the over-loaded burdens of treating patients and capacities to contain the outbreak (i.e., outbreak [LSBHRS]) may increase the CFR.
- Suggesting the interception of the 2 trends of the combined Bayes and deaths/(cases 7 days ago) and the Bayes based CFR as a reference to the final CFR.
- An app developed for displaying the provisional CFR with the 2 CFR trends to modify the underestimated CFR reported by WHO and media.

1. Introduction

Since the outbreak of Covid-19 in Wuhan City, China, on January 30, 2020,^[1,2] a total of 128,343 confirmed cases, 4720 death, and 68,324 recovered and discharged patients have been reported as of March 13, 2020,^[3] involving 31 provinces or cities in China as well as 129 countries/regions out of China.^[4] The total number of deaths (=4720) surpassed the severe acute respiratory syndrome (SARS) in 2003 (final toll of 774 deaths) and the Middle East respiratory syndrome (MERS) in 2012 (final toll of 858 deaths).^[5–7]

1.1. Concerns on case fatality rate

When a new disease (e.g., COVID-19) starts to spread, many questions emerge: how is the virus transmitted, how long is the incubation period, what is the influence of those asymptotically infected, what is the definite reproductive number (i.e., so-called R_0), how long does viral shedding persist after symptoms faded, who is at risk for a severe course, and how high is the case fatality rate (CFR), defined as the total number of deaths known by today divided by the total number of known cases confirmed by today.^[7] The CFR is related to the questions: how deadly is this? and how many people will die in this outbreak? SARS, MERS, Ebola, and H1N1 yielded a real CFR of 9.6%, 34.4%, 73%, and 0.4%, respectively,^[5–9] the CFR for COVID-19 is of interest to numerous researchers.^[7,10,11]

Due to the underestimated CFR using the formula (=deaths/confirmed),^[5] particularly at the fast-spreading period, several other calculations of CFR were proposed,^[12–15] such as deaths/(deaths+ recoveries) and deaths/(cases x days ago), where x denotes the lag days from symptom onset to death.^[3] None of the provisional CFRs can accurately predict the real (final) CFR^[7] based on the preliminary COVID-19 data.^[6] Ignoring the underline meaning of the CFR and carelessly reporting imprecise information with CFR (=deaths/cases) might be biased; see the next section.

1.2. CFR of 2% reported by WHO and media

The World Health Organization (WHO), in a press conference on January 29, 2020, announced the death rate of COVID-19 was

2% based on the calculation of CFR (=deaths/cases).^[16,17] It was substantially underestimated due to assuming no lag days from symptom onset to death and all currently infected cases totally recovered by computing the CFR, particularly, in the early stages of this COVID-19 epidemic.^[11]

Accordingly, many big media platform released news with such statements as below:

“The mortality rate for the new coronavirus is about 2.1%, currently far lower than the 9.6% of SARS.”^[18] “With a fatality rate of around 2%, which experts agree appears to be the current level for the virus.”^[19] Health experts say they are encouraged by the steady rise in the number of recoveries. They take it as evidence that the treatments meted out have been effective and that the virus does not appear to be as deadly as SARS. SARS had a mortality rate of 9.6 percent, and about 2 percent of those reported to have been infected with the new coronavirus have died.”^[20] “The Wuhan coronavirus seems to have a low fatality rate, and most patients make full recoveries.”^[21] “The number of confirmed cases and deaths indicates that it is around 2 percent, significantly lower than SARS’ 10 percent.”

However, “many epidemiologists and people who are following this outbreak closely are assuming that it’s probably quite a bit more widespread than the case counts suggest.”^[22] The statements of underestimated CFR that was false and^[16] motivated us to ask whether this provisional CFR was truly underestimated with false messages to the public and intrigued us to analyze the data with various CFR calculations in this study.

1.3. Bayes theorem used for computing CFR

Bayes theorem (alternatively Bayes law or Bayes rule) describes the probability of an event, based on prior knowledge of conditions that might be related to the event.^[23] For instance, the equation, $P(A1|B) = \frac{P(B|A1)P(A1)}{P(B)}$, is to estimate the probability of ($P(A1)$) based on the condition B . Someone have cancer ($A1$, or death in this study) caused by their age ($P(B)$). The age can be used to more accurately assess the probability of cancer than can be done without the knowledge of the age using Bayes theorem to estimation.

In the current study, we are motivated to apply Bayes theorem to estimate the CFR probability ($P(A1|B)$) through the likelihood of the observed events ($P(B)$), such as the level of staff burnouts and healthcare resource scarcities (LSBHRS) across countries/regions of COVID-19.

We assumed that supportive treatment is crucial for severe respiratory disease. Differences in CFRs may be caused by differences in medical resources among a large epidemic versus care of individual cases. Hence, it is hypothesized that the over-loaded burdens of treating patients and capacities to contain the outbreak (i.e., LSBHRS) may increase the CFR probability ($P(A1|B)$), where $A1$ denotes the death, and B stands for the factor (or cause) of the death.

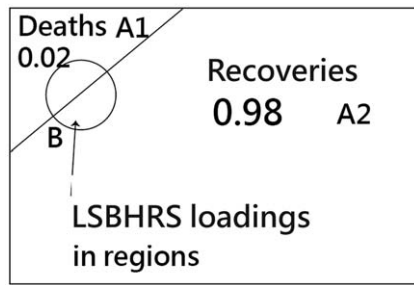


Figure 1. The CFR calculation based on Bayes' theorem when the prior probability was known. CFR=case fatality rate.

1.4. The aims of this study

The aims of the current study are to compare the differences of the proposed CFR formula and design an app that allows better understandings of the daily CFR on the COVID-19 situations.

2. Methods

2.1. Data source

We downloaded COVID-19 outbreak numbers in countries/regions on February 14, 2020, from the Github website^[24] that contains information on confirmed cases in >30 Chinese locations and other countries/regions. All downloaded data were publicly deposited on the website.^[24] Ethical approval was not necessary for this study because all the data were obtained from the Github website on the Internet.

2.2. Bayes theorem used for estimating CFRs

We designed the study structure of Bayes theorem used for estimating the CFRs of COVID-19. The research question is to investigate the post-CFR probability of the confirmed cases (i.e., $P(A1|B)$ in Eq. (1) caused by the prior probabilities (i.e., $P(A1)$ and $P(A2)$), outbreak (LSBHRS) (i.e., the cause= $P(B)$), and the conditional probabilities (i.e., $P(B|A1)$ and $P(B|A2)$), based on Bayes theorem, see Eqs. (1) and (2) and Fig. 1.

$$P(A1|B) = \frac{P(A1 \cap B)}{P(B)} = \frac{P(B|A1) \times P(A1)}{\sum_{i=1}^2 P(B|Ai) \times P(Ai)} \quad (1)$$

$$P(B) = P(B \cap A1) + P(B \cap A2) = \sum_{i=1}^2 P(B|Ai) \times P(Ai) \quad (2)$$

where $P(A1|B)$ denotes the post-CFR, $P(B)$ stands for the cause of outbreak (LSBHRS), and $P(B|Ai)$ represents the conditional probabilities surveyed from deaths (=A1) and recoveries (=A2), see Fig. 1.

2.2.1. The prior CFR and 1-CFR. The prior CFR probability (e.g., 0.02) is denoted by the deaths (A1). The remainder (A2) is defined by (1 - CFR) (e.g., 0.98), see Table 1.

The goal is to compute the $P(A1|B)$ (i.e., post-CFR) in terms of the outbreak (LSBHRS) (=P(B)), see Fig. 1. The LSBHRS is referred to as the next section.

2.2.2. The proportion of LSBHRS for each region. The shares of the outbreak (LSBHRS) (=P(B)) are proportioned in countries/regions based on currently infected cases (defined as the confirmed cases subtracted by deaths and recoveries), indicating the medical loadings on COVID-19. For instance, the daily currently infected cases are reported by 4 designated regions as 20, 30, 50, and 100, respectively. The outbreak (LSBHRS) proportions for 4 regions (i.e., just exemplified in this section) are thus computed as $0.1(=20/[20 + 30 + 50 + 100])$, 0.15, 0.25, and 0.5 based on Eq. (3), see the computation in Table 1.

$$B = \sum_j^n V_j = \sum_j^n \frac{LSBHRS_j}{TotalLSBHRS} = 1.0, \quad (3)$$

where n is the number of regions, Total_LSBHRS means the summation of the outbreak (LSBHRS) (i.e., the currently infected cases) in all regions, and $LSBHRS_j$ stands for the outbreak (LSBHRS) in a particular region.

2.2.3. Proportions of deaths and recoveries for regions. The summation of proportions of deaths (i.e., A1) in all regions is defined in Eq. (4).

$$A1 = \sum_j^n W_j = \sum_j^n \frac{Deaths_j}{TotalDeaths} = 1.0 \quad (4)$$

where n is the number of regions, Total_deaths means the summation of death tolls in all regions, and deaths stand for the number of deaths in a particular region.

Similarly, the summation of all recovered proportions (i.e., A2) in all regions is defined in Eq. (5), see Table 1.

$$A2 = \sum_j^n W_j = \sum_j^n \frac{Recoveries_j}{Totalrecoveries} = 1.0, \quad (5)$$

Both proportions (i.e., shares) for the deaths (i.e., A1) and the recoveries (i.e., A2) in each region are assumed (i.e., just exemplified in this section) and allocated as the 2 of {0.1, 0.2, 0.3, 0.4} and {0.15, 0.25, 0.3, 0.3} (Eq. (4)), see Table 1.

2.2.4. To compute the provisional CFR. From Eq. (1), the CFR can be obtained in Eq. (6), see Table 1.

$$CFR = P(A1|B) = \frac{P(B|A1) \times P(A1)}{P(B|A1) \times P(A1) + P(B|A2) \times A(A2)}, \quad (6)$$

Referring to the example in this section, the CFR can be computed by the formula ($= \frac{P(B|A1) \times P(A1)}{P(B|A1) \times P(A1) + P(B|A2) \times 0.98} = \frac{0.315 \times 0.02}{0.315 \times 0.02 + 0.2775 \times 0.98} = 0.023$) where $P(B|A1) = (0.1 \times 0.1 + 0.15 \times 0.2 + 0.25 \times 0.3 + 0.5 \times 0.4) = 0.315$, $P(B|A2) = (0.1 \times 0.15 + 0.15 \times 0.25 + 0.25 \times 0.3 + 0.5 \times 0.3) = 0.2775$, Eq.(3).

We can see that the post probability of CFR is increased from the prior probability of 0.02 to 0.023 due to the unequal weights in regions, see scenario (A) in Table 1. The final CFR equals the prior CFR at the end of the outbreak, see scenario (B) in Table 1.

Similarly, the post-CFR for a region can be obtained by Eq. (7)

$$CFR = P(A1|B) = \frac{P(B|share_{A1}) \times P(A1)}{P(B|share_{A1}) \times P(A1) + P(B|share_{A2}) \times A(A2)}, \quad (7)$$

where share_A1 is the product of LSBHRS% and deaths% [= LSBHRS(j) × deaths(j)] and share_A2 equals the product of LSBHRS% and deaths% [= LSBHRS(j) × recoveries(j)], see the example (1) in the next section.

Table 1
The computation of CFR using an example in this section.

Prior Prob.	LSBHRS(%)	Deaths	Recoveries
		0.02	0.98
		A1	A2
Daily currently infected cases (weighted)	Weighted	Weighted	Weighted
Region A(20)	0.10	0.1	0.15
Region B(30)	0.15	0.2	0.25
Region C(50)	0.25	0.3	0.30
Region D(100)	0.50	0.4	0.3
Eq.	(3)	0.315	0.2775
Eq.		(4)	(5)
CFR = (0.315 × 0.02)/(0.315 × 0.02 + 0.2775 × 0.98)		0.023	
Eq.		(6)	
B. No daily currently infected case (unweighted)	Unweighted	Weighted	Weighted
Region A(0)	0.25	0.1	0.15
Region B(0)	0.25	0.2	0.25
Region C(0)	0.25	0.3	0.30
Region D(0)	0.25	0.4	0.3
Eq.	(3)	0.25	0.25
Eq.		(4)	(5)
CFR = (0.25 × 0.02)/(0.25 × 0.02 + 0.25 × 0.98)		0.02	
Eq.		(6)	
C. Attributes:			
1. CFR for the first region = (0.5 × 0.4 × 0.04)/(0.5 × 0.4 × 0.04 + 0.5 × 0.3 × 0.96) = 0.053			
2. CFR = prior CFR because all LSBHRS loadings in regions are zero (see unweighted scenario)			
3. Provisional CFR = prior CFR due to weighted A1 and A2 ignored (see scenario D)			
4. CFR <> prior CFR when weighted LSBHRS loadings are substantially different			
D. Unweighted deaths and recoveries	Weighted	Unweighted	Unweighted
Region A(20)	0.10	0.25	0.25
Region B(30)	0.15	0.25	0.25
Region C(50)	0.25	0.25	0.25
Region D(100)	0.50	0.25	0.25
Eq.	(3)	0.25	0.25
Eq.		(4)	(5)
CFR = (0.25 × 0.02)/(0.25 × 0.02 + 0.25 × 0.98)		0.02	
Eq.		(6)	

CFR = case fatality rate, LSBHRS = the level of staff burnouts and healthcare resource scarcities.

2.2.5. Attributes of the provisional CFR. Four scenarios are illustrated in the scenario (C) in Table 1.

(1) The post CFR for the forth region = $\frac{0.5 \times 0.4 \times 0.04}{0.5 \times 0.4 \times 0.04 + 0.5 \times 0.3 \times 0.96} = 0.053$ when the prior CFR = 0.04. Tj

(2) The final CFR = $\frac{0.25 \times 0.02}{0.25 \times 0.02 + 0.25 \times 0.98} = 0.2$ = the prior CFR when all LSBHRS loadings in regions are zero, see scenario (B) in Table 1, and then let $P(B|A1) = (0.25 \times 0.1 + 0.25 \times 0.2 + 0.25 \times 0.3 + 0.25 \times 0.4) = 0.25$ and $P(B|A2) = (0.25 \times 0.15 + 0.25 \times 0.25 + 0.25 \times 0.3 + 0.25 \times 0.3) = 0.25$;

(3) The provisional CFR = $\frac{0.25 \times 0.02}{0.25 \times 0.02 + 0.25 \times 0.98} = 0.2$ = the prior CFR when the unweighted deaths and the weighted recoveries are applied (see scenario D in Table 1):

$$P(B|A1) = (0.25 \times 1.0) = 0.25 \text{ and } P(B|A2) = (0.25 \times 1.0) = 0.25;$$

(4) The provisional CFR = $\frac{0.25 \times 0.02}{0.6 \times 0.02 + 0.1 \times 0.98} = 0.45$ if the surveyed $P(A1)$ and $P(A2)$ of weighted LSBHRS are substantially different such as $P(B|A1) = 0.6$ and $P(B|A2) = 0.1$, where surveyed $P(A1)$ (or $P(A2)$) means the proportion from a survey (e.g., how many percentages are experienced with the severe outbreak [LSBHRS] [e.g., intensive care unit [ICU]] in deaths or in recoveries).

As such, we can apply the formula from (1) to (6) to estimate the post-CFR and verify that unequally weighted LSBHRS loadings cause the CFR probability ($P(A1|B)$) to increase when $P(B|A1) > P(B|A2)$ in Eq. (6) or $P(B|shareA1) > P(B|shareA2)$ in Eq. (7).

2.3. Task 1: Deaths on a choropleth map

We created HTML pages showing the death counts on a choropleth map. All relevant COVID-19 information on the countries/areas can be linked to dashboards on Google Maps when the color is clicked.^[2,5]

2.4. Task 2: Using the coefficient of variance (CV) to compare the features of formula on CFRs

The features and limitations were compared among the 5 CFR formulae, including

- (A) deaths/confirmed;
- (B) deaths/(deaths + recoveries);
- (C) deaths/(cases 7 days ago);

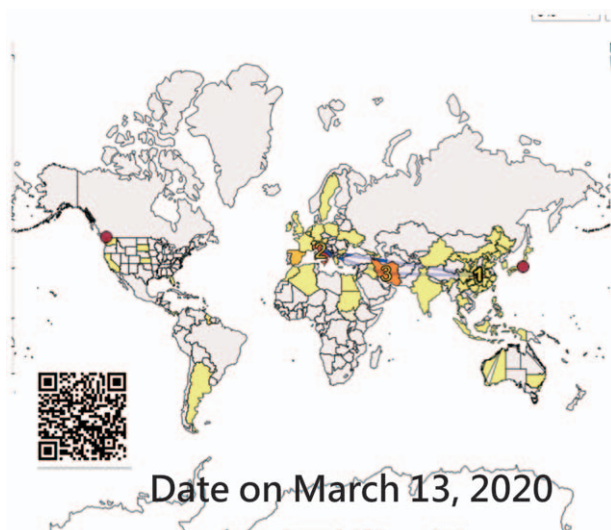


Figure 2. The distribution of the deaths on a choropleth map.

- (D) Bayes estimation on outbreak (LSBHRS) (see Eqs. from (1) to (6) and Table 1) based on the prior CFR of (A) deaths/confirmed;
- (E) Bayes based on the prior CFR of (C) deaths/(cases \times days ago).

The coefficients of variance (CV=the ratio of the standard deviation to the mean) were applied to measure the relative variability for each CFR. The lower is the CV; the higher stable is the feature of the CFR computation.

2.5. Task 4: A dashboard on Google Maps to present the daily CFR

A dashboard was developed for daily display of the CFR across various regions using the Kano model.^[26,27] In which 3 features are displayed, including the wow feature (top), one dimension quality (middle), and the essential requirement (bottom).

2.6. Task 3: The changes between prior-/post-CFRs

The changes between prior and post-CFRs were compared using the Kano model^[26,27] when the Bayes theorem was applied to compute the post-CFR. Details are included in Additional File 1 with an MS-Excel dataset.

3. Results

3.1. Deaths on a choropleth map

A choropleth map about the distribution of the deaths in regions is present in Fig. 2. The darker color means a large number of deaths in the region. We can see that Hubei province in China with a darker red color. When the color is clicked, a box will pop up and shows that infected=51,986, deaths=1318 (0.03%), and recoveries=3900. All data were extracted on February 14, 2020.

Three line graphs regarding the cumulative confirmed cases, the daily confirmed cases in the recent 7 days, and the trend was linked for interested readers to examine the details. For

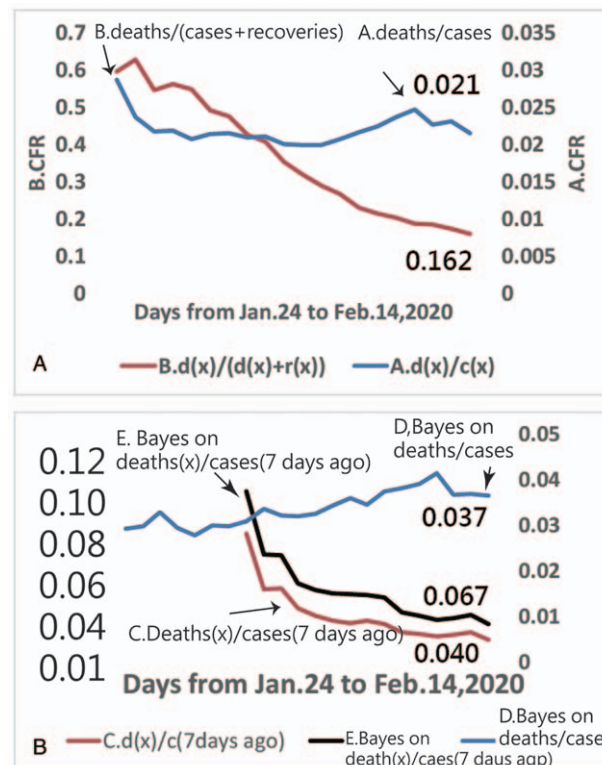


Figure 3. Comparison of methods used for computing the CFRs of Covid-19. CFR=case fatality rate.

more information about the recent COVID-19 situations, readers are invited to click on the link^[28] or scan the QR-code in Fig. 2.^[29]

3.2. Using CV to compare formula on CFRs

In Fig. 3 we can see that the Bayes estimation (D) has the lowest CV (=0.10) followed by the deaths/confirmed (=0.11) (A), deaths/(deaths+recoveries) (=0.42) (B), combined Bayes and deaths/(cases 7 days ago) (=0.49) (E), and deaths/(cases \times days ago) (=0.59) (C), see Table 2. In which, we explained the equations from A to E in details.

Each method will eventually yield an identical CFR over the course of the outbreak. Some converged from a lower percentage of CFR to a higher one (e.g., (A) and (D)); other three (e.g., (B), (C), and (E)) dropped from a higher CFR to a lower one. From Fig. 3, the combined Bayes and deaths/(cases 7 days ago) was recommended instead of the lower CFR WHO claimed and reported by media.

3.3. A dashboard to present the daily CFR

A dashboard was developed in responding to the daily CFR in comparison to each region. The CFR (=3.7% with purple color) greater than the WHO's 2% was shown on our developed dashboards based on data of February 14, 2020, using Bays estimation on axis y and prior CFR on axis x in Fig. 4. From this, we can compare the daily CFRs for each region. Other CFRs (e.g., deaths/confirmed cases on 7 days ago and deaths/[deaths+ recoveries]) are present when the bubble of interest is clicked.

Table 2
Comparison of CFR computations.

Label	CFR formula	CV	CFR
A	$d(x)/c(x)$ = deaths/confirmed cases on day x	0.11	0.021
B	$d(x)/(d(x) + r(x))$ = deaths/(deaths + recoveries) on day x	0.42	0.162
C	$d(x)/c(7 \text{ days ago})$ = deaths/confirmed cases at 7 days ago	0.59	0.040
D	Bays = Bays based on deaths/confirmed cases on day x	0.10	0.037
E	C&Bays = Bays based on deaths/confirmed cases at 7 days ago	0.49	0.067

Note. The date was on February 14, 2020; (D) Bays considers the staff burnout and resource scarcity (i.e., LSBHRS). CFR = case fatality rate.

Interested readers are invited to scan the QR-code in Fig. 4 and see the details by clicking the bubble of interest.

From Fig. 4, we can see that all bubbles sized by the post-CFR were within the one-dimensional band using the Kano model. Indicating both prior- and post-CFRs were almost equal to each other. The black bubbles would denote the post-CFRs equal to zero even if the prior-CFRs were significantly higher toward the right-hand side on axis x . The highest prior-CFR (=33%) is the Philippines (on February 14, 2020) because 3 were infected and 1 died. The highest post=CFR was Hubei (China) with post-CFR = 3.66% higher than the prior-CFR = 2.54%, indicating the heavy loadings of the outbreak (LSBHRS) leads to a higher probability of the post-CFR.

It is worth noting that the prior-CFR is derived from the deaths/confirmed by today, and the post-CFR is yielded by Eq. (6). The post-CFR (3.7%) greater than the prior CFR of 2.2% on February 14 and equal to 3.7% on March 13 was evident of the outbreak (LSBHRS) increasing the CFR due to weighted scheme applied, see Table 1.

3.4. The recent CFRs using Bayes theorem

The recent CFRs were online computed in Fig. 5, where prior CFRs are shown on axis x and post-CFRs on axis y . The highest prior-CFR was located in Suan with 1 died and 1 confirmed case on March 13, 2020, see the right-top corner, Fig. 5. All bubbles are sized by the post-CFRs and colored the areas in the Kano diagram. For instance, bubbles in the one-dimensional area are colored in blue.

The global with the big purple bubble had an increase from the prior CFR (=3.71%) to the post CFR (3.27%) in comparison to

Hubei (China) having a decrease from the prior CFR = 4.521% to the post CFR = 3.52% on March 13, 2020, due to the outbreak (LSBHRS) in deaths relatively lower than that in recoveries.

Three countries outside the one-dimensional area are Italy, Spain, and France with significantly relative post-CFRs more than the prior-CFRs. The recent CFRs are shown on Google Maps when the QR-code is scanned in Figs. 4 and 5.

4. Discussion

4.1. Principle findings

We observed that the Bayes estimation [D] has the lowest CV (=0.10) followed by the deaths/confirmed (=0.11), deaths/(deaths + recovered) (=0.42) [B], Bayes based on deaths/(cases \times days ago) (=0.49) [E], and deaths/(cases 7 days ago) (=0.59) [C]. All final CFRs will be equal. The CFR (3.7%) greater than the prior-CFR of 2.2% was evident in LSBHRS impact on the CFR. A dashboard was created to present the CFRs.

A dashboard was designed to respond to the daily overall CFR. Each individual CFR can be viewed by clicking the bubble of interest.^[30,31]

We found that the method (D) has the lowest CV (=0.10), and the (C) has the highest CV (=0.59). The (E), Bayes based on deaths/(cases 7 days ago) (=0.49), is suggested to calculate the provisional CFR to the general public. The reason for use is illustrated in the next section.

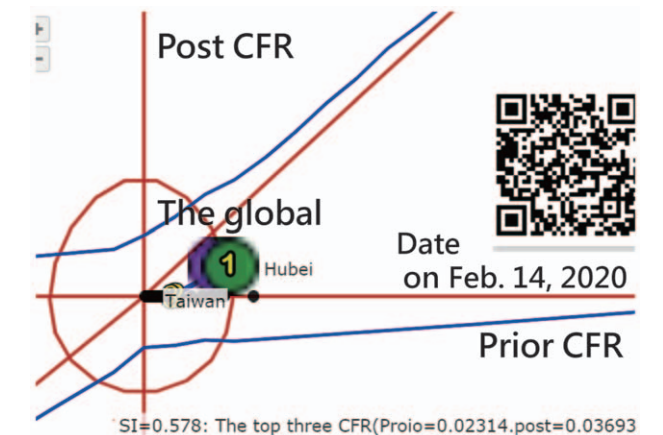


Figure 4. A dashboard showing the CFR on Google Maps ended at February 11, 2020. CFR = case fatality rate.

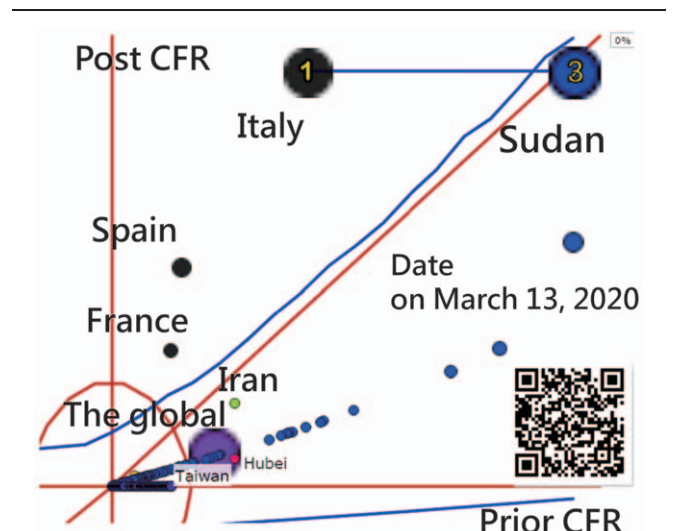


Figure 5. A dashboard showing the CFR on Google Maps ended at March 13, 2020. CFR = case fatality rate.

4.2. What this finding adds to what we already knew

CFR is a measure of the “severity” of a disease and is defined as the proportion of cases of a specified disease or condition which are fatal.^[16] We believe that WHO voiced the CFR (=2%)^[17] with misinformation (i.e., underestimation) which is only based on confirmed cases, ignoring the lag time between the onset of symptoms and progression to deaths, as well as in mild cases assuming the LSBHRS being equal across countries/regions.

WHO did not explain to the media and public what CFR implicates. The understated CFR has been shared worldwide.^[18–22] It is clear that the conventional CFR defined as deaths/cases in the current fast-growing epidemic gives us a totally useless and unreliable figure when it comes to an outbreak like COVID-19 in the past 50 days since February 14, 2020.

The real fatality rate is hard to estimate because we have many currently infected cases yet known the distributions on age, sex, and other factors such as the proportion of severity in intensive care units (ICU) and patients with multiple comorbidities, and the CFR taking into account those clinically diagnosed cases as confirmed cases determined in Hubei since February 14, 2020.^[32]

What WHO and media did by publishing CFR (=2%) has served to avoid public panic for a period of time by displaying the comparison of data with the true CFRs of SARS (9.6%), MERS (34.4%), Ebola (73%), and H1N1 (0.4%).^[5–9]

However, true and final CFRs are usually much bigger than the provisional CFRs according to the SARS experience: when WHO first reported daily statistics, the SARS death rate was about 2%. It was 2.4% on March 17 and 1.8% on March 18 (in 2003).^[33] In April 2003, CFR was increased to 5.6% instead of an initial 1.8%. The final CFR in the middle of the summer (2003) raised up to almost 10%.^[33]

That is the reason why we suggested using the Bayes based on deaths/(cases 7 days ago) in reporting the provisional CFR to the general public. Through this, we can avoid the negative effect of the underestimated CFR and prepare the population, including doctors, authorities, and epidemiologists, etc., to face the real threat of COVID-19.

4.3. What it implies and what should be changed

Dr Michael Ryan, Executive Director of WHO, addressed that the COVID-19 fatality rate is 2% in a press conference on January 29, 2020.^[16] However, most of the confirmed cases of this novel coronavirus are still hospitalized, many of them in ICUs. The virus does not kill a person immediately, hence creating a lag between infection confirmation and death.^[16] For example, only patients that had died by today (February 14, 2020) would have been taken into account in the provisional CFR, but these individuals would have been registered as “confirmed cases” a few days or more than a week ago. The provisional CFR is, therefore, underestimated due to the potential deaths of currently confirmed cases in the next couple of days (a.k.a. lag days).^[34] As the number of death increased, the death rate has also steadily risen, leaving the officials worried, based on the CFR trend calculated as deaths/cases, see Fig. 3 (upper). Due to the imprecise and carelessness in the calculation, the media has misled the public that “CFR of 2% on COVID-19 is less than the 10% in SARS.”^[16–21]

Although we compared the final CFR for the SARS epidemic with current CFR for COVID-19 on the same basis of deaths/cases, what has been forgotten is that CFR was 1.8% for SARS on an early stage in 2003 epidemic calculated using this underestimated method. The current calculation of CFR for a growing epidemic

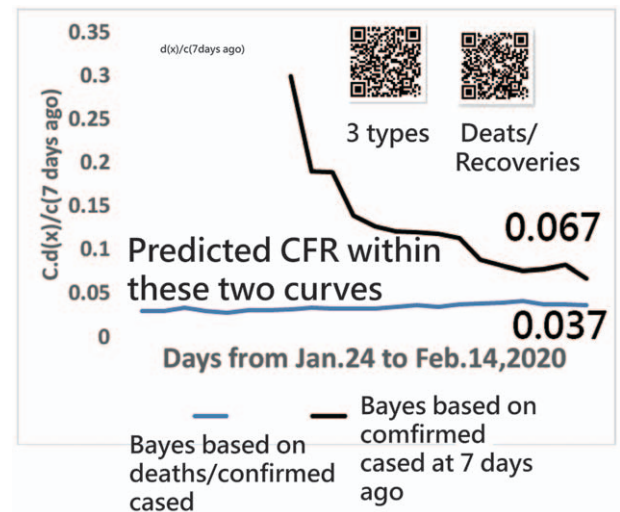


Figure 6. To examining the interception of the 2 trends.

stage is not suitable for representing the final CFR. An adjusted CFR is required in an epidemic to better elucidate the outbreak.

We suggest that in a fast-growing epidemic, CFR cannot be used at all when referring to CV or the slope of the CFR trend (e.g., in Fig. 3). If the true CFR managed to reach over 10% while the provisional CFR was just under 3%, it makes no sense to use such definition in the calculation of CFR (deaths/cases).

The best way to calculate the provisional CFR is based on cases that already concluded by either death or recovery. However, the epidemic is spreading fast, and there is not enough time to collect data on how many patients of first, say 100, patients recovered and died by any particular day. Examining the interception of the 2 trends (Fig. 6) might be a reference to predict the final CFR, which should be evident in the future. With this change, it is hopeful that the public would not be misinformed again.

4.4. Strengths of this study

In our suggested CFR calculation of the method of Bayes based on deaths/(cases 7 days ago), 3 pivotal components have been focused in the current study: assuming unequal LSBHRS proportions (based on the daily no-outcome-yet cases) across all countries/regions in COVID-19; using Bayes theorem to adjust the daily provisional CFR and observing whether the CFR trend has been flattened at stationarity phase (i.e., approaching real CFR); considering the lag time between symptom to death as the thread alerted at the fast-growing (i.e., beginning) period.

Although it is hard to accurately estimate the provisional CFR because we do not know the distributions of age, severity, number in ICU, etc., an appropriate death lag time (e.g., 7 days) was designed in this study. The longer this estimated period is, the higher the provisional CFR will be.^[3,16]

As with all forms of Web-based technology, advances in health communication technology are happening every moment.^[35] The provisional CFR shown on a dashboard is practical and worth replicating in other fields of a disease outbreak, as we demonstrated in Figs. 4 and 5. Interested readers are recommended to scan the QR codes in Figs. 4–6 and read the details about the corresponding line graph on trends of the outbreak.

4.5. Limitations and future studies

Our study has some limitations. First, although the data were downloaded from Google Sheet on a daily basis, we are concerned with different criteria for determining confirmed cases among regions that affect the CFR. For instance, the confirmed cases are from clinically diagnosed cases in Hubei since February 14, 2020.^[32]

Second, although we applied Bayes theorem to estimate the provisional CFR using Eqs. from (1) to (5), the post-CFR calculated by Bayes theorem (Eq. (6)) is strongly related to the prior = CFR. The result due to LSBHRS among regions drives a little bit of increase compared with the CFR calculated by the deaths over cases assuming all medical resources and staff burnout are equal. The Bayes based on deaths/(cases 7 days ago) suggested in this study is worthy of investigation in the future.

Third, the study was based on the data from the Github website.^[24] The Google team has announced that the service will be terminated and transformed into Github,^[24] providing the same information. Due to the download limit, the Google team will soon stop the update of this Google Sheet, which might affect the data being transferred effectively and efficiently on the daily calculated CFR (e.g., in Figs. 4 and 5).

Fourth, the post-CFRs (calculated by Bayes theorem) in Fig. 4 confused us with the gap from the prior-CFRs (resulted from the deaths/confirmed by today), such as the Philippines (0%: 33%) with confirmed = 3, deaths = 1, and recovered = 1. Owing to the LSBHRS ($= [3-1-1]/64,460$) substantially relatively less than Hubei (China) ($= [51,986-1318-3900]/64,460$) (3.66%: 2.54% with confirmed = 51,986, deaths = 1318, recovered = 3900), the expected CFR for the Philippines thus sharply decreases after computation based on Bayes theorem.

Finally, we suggest examining the interception of the 2 trends (Fig. 6) as a reference to the final CFR, which should be further verified in the future.

5. Conclusion

We suggest examining the interception of the 2 trends of the Bayes based on both deaths/(cases 7 days ago) and death/confirmed cases as a reference to the final CFR. An app developed for displaying the provisional CFR with the 2 CFR trends is encouraged to ameliorate the underestimated CFR reported by WHO and media.

Acknowledgments

The authors thank Enago (www.enago.tw) for the English language review of this manuscript.

Author contributions

Tsair-Wei Chien developed the study concept and design. Tsair-Wei Chien, Chi-Sheng Chang, Jui-Chung John Lin, and Yu-Tsen Yeh analyzed and interpreted the data. Shu-Chun Kuo monitored the process of this study and helped in responding to the reviewers' advice and comments. Tsair-Wei Chien drafted the manuscript, and all authors provided critical revisions for important intellectual content. The study was supervised by Shu-Chun Kuo. All authors read and approved the final manuscript.

References

[1] Nishiura H, Kobayashi T, Yang Y, et al. The rate of underascertainment of Novel Coronavirus (2019-nCoV) infection: estimation using Japanese passengers data on evacuation flights. *J Clin Med* 2020;9:419. doi:10.3390/jcm9020419.

[2] Zhao S, Lin Q, Ran J, et al. Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: a data-driven analysis in the early phase of the outbreak. *Int J Infect Dis* 2020;92:214–7.

[3] JHC. Coronavirus disease 2019(COVID-19) outbreak. Available at <http://www.healthup.org.tw/kpiall/covid19dashboard.asp>. Accessed April 14, 2020

[4] Leszkiewicz A. Coronavirus disease 2019(COVID-19) outbreak. Available at <http://www.healthup.org.tw/kpiall/covid19dashboard2.asp>. Accessed April 14, 2020

[5] Suwatarat N, Apisarnthanarak A. Risks to healthcare workers with emerging diseases: lessons from MERS-CoV, Ebola, SARS, and avian flu. *Curr Opin Infect Dis* 2015;28:349–61.

[6] NBC news. Are coronavirus diseases equally deadly? Comparing the latest coronavirus to MERS and SARS. Available at <https://www.nbcnews.com/health/health-news/coronavirus-diseases-comparing-covid-19-sars-mers-numbers-n1150321>. Accessed May 4, 2020

[7] Majumder MS, Rivers C, Lofgren E, et al. Estimation of MERS-Coronavirus reproductive number and case fatality rate for the Spring 2014 Saudi Arabia outbreak: insights from publicly available data. *PLoS Curr* 2014; doi:10.1371/currents.outbreaks.98d2f8f3382d84f390736cd5f5fe133c.

[8] Fraser C, Donnelly CA, Cauchemez S, et al. Pandemic potential of a strain of influenza A (H1N1): early findings. *Science* 2009;324:1557–61.

[9] Singh SK. Middle East respiratory syndrome virus pathogenesis. *Semin Respir Crit Care Med* 2016;37:572–7.

[10] Kamigaki T, Oshitani H. Epidemiological characteristics and low case fatality rate of pandemic (H1N1) 2009 in Japan. *PLoS Curr* 2009; doi: 10.1371/currents.RRN1139.

[11] Battegay M, Kuehl R, Tschudin-Sutter S, et al. 2019-novel Coronavirus (2019-nCoV): estimating the case fatality rate - a word of caution. *Swiss Med Wkly* 2020;150:w20203. doi:10.4414/SMW.2020.20203.

[12] Schwartz DA, Graham AL. Potential maternal and infant outcomes from (Wuhan) coronavirus 2019-nCoV infecting pregnant women: lessons from SARS, MERS, and Other Human Coronavirus Infections. *Viruses* 2020;12:194. doi:10.3390/v12020194.

[13] Chen J. Pathogenicity and transmissibility of 2019-nCoV-A quick overview and comparison with other emerging viruses. *Microbes Infect* 2020;22:69–71.

[14] Garske T, Legrand J, Donnelly CA, et al. Assessing the severity of the novel influenza A/H1N1 pandemic. *BMJ* 2009;339:b2840. doi:10.1136/bmj.b2840.

[15] Yoshikura H. On case-fatality rate: review and hypothesis. *Jpn J Infect Dis* 2012;65:279–88.

[16] Leszkiewicz A. Novel Coronavirus (2019-nCoV) fatality rate: WHO and media vs logic and mathematics; 2020. Available at: <http://avatorl.org/en/novel-coronavirus-2019-ncov-fatality-rate-who-and-media-vs-reality/>. Accessed April 20, 2020

[17] Ryan M. Novel Coronavirus (2019-nCoV) fatality rate is 2%; 2020. Available at: <https://twitter.com/WHO/status/1222550294942191623>. Accessed April 20, 2020

[18] BBC. Coronavirus: China admits 'shortcomings and deficiencies'; 2020. Available at: <https://www.bbc.com/news/world-asia-china-51362336>. Accessed April 20, 2020

[19] CNN. January 30 coronavirus news; 2020. Available at: https://edition.cnn.com/asia/live-news/coronavirus-outbreak-01-30-20-intl-hnk/h_a76088014244cc478838aa9d73a3440b. Accessed April 20, 2020

[20] The New York Times. Hong Kong Reports First Death From Coronavirus Outbreak; 2020. Available at: <https://www.nytimes.com/2020/02/03/world/asia/coronavirus-china.html>. Accessed April 20, 2020

[21] Business Insider. The Wuhan coronavirus seems to have a low fatality rate, and most patients make full recoveries. Experts reveal why it's causing panic anyway; 2020. Available at: <https://www.businessinsider.sg/wuhan-coronavirus-unnecessary-panic-experts-say-2020-1/#XjTXVCqDSgy.twitter>. Accessed April 20, 2020

[22] The Harvard Gazette. Harvard epidemiologist Mina says outbreak more widespread than thought, and uncertainties abound; 2020. Available at: <https://news.harvard.edu/gazette/story/2020/02/as-confirmed-cases-of-coronavirus-surge-path-grows-uncertain/>. Accessed April 20, 2020

[23] Joyce J, Zalta EN. Bayes' Theorem. *The Stanford Encyclopedia of Philosophy* (Spring 2019 ed.), Metaphysics Research Lab, Stanford University; 2003

[24] Google Team. 2019 Novel Coronavirus (nCoV) Data Repository; 2020. Available at: <https://github.com/CSSEGISandData/2019-nCoV>. Accessed April 9, 2020

[25] Chien TW, Wang HY, Hsu CF, et al. Choropleth map legend design for visualizing the most influential areas in article citation disparities: a bibliometric study. *Medicine* (Baltimore) 2019;98:e17527.

- [26] Kano N, Seraku N, Takahashi F, et al. Attractive quality and must-be quality. *J Jpn Soc Qual Control* 1984;41:39–48.
- [27] Lin CH, Chou PH, Chou W, Chien TW. Using the Kano model to display the most cited authors and affiliated countries in schizophrenia research [published online ahead of print, 2019 Dec 17]. *Schizophr Res* 2019; S0920-9964(19)30493-1.
- [28] Chien TW. Three types of case fatality rates of COVID-19 on a dashboard; 2020. Available at: <http://www.healthup.org.tw/kpiall/covid19manu.htm>. Accessed April 13, 2020
- [29] Chien TW. Death tolls in regions on a choropleth map; 2020. Available at: <http://www.healthup.org.tw/kpiall/ncovworld.asp?mtypeabc=2>. Accessed April 13, 2020
- [30] Chien TW. Three types of case fatality rates of COVID-19 on a dashboard; 2020. Available at: <http://www.healthup.org.tw/kpiall/wuhenddeath.asp>. Accessed April 13, 2020
- [31] Chien TW. Case fatality rates of COVID-19 using death divided by confirmed cases on a dashboard; 2020. Available at: <http://www.healthup.org.tw/kpiall/wuhenddeath2.asp>. Accessed April 13, 2020
- [32] BBC. Coronavirus: sharp increase in deaths and cases in Hubei; 2020. Available at: <https://www.bbc.com/news/world-asia-china-51482994>. Accessed April 14, 2020
- [33] Altman LK. Death Rate From Virus More Than Doubles, Varying Sharply by Country; 2020. Available at: <https://www.nytimes.com/2020/04/22/world/death-rate-from-virus-more-than-doubles-varying-sharply-by-country.html>. Accessed April 14, 2020
- [34] Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020;395:497–506.
- [35] Mitchel SJ, Godoy L, Shabazz K, et al. Internet and mobile technology use among urban African American parents: survey study of a clinical population. *J Med Internet Res* 2014;16:e9.