Estimation of standardized real-time fatality rate for ongoing epidemics

Yuanke Qu[1], Chun Yin Lee [2]*

1 Department of Computer Science and Engineering, Guangdong Ocean University, Zhanjiang, People's Republic of China, **2** Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong

* james-chun-yin.lee@polyu.edu.hk

Abstract

Background

The fatality rate is a crucial metric for guiding public health policies during an ongoing epidemic. For COVID-19, the age structure of the confirmed cases changes over time, bringing a substantial impact on the real-time estimation of fatality. A 'spurious decrease' in fatality rate can be caused by a shift in confirmed cases towards younger ages even if the fatalities remain unchanged across different ages.

Methods

To address this issue, we propose a standardized real-time fatality rate estimator. A simulation study is conducted to evaluate the performance of the estimator. The proposed method is applied for real-time fatality rate estimation of COVID-19 in Germany from March 2020 to May 2022.
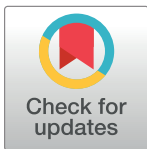
Findings

The simulation results suggest that the proposed estimator can provide an accurate trend of disease fatality in all cases, while the existing estimator may convey a misleading signal of the actual situation when the changes in temporal age distribution take place. The application to Germany data shows that there was an increment in the fatality rate at the implementation of the 'live with COVID' strategy.

Conclusions

As many countries have chosen to coexist with the coronavirus, frequent examination of the fatality rate is of paramount importance.

Introduction

As the prospect of herd immunity fades, achieving zero COVID-19 infections appears unfeasible without extremely stringent public health measures, which will come at the cost of a

deteriorating economy. After more than two years of combating the COVID-19 pandemic, many parts of the world, including Europe, the United States, and much of Asia, have phased out hard restrictions and returned to normal life since early 2022. While a surge in COVID-19 cases was expected following the reopening of borders and the easing of public health restrictions, the rate of death, hospitalization, and serious illness have remained low in most countries. This suggests that the number of COVID-19 cases is no longer an appropriate metric for determining which and when the control measures are implemented. Health professionals advise against solely relying on the daily number of COVID-19 cases to inform public health policies but to adopt a more comprehensive approach tracking the disease severity, which provides a more accurate picture of disease burden and fatalities, especially in the era of vaccination [1, 2].

The literature on the severity of COVID-19 has primarily focused on the case fatality rate (CFR), which measures the proportion of the number of deaths among confirmed cases up to a specific time, for different populations. Considering that there is a significant time lag between disease onset and death for COVID-19 infection, various methods have been proposed to reduce the bias in the estimation of CFR. These include restricting the analysis to cases with definitive clinical outcomes (died or recovered) or adjusting the value of the cumulative number of cases in the denominator, using the information on the epidemic growth rate and the distribution of time from disease onset to death [3–5].

Nevertheless, estimators developed based on aggregated counts, namely the cumulative numbers of cases and deaths, are generally unable to capture changes in disease fatality over time. In the context of the persisting COVID-19 pandemic involving frequent viral mutations, calculating the overall fatality rate up to a certain point in time may not provide meaningful insights. Instead, continuously monitoring the real-time fatality rate, or in other words, tracking changes in disease fatality rate, is crucial. This metric often yields more valuable information that can inform public health policies, including shifts in disease virulence, the efficacy of response strategies, and the quality of healthcare services. To this end, Qu et al. [6] proposed a real-time fatality rate estimator adjusted for reporting delay (rtaCFR) based on the fused lasso technique. The rtaCFR was shown to outperform other fatality rate estimators in tracking disease severity throughout the COVID-19 pandemic in Germany. They considered the rtaCFR a key metric to guide public health policies as it can capture the changes in disease fatality in real-time, such as a decreasing trend due to the implementation of effective interventions, or an upward trend due to the emergence of a more virulent or vaccine-resistant variant.

However, a limitation of the study is that the effects of explanatory variables, such as age, gender and other demographic variables are not considered. Since the onset of the COVID-19 pandemic in 2020, it has been well-known that the disease outcome of COVID-19 patients varies markedly by age, with older patients facing a higher risk of severe illness and death [7–9]. Therefore, age demographic has long been highlighted as a crucial factor in elucidating the disparities in COVID-19 fatality rates among different countries [10–15]. For instance, India cited its low case fatality rate as evidence of its success in containing the COVID-19 crisis. However, the seemingly low fatality is presumably attributed to its young-skewed population, who are less susceptible to severe disease outcomes [16]. When accounting for age distribution, India could be in a much worse position than other nations. This example illustrates how the population age structure can affect the overall fatality rate. Nevertheless, limited attention has been paid to the fact that the age distribution among confirmed cases also changes throughout the epidemic. Staerk et al. [17] showed that variations in the estimated effective infection fatality rates could be attributed to a shift in the age distribution of confirmed cases towards older age groups. Such an impact could lead to misinterpretations of the real-time fatality rates, and misguided public health policies.

In this article, we propose a standardized real-time adjusted case fatality rate estimator (srtaCFR) that can serve as an indicator for the government to better understand the situation in the midst of a pandemic. The proposed method corrects the bias of the commonly used CFR by adjusting for the time lags between illness onset and death, and is more sensitive in picking up the changes in fatality rate during an ongoing epidemic. A distinguished feature of the new estimator compared to the existing rtaCFR in [6] is its ability to remove the effects of some time-varying confounding factors attributed to the changes in disease severity when the relevant information is available. One typical application of the proposed method is to adjust for the impact induced by the changes in the age distribution of confirmed cases when monitoring changes in disease virulence over time. For instance, at the beginning of the COVID-19 epidemic, the majority of cases were concentrated in the older and more vulnerable population, which partly contributed to the relatively high fatality rate therein. Hence, the overall fatality could be overestimated even if the disease virulence remains unchanged. Similarly, if there is a shift towards younger individuals being diagnosed with COVID-19, it could potentially lead to a spurious decline in disease severity within that period. Therefore, adjusting for the effects of age distribution helps to provide a more accurate picture of COVID-19, and is crucial in informing policy decisions. Simulation studies demonstrate the superiority of the proposed srtaCFR in terms of high accuracy and sensitivity in capturing the changes in disease severity throughout the epidemic. Under the influence of the temporal age distribution, the proposed method is shown to be empirically unbiased in all scenarios, whereas the existing rtaCFR performs well only when the underlying age-specific fatality rates are constant and identical over time; otherwise, rtaCFR may result in a delayed or inaccurate representation of the trends in disease severity. We illustrate the usefulness of the proposed method through an application to the COVID-19 data in Germany from March 2020 to May 2022.

Methods

Basic setting

Assume that epidemiological data are collected regularly, such as on a daily basis, during an emerging epidemic. Let τ be the current time point or the end of an epidemic. Throughout the article, we always label a time point in the subscript of a notation, if any. For $t = 1, \ldots, \tau$, the basic data collected at time t includes the number of confirmed cases c_t and the number of reported deaths d_t. It is assumed that each confirmed case will eventually die or recover. For infectious diseases, there will be a significant time lag between illness onset and death. Let F denote the cumulative distribution function of the time from disease onset to death, which can be informed by some prior knowledge, such as estimates obtained from previous outbreaks or from the analysis of some hospitalized cases where individual-level data are available. Given an outcome of death, we assume that the death is reported s days after the time of diagnosis with probability $f_s = F_s - F_{s-1}$, the difference of F evaluated at time s and $s - 1$ respectively.

Real-time fatality rate estimator

A real-time fatality rate estimator taking into account reporting delay in death has been recently proposed by Qu et al. [6], where minimal epidemiological data, namely the observed time series of c_t and d_t, and the distribution F, are required as inputs. We will briefly mention the existing estimator to facilitate the discussion of its standardized version. Let p_t denote the (unknown) probability that a confirmed case reported on day t will eventually die from the disease, for $t = 1, \ldots, \tau$. Clearly, the expected total number of deaths at time t ignoring reporting delay is given by $c_t p_t$. In the presence of reporting delay, these $c_t p_t$ individuals will die subsequently with a delay distributed according to F. In this model, the expected number of deaths

at time t, which can be regarded as a function of $(p_1, \ldots, p_t)$, is given by

$$d_t^* = \sum_{s=0}^{t-1} p_{t-s} c_{t-s} f_{s+1}, \qquad t = 1, \ldots, \tau. \tag{1}$$

Let $\boldsymbol{d}^* = (d_1^*, \ldots, d_\tau^*)^T$, $\boldsymbol{c} = (c_1, \ldots, c_\tau)^T$ and $\boldsymbol{p} = (p_1, \ldots, p_\tau)^T$. By considering the temporal structure of the data, $\boldsymbol{d}^*$ can be expressed in matrix form $\boldsymbol{Qp}$ where

$$\boldsymbol{Q} = \begin{pmatrix} f_1 & 0 & \cdots & 0 \\ f_2 & f_1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ f_\tau & \cdots & f_2 & f_1 \end{pmatrix} \begin{pmatrix} c_1 & 0 & \cdots & 0 \\ 0 & c_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & c_\tau \end{pmatrix}$$

is a deterministic matrix which can be constructed by $\boldsymbol{c}$ and F. Thus, the estimation of $\boldsymbol{p}$ can be simplified to a linear regression problem. By setting $p^* = p_1 = \cdots = p_\tau$ in (1), the above framework is reduced to that of the time-delay adjusted case fatality rate estimator studied in Nishiura et al. [18], denoted as p^* in this article. Given the strong serial correlation in the time series of p_t and the fact that typically only the trend of the time series is of interest, Qu et al. [6] proposed to apply the fused lasso technique [19, 20] to obtain a smoothed estimator of $\boldsymbol{p}$, namely the real-time delay-adjusted case fatality rate, henceforth rtaCFR$_t$ for $t = 1, \ldots, \tau$. Let $\boldsymbol{d} = (d_1, \ldots, d_\tau)^T$ and

$$\boldsymbol{D} = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & -1 \end{pmatrix}$$

be a $(\tau - 1) \times \tau$ penalty matrix. That is to solve the following minimization problem

$$\textbf{rtaCFR} = \widehat{\boldsymbol{p}} = \underset{p}{\arg\min} \frac{1}{2} \| \boldsymbol{d} - \boldsymbol{Qp} \|_2^2 + \lambda \| \boldsymbol{Dp} \|_1, \tag{2}$$

where $\textbf{rtaCFR} = (\text{rtaCFR}_1, \ldots, \text{rtaCFR}_\tau)^T$, $\lambda > 0$ is a trade-off parameter that weights the accuracy of the solution and its sparsity level. As λ increases, the adjacent coordinates in $\widehat{\boldsymbol{p}}$ are pooled towards each other, which results in smoother estimates of real-time fatality rates. Given the expression in (2), the solution path of the generalized lasso problem can be computed using the R package `genlasso` [21] where the residual sum of squares value associated with each pre-assigned λ value can be obtained. In real data analysis, we can first consider the candidates in the solution path which satisfies $0 \leq \widehat{p}_t \leq 1$ for all t, then choose $\lambda = \lambda^*$ that yields the smallest residual sum of squares.

Standardized real-time fatality rate estimator

The aforementioned rtaCFR estimator is derived based on the minimal data requirements, without considering the effects of explanatory variables such as age, sex and some healthcare burden indicators are not considered. It is a composite index that takes into account a basket of time-varying indistinguishable factors when interpreting trends of the fatality rate over time. In other words, the effects of various underlying factors are mixed together, and we cannot segregate the effect of one factor from that of others. Consequently, the effect of a factor of

interest may be masked by the effect of another less important confounding factor. For instance, it is widely accepted that the fatality rate of a COVID-19 patient increases with age. If the effectiveness of a certain public health policy is evaluated based on a composite index that does not account for age, the actual effect of a successful policy may be obscured or confounded by the increasing proportion of elderly individuals among the confirmed cases over the period of policy implementation, thus providing misleading signals to the policymakers. This raises the question of whether a standardized version of rtaCFR can be developed when additional information on a confounding variable is collected along with the minimal epidemiological data.

Suppose that an additional individual-level variable with J ($J > 1$) categories, denoted by X, is available throughout the observation period. Then, for $j = 1, \ldots, J$, the stratified epidemiological data may be represented by $\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(J)}$ and $\boldsymbol{d}^{(1)}, \ldots, \boldsymbol{d}^{(J)}$, where $\boldsymbol{c}^{(j)} = (c_1^{(j)}, \ldots, c_\tau^{(j)})^T$ and $\boldsymbol{d}^{(j)} = (d_1^{(j)}, \ldots, d_\tau^{(j)})^T$ denote the time series of confirmed cases and deaths in the jth category, respectively. Furthermore, let $\boldsymbol{p}^{(j)} = (p_1^{(j)}, \ldots, p_\tau^{(j)})^T$ be the fatality rates for group j. Based on the formulation in (2), a group-specific real-time fatality rate estimator, denoted as **rtaCFR**$^{(j)}$, of $\boldsymbol{p}^{(j)}$ can be obtained for the jth category based on the stratified data $\boldsymbol{c}^{(j)}$, $\boldsymbol{d}^{(j)}$ and F, for $j = 1, \ldots, J$. Thus, a standardized real-time delay-adjusted case fatality rate estimator is given by

$$\text{srtaCFR}_t = \sum_{j=1}^{J} q_t^{(j)} \text{rtaCFR}_t^{(j)}, \qquad \text{for } t = 1, \ldots, \tau, \qquad (3)$$

where $q_t^{(j)}$ denotes a deterministic, possibly time-dependent reference weight for confirmed cases belonging to the jth category at time t, such that $\sum_{j=1}^{J} q_t^{(j)} = 1$ for all t. Similar to (2), we define **srtaCFR** $= (\text{srtaCFR}_1, \ldots, \text{srtaCFR}_\tau)^T$. It follows that one can easily remove the temporal effects of X on the real-time fatality rates by assigning a time-independent sequence $(q^{(1)}, \ldots, q^{(J)})$ in (3). For instance, to control for the effects on fatality rates induced by a shift in the age distribution of the confirmed cases (where X represents the age group), one can assign $(q^{(1)}, \ldots, q^{(J)})$ to be the time-independent population age structure at time 0 as the reference distribution. In this case, srtaCFR pertains to the weighted average of the age-specific real-time fatality rates for the reference distribution, at which the infections are assumed to always occur randomly and uniformly over the population throughout the observation period.

Simulation studies

We investigate the empirical properties of the proposed srtaCFR estimator in contrast to the existing rtaCFR estimator under different hypothetical scenarios, to provide some insights into why the former metric would be more appropriate for assessing the severity of an infectious disease in real-time. Consider a scenario where minimal epidemiological data is available over a period of $\tau = 200$ intervals (e.g. in days). This data consists of the observed time series of c_t and d_t for $t = 1, \ldots, 200$, while the exact infection and death times of each individual are not obtainable. Furthermore, the collected time series data can be stratified based on a demographic variable X. This mimics the real-world situation of most epidemic outbreaks, where detailed individual-level data may not be readily accessible, as an infected subject may not be followed up regularly as in the clinical setups. However, demographic information such as age, gender or health conditions may still be accessible. To facilitate interpretation, we can consider treating X as the age variable, as age is a well-known influential factor in COVID-19 fatality. Specifically, the number of confirmed cases at each time point t can be stratified into $J = 3$ broad age groups, namely (1) young, (2) middle, and (3) old age groups.

Let $q_t^{*(1)}$, $q_t^{*(2)}$ and $q_t^{*(3)}$ be the underlying proportions of confirmed cases in young, middle and old age groups at time t, respectively, for $t = 1, \ldots, 200$. We consider four typical scenarios, where the temporal age distributions, namely $\{q_t^{*(1)}, q_t^{*(2)}, q_t^{*(3)}\}_{t=1,\ldots,200}$, and the age-specific fatality rates, namely $\{p_t^{(1)}, p_t^{(2)}, p_t^{(3)}\}_{t=1,\ldots,200}$ are depicted in Fig 1. Specifically, we set the proportion of the middle age group to be constant at $q_t^{*(2)} = 0.2$ for all t in all scenarios. In scenarios I and II, the proportion of elderly cases switches from $q_t^{*(3)} = 0.2$ to $q_t^{*(3)} = 0.6$ at $t = 100$. This shift mimics the real-world situation that the emergence of a novel variant tends to have a greater impact on more vulnerable populations. In scenarios III and IV, we assume a shift towards younger age groups by setting a linearly decreasing proportion of elderly cases, $q_t^{*(3)} = 0.6 - 0.002t$. This pertains to the situation where the disease is spreading more and more rapidly among younger individuals, possibly due to various factors, such as changes in population behaviour or increased susceptibility among younger age groups. Regarding the fatality rates, we set $p_t^{(1)} = p_t^{(2)} = p_t^{(3)} = 0.03$ in scenario I, and $p_t^{(1)} = 0.1, p_t^{(2)} = 0.02, p_t^{(3)} = 0.06$ in scenario II. In scenario III, the fatality rates increase exponentially over the entire period, with $p_t^{(1)} = 0.1 \exp(0.004t)$, $p_t^{(2)} = 0.02 \exp(0.004t)$, and $p_t^{(3)} = 0.06 \exp(0.004t)$. In scenario IV, the fatality rates are set to be $p_t^{(1)} = 0.1 \exp\{0.004(t - 100)_+\}, p_t^{(2)} = 0.02 \exp \{0.004(t - 100)_+\}, p_t^{(3)} = 0.06 \exp\{0.004(t - 100)_+\}$, where $a_+ = a$ if the constant $a > 0$, and $a_+ = 0$ otherwise. In summary, the age-specific fatality rates remain constant and identical in scenario I, constant but different in scenario II, while in scenario III, they increase exponentially throughout the observation period. In scenario IV, the fatality rates remain constant initially but begin to increase exponentially after $t > 100$, indicating the emergence of a more virulent mutant variant.
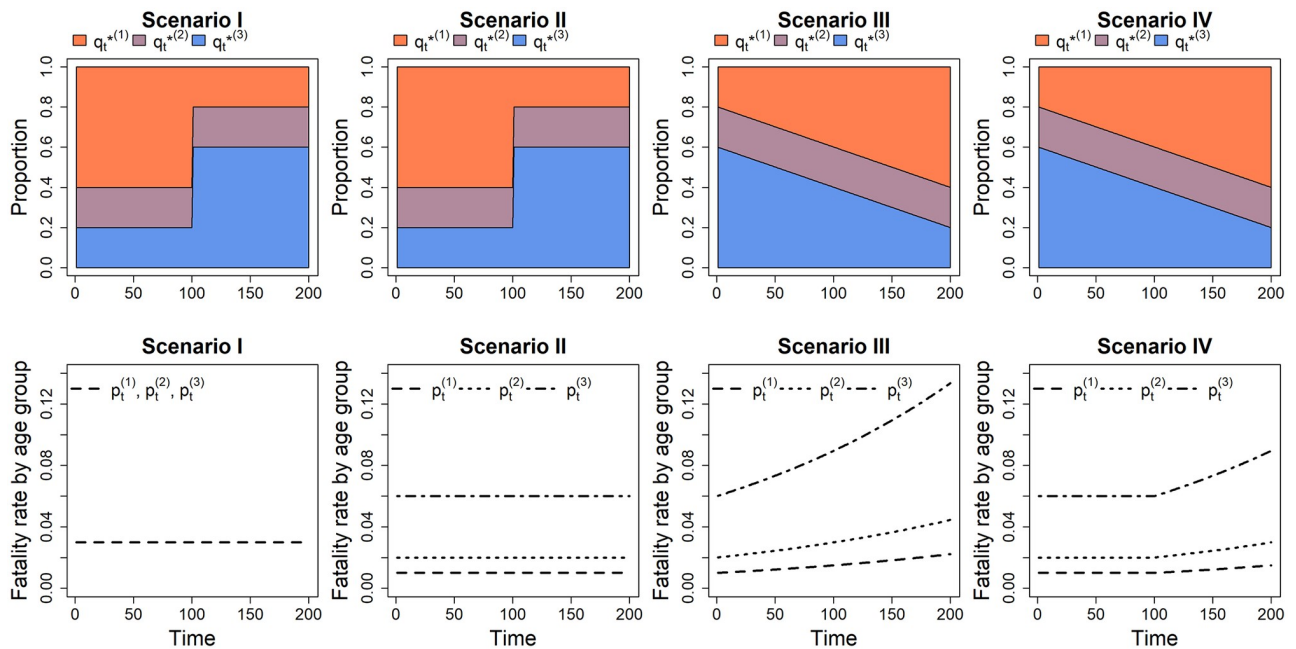
Fig 1. Simulation settings in scenarios I–IV. Top panel: temporal age distributions of confirmed cases, where $q_t^{*(1)}$, $q_t^{*(2)}$, and $q_t^{*(3)}$ represent the proportion of young, middle and old age groups, respectively. Bottom panel: patterns of the temporal age-specific fatality rates, where $p_t^{(1)}$, $p_t^{(2)}$ and $p_t^{(3)}$ represent the fatality rates of the young, middle and old age groups, respectively.

The following briefly describes the simulation setup. The non-stratified number of confirmed cases is set to $c_t = 10000 - 50 \cdot |100 - t|$ for $t = 1, \ldots, 200$. This mimics the real-world situation where a surge in confirmed cases is typically observed in a new wave of infection, followed by a decline resulting from the implementation of some effective control measures. Given c_t at time t, we then generate the age-stratified confirmed cases $(c_t^{(1)}, c_t^{(2)}, c_t^{(3)})$ based on the multinomial distribution with probabilities $(q_t^{*(1)}, q_t^{*(2)}, q_t^{*(3)})$ provided in the top panel of Fig 1. For a particular group j, the total number of deaths contributed by $c_t^{(j)}$ is generated based on the binomial distribution with probability $p_t^{(j)}$ provided in the bottom panel of Fig 1. The death figures will be reported (realized) in future time points ($t + t_{delay}$), where t_{delay} is distributed according to F to capture the delay between the onset of disease and death. We assume that F follows a gamma distribution with a mean of $\mu = 15.43$ days and shape parameter $\gamma = 2.03$, referring to the situation of COVID-19 outbreak estimated by recent studies [22]. Therefore, we obtain $\boldsymbol{d}^{(j)} = (d_1^{(j)}, \ldots, d_\tau^{(j)})^T$ by summing up the reported deaths at each time point t for groups $j = 1, 2, 3$. The total number of reported deaths at time t under this case is simply given by $d_t = \sum_{j=1}^J d_t^{(j)}$.

We consider 1000 replications of the above simulation procedure. For each replication, we compute the age-specific real-time fatality rate $\textbf{rtaCFR}^{(j)}$ for each of the three age groups ($j = 1, 2, 3$) based on the simulated age-stratified confirmed cases $\boldsymbol{c}^{(j)}$ and deaths $\boldsymbol{d}^{(j)}$. Then, we compute the proposed standardized estimator $\textbf{srtaCFR}$ given in (3) based on the time-invariant reference age distribution $(q^{(1)}, q^{(2)}, q^{(3)}) = (1/3, 1/3, 1/3)$, which pertains to the situation that infection occurs uniformly across different age groups. According to (3), $\textbf{srtaCFR}$ estimates the weighted averaged fatality rates $\bar{p}_t = \sum_{j=1}^3 q^{(j)} p_t^{(j)} = \sum_{j=1}^3 p_t^{(j)}/3$, allowing us to comprehensively capture the changing trend of the three pre-specified fatality rates $p_t^{(j)}$'s depicted in Fig 1. In contrast, the $\textbf{rtaCFR}$ is computed based on the total (non-stratified) cases $\boldsymbol{c}$ and deaths $\boldsymbol{d}$, which is a composite index measuring the overall real-time fatality rates $\tilde{\boldsymbol{p}} = (\tilde{p}_1, \ldots, \tilde{p}_\tau)$ where $\tilde{p}_t = \sum_{j=1}^3 q_t^{*(j)} p_t^{(j)}$ for $t = 1, \ldots, \tau$. The average estimates of $\textbf{rtaCFR}_t^{(j)}$, rtaCFR$_t$ and srtaCFR$_t$ over the 1000 replications, along with $\tilde{p}_t$ are provided in Fig 2.

As shown in the top panel of Fig 2, the estimated curves for the age-specific real-time fatality rates, $\textbf{rtaCFR}_t^{(j)}$, align closely with their respective true values across all scenarios, showing that the age-specific estimators are unbiased. We then evaluate the performance of the composite estimator, rtaCFR, and the standardized estimator, srtaCFR, across different scenarios. In scenario I where the age-specific fatality rates are constant and identical across all age groups, the temporal age distribution of confirmed cases has no effect on the estimates of disease fatality. We can see in the bottom panel of Fig 2 that rtaCFR$_t$ and srtaCFR$_t$ are almost identical and stay at the same level across the time axis in scenario I. However, in scenario II, where the age-specific fatality rates are constant with $p_t^{(1)} < p_t^{(2)} < p_t^{(3)}$, a sudden shift in the age distribution towards the elderly (i.e., a sudden increase in $q^{*(3)}$) at $t = 100$ will result in an increase in the overall fatality rate, $\tilde{p}_t$, even though the age-specific fatalities remain unchanged throughout the observation period. As expected, the composite estimator rtaCFR$_t$ aligns closely with $\tilde{p}_t$, indicating an increase in disease virulence. On the contrary, the proposed standardized fatality rate estimator srtaCFR$_t$ is totally unaffected by the shift in the age distribution and stays at a constant level because $p_t^{(1)}, p_t^{(2)}, p_t^{(3)}$ are constants, and the computation of estimator relies on a time-invariant reference age distribution.

In scenario III, where the age distribution gradually shifts towards the youth, and all the age-specific fatality rates are increasing exponentially, the rtaCFR$_t$ (and also $\tilde{p}_t$) is almost flat, as the increase in age-specific fatality rates is offset by the shift in the age distribution.
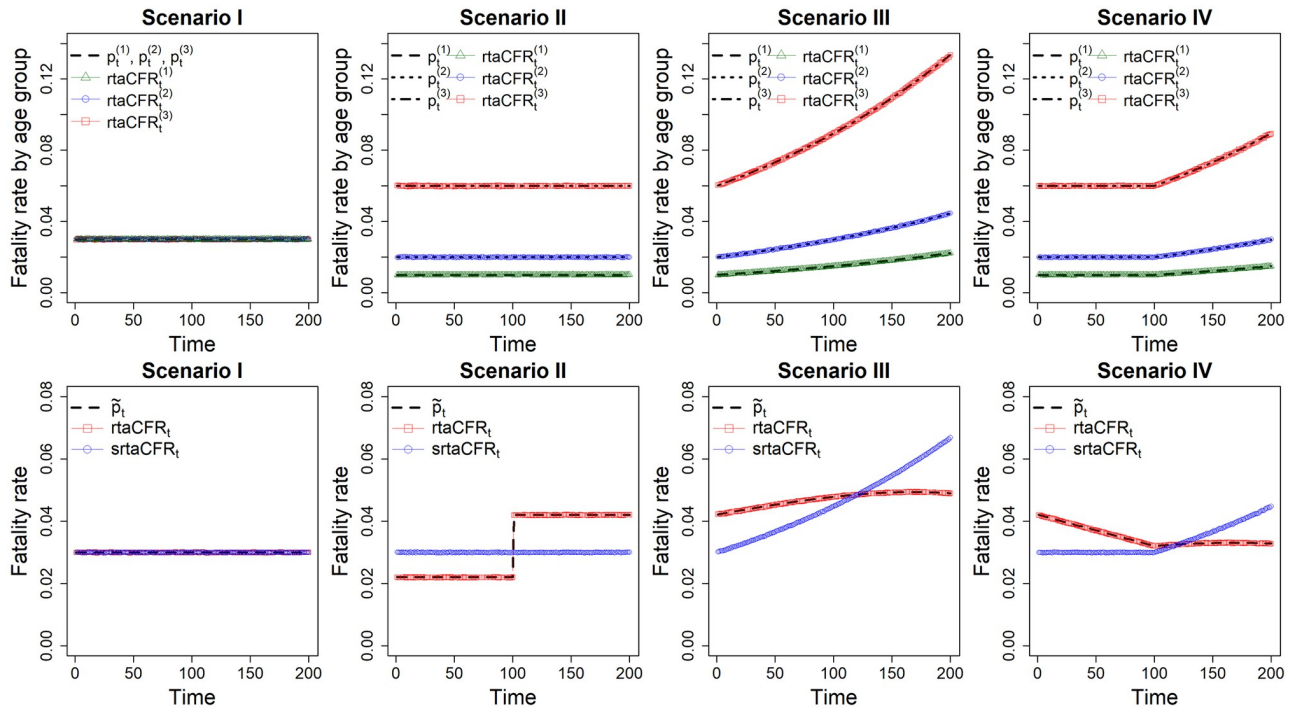
Fig 2. Simulation results under scenarios I–IV. Top panel: empirical averages of the age-specific real-time fatality estimates **rtaCFR**$_t^{(j)}$, $j = 1, 2, 3$, overlayed with the true age-specific fatality rates. Bottom panel: the underlying overall fatality rate $\tilde{p}$, and empirical averages of the estimates rtaCFR$_t$ and srtaCFR$_t$.

However, the standardized fatality rate estimator srtaCFR$_t$ is increasing exponentially over time, reflecting the actual increase in age-specific fatality rates. Scenario IV is similar to scenario III in terms of the shifting age distribution, but the exponential increase in age-specific fatalities occurs only after $t > 100$. Consequently, the effects of an increase in age-specific fatalities on the overall fatality rate, $\tilde{p}_t$, are offset or even surpassed by the effects of the skewed young confirmed cases. This leads to a conflicting phenomenon where the non-standardized fatality rate estimator, rtaCFR$_t$, is decreasing, but the standardized fatality rate estimator, srtaCFR$_t$, is increasing.

The simulation results highlight the importance of accounting for the age distribution when estimating real-time fatality rates. The impact of the temporal age distribution on the estimates of the fatality rate also depends on the underlying fatality rates for each age group. If the fatality rates are identical across age groups, as in scenario I, both rtaCFR and srtaCFR estimators perform well in capturing the disease virulence in a timely manner. However, it is important to note that this scenario is unlikely to occur in the context of infectious disease, as the fatality rate typically increases with age. In situations where the fatality rates differ across age groups, as seen in scenarios II to IV, the changing age distribution can have a significant impact on the estimation of disease fatality. As demonstrated in scenario II, a sudden shift in the age distribution towards the elderly can induce a false signal of increasing virulence of the disease. This is because the rtaCFR is a weighted average of the age-specific fatality rates, where the weights are determined by the age distribution of confirmed cases. When the age distribution shifts towards the elderly, the fatality rate will increase, even if the underlying age-specific fatality rates remain constant over time. We can see analogous results in scenarios III

and IV, where the increase in fatality rate is masked or even surpassed by the decreasing age of confirmed cases.

In real-world applications, policymakers are often interested in estimating the virulence of a disease or evaluating the effectiveness of a control measure, while the impact of changes in age distribution on overall fatality rates is typically considered a nuisance factor. Consequently, the composite index rtaCFR that includes a basket of unknown factors may not be desirable in fatality estimation provided that information on demographic variables, such as the patient's age, is available. In this simulation study, we provide a very basic example to illustrate that the proposed age-standardized fatality rate estimator can be more appropriate and accurate in reflecting the actual changes in disease fatality, by removing the confounding effects of the age variable. Therefore, the government can have a better evaluation of the true trajectory of the fatality rates promptly when informing public health policies. The real-world epidemiological data structure is far more complex than that used in our simulation settings. Presumably, we may also consider a standardized fatality rate estimator with data stratified on some important factors other than age, given the availability of those data.

The COVID-19 pandemic in Germany

The COVID-19 pandemic, caused by the novel coronavirus SARS-CoV-2, emerged in late 2019 in Wuhan, China [23]. Then it spread rapidly across the world, causing sustained outbreaks in early 2020. On March 11, 2020, the World Health Organization declared the COVID-19 outbreak a global pandemic [24]. By the end of March 2020, more than 750,000 cases with around 37,000 deaths have been reported in 200 countries and territories, with Europe and the Americas becoming the new epicentre [25]. Compared to other European nations, Germany was considered a role model that handled the COVID-19 pandemic relatively well during its early stages. The daily numbers of confirmed cases and deaths in Germany with age demographics were collected from a public demographic database [26]. Given that the fatality rate remained at a very low level since the second half of 2022 with the emergence of a milder Omicron subvariant, we aim to examine changes in disease fatality in Germany prior to this period. Fig 3 shows the seven-day moving averages of daily confirmed cases and deaths from March 2020 to May 2022.

The COVID-19 pandemic reached Germany in late February 2020, resulting in a notable surge in the number of cases and deaths. The government reacted quickly, and a range of measures including widespread testing, social distancing, school closure and travel bans have been introduced since early March [17, 27, 28]. Despite these efforts, the first wave of the pandemic hit Germany hard, particularly in the older population due to several outbreaks originating from nursing homes and religious events [28]. The infection reached its peak in early April 2020, with the country reporting over 100,000 cases and more than 2,000 deaths related to COVID-19. Owing to strict measures put in place, especially the nationwide lockdown introduced in late March 2020, Germany was able to successfully reduce the number of COVID-19 infections in the subsequent weeks [29]. By the summer of 2020, cases had significantly decreased, leading to a loosening of restrictions starting in May 2020. However, we can see from Fig 3 that Germany experienced a new surge in COVID-19 cases after the summer holidays. The increased travel demand and social gatherings during the holiday seasons led to a record-high infection within communities in the fall of 2020. In response, the government reintroduced stricter measures in November 2020, including social distancing, closing non-essential businesses and enforcing mandatory mask-wearing requirements in public places [30]. Concurrently, the vaccination campaign in Germany was initiated at the end of 2020
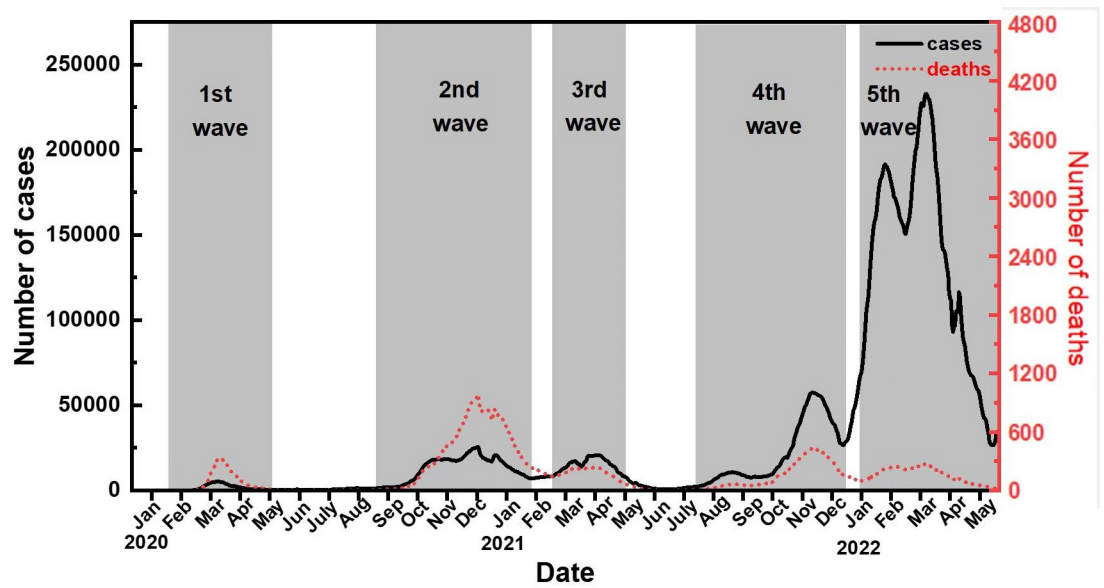
**Fig 3. The daily numbers of confirmed cases and deaths in Germany.**

with priority given to the residents of nursing homes and health workers who were at higher risk of severe illness or exposure to the virus [31].

While the situation improved towards the end of 2020, Germany experienced a third wave of the pandemic in early 2021 due to the emergence of the more transmissible Alpha variant, initially identified in the United Kingdom. In response, Germany implemented various restrictive measures and intensified its vaccination campaign, prioritizing older adults and individuals with underlying health conditions. By May 2021, there was an evident decline in the number of cases and hospitalizations since the peak in late March 2021. With infection rates remaining low and an increased vaccination rate among citizens, the government introduced a five-step plan in July 2021 to gradually lift restrictions on businesses and public life. This approach aimed to strike a balance between controlling the spread of the virus and promoting long-term economic growth. It reflected an adaptation to living with COVID after a period of stringent measures [32, 33].

However, the emergence of the Delta variant in the summer of 2021 led to a fourth wave of the pandemic. The relaxed "live with COVID" policy in July was believed to have contributed to the worsening situation. Towards the end of 2021, the already strained healthcare system faced considerable pressure with the rise of the highly contagious Omicron variant, resulting in a large number of infections [34]. As shown in Fig 3, the number of cases increased, with daily counts surpassing 200,000 by mid-February 2022. Notably, COVID-19-related deaths remained relatively low during the fifth wave, attributed to high vaccination rates among the population, particularly among the elderly. As of September 2022, over 90% of individuals aged 60 and older have been fully vaccinated [35].

We apply **rtaCFR** in (2) and **srtaCFR** in (3) to the COVID-19 data in Germany from March 2020 to May 2022, with an observation period of $\tau = 833$ days. The tunning parameter λ^* obtained in **rtaCFR** is 28654, while the λ^* values for **rtaCFR**$^{(1)}$, **rtaCFR**$^{(2)}$, and **rtaCFR**$^{(3)}$ are 9625, 21771 and 8220, respectively. We assume that F follows a gamma distribution with a mean of 15.43 days and a shape parameter of 2.03 [22]. To gain an insight into the trend of fatality rate over time, we apply a Gaussian kernel smoother to the fatality rate estimators with
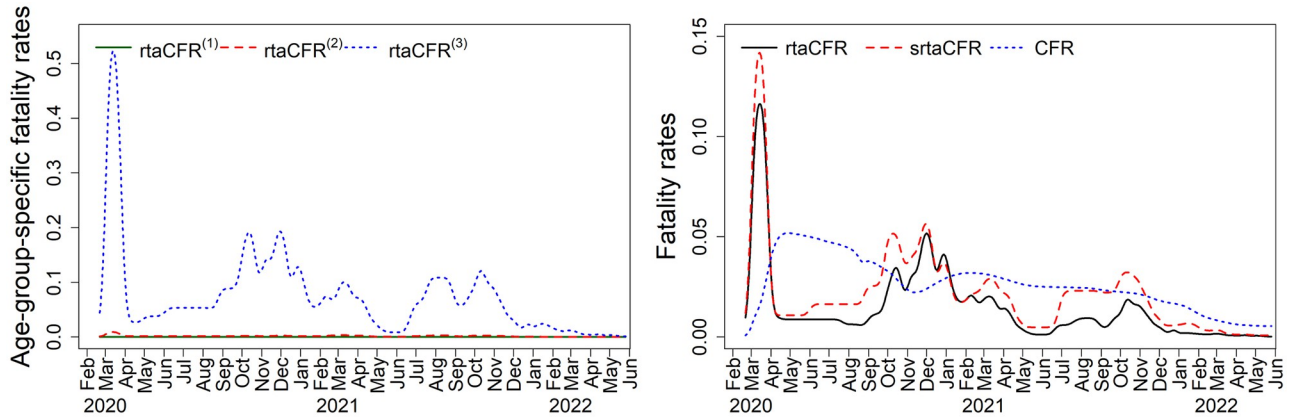
Fig 4. Estimation of the real-time fatality rates for the COVID-19 data in Germany. Left panel: Age-specific fatality rates for groups (1) $0-20$ years old, (2) $20-60$ years old, and (3) 60+ years old. Right panel: traditional CFR, rtaCFR$_t$ and srtaCFR$_t$.

bandwidth 15 days. The left panel of Fig 4 plots the smoothed age-specific fatality rates (i.e. **rtaCFR**$_t^{(j)}$ for $j = 1, 2, 3$) for ages 0-20, 20-60, and 60+ years. Meanwhile, the red dashed line in the right panel of Fig 4 shows the srtaCFR$_t$, which is the weighted average of **rtaCFR**$_t^{(j)}$ with time-independent weights $(q^{(1)}, q^{(2)}, q^{(3)}) = (0.20, 0.51, 0.29)$ corresponding to the population age distribution of Germany as of December 31, 2021. For comparison, we also included the rtaCFR$_t$ as a black solid line and the traditional CFR as a blue dotted line on the same graph. To investigate the effect of the penalty term λ^* in the fused lasso regression, Fig 5 presents a sensitivity analysis showing the smoothed rtaCFR$_t$ and srtaCFR$_t$ obtained using a range of different values for λ^*. As expected, when λ^* increases, a larger penalty is imposed and the neighbouring estimates are getting closer to each other across the time axis, yielding a smoother estimated curve for the real-time fatality rate. The sensitivity analysis demonstrates that the proposed method is robust against the changes in λ^*, as the estimated trends of the fatality rates are shown to be insensitive to the selection of λ^*.

As illustrated in Fig 4, the CFR increased significantly during the first wave of infection, reaching its peak in May 2020, and it dropped gradually to a low level in 2022 since the emergence of the Omicron variant. While the CFR provides a useful measure of overall disease
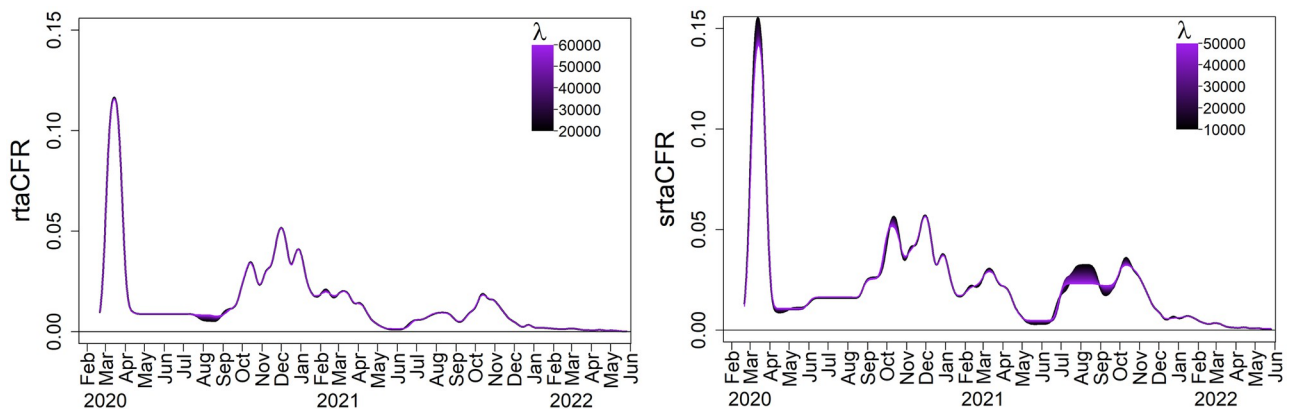

Fig 5. Sensitivity analysis.

severity, it is important to recognize that this measure has limitations in capturing the actual variability of the fatality rate in real-time, as it is computed based on cumulative epidemiological figures. As remarked in Qu et al. [6], the CFR is better served as a measure of the severity of a disease with a stable fatality rate spanning a short time frame. On the contrary, both rtaCFR and srtaCFR estimators are competent to capture the changes in disease fatality during the ongoing epidemic, with four peaks of fatality rates attained in April 2020, December 2020, April 2021, and November 2021, respectively. This finding aligns closely with the progression of the first four pandemic waves in Fig 3.

The estimated trends provided by $rtaCFR_t$ and $srtaCFR_t$ are largely similar within the observation period. Nevertheless, there was a noticeable difference between the $rtaCFR_t$ and $srtaCFR_t$ during the summer of 2020, with the latter being higher. This difference can be attributed to the fact that the age structure of confirmed cases during the summer of 2020 was different from that of the baseline period. The rtaCFR estimator does not account for changes in age distribution over time, while the srtaCFR estimator adjusts for such changes. The departure is evident based on Fig 6 that the proportion of confirmed cases among young people increased quickly during that period. This shift was possibly induced by the lifting of lockdown measures and the reopening of schools and universities, which led to more social activities among younger people. A similar shift towards young confirmed cases was seen in the summer of 2021 after the government had decided to implement the "live with COVID" policy. The observed pattern is analogous to scenario III in the simulation study, where the actual increasing trend of the fatalities is masked by a younger structure of confirmed cases when one considers rtaCFR.

The "live with COVID" strategy has been widely recognized as a great attempt to return to normalcy, as most of the first countries that adopted this approach in 2021 have observed a stabilization or even decline in COVID-related fatalities while simultaneously rebooting their
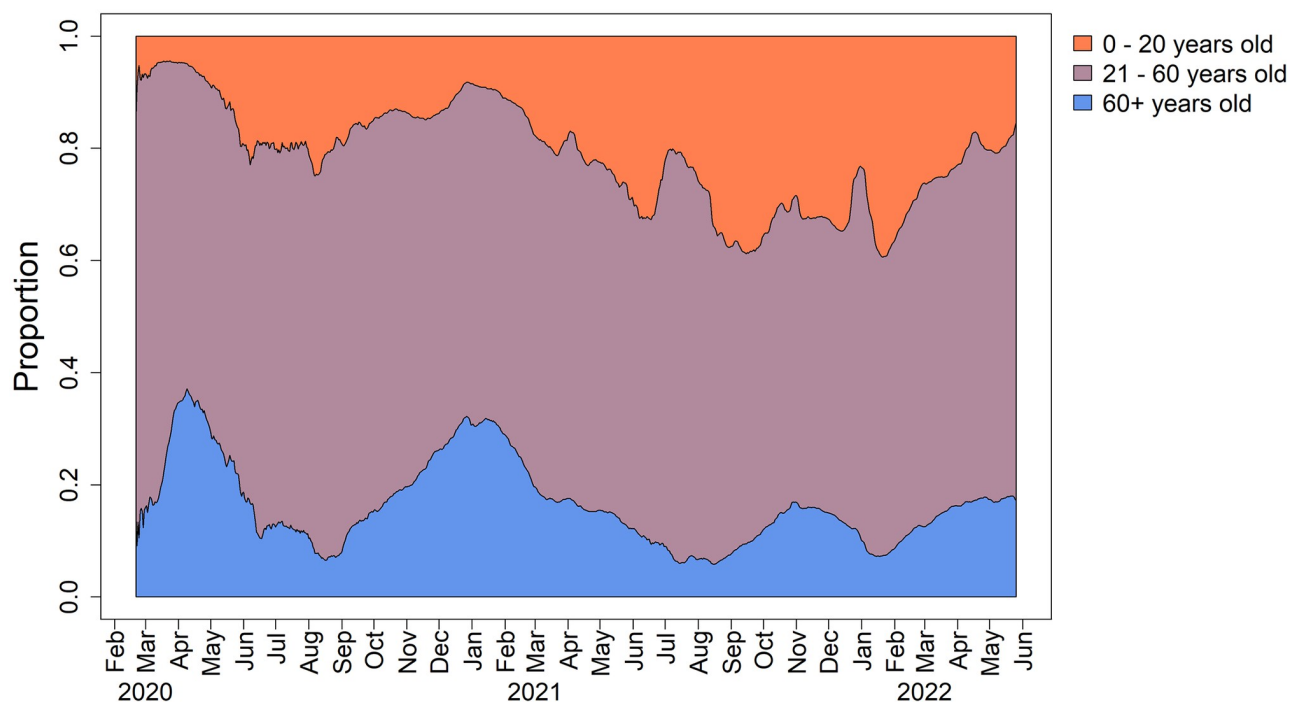
Fig 6. Temporal age distribution of the confirmed cases in Germany.

economies [36]. However, the success of this strategy in improving the case fatality rate may be an illusion due to bias in the estimation of CFR. As demonstrated by the right panel of Fig 4, the implementation of the "live with COVID" policy in June 2021 was associated with a spike in both rtaCFR and srtaCFR estimators. In particular, srtaCFR_t has risen to a value twice as much as rtaCFR_t in the second half of 2021. The observed false decreasing trend of the CFR after the implementation of the "live with COVID" policy suggests that the widely-used CFR estimator may not be an appropriate metric for informing policies. Relying on the CFR alone to assess the impact of public health measures may lead to a false sense of security and complacency among the population and policymakers. For example, if the CFR appears to be decreasing, policymakers may assume that the situation is improving and ease public health measures prematurely, potentially leading to a resurgence of the virus. It is also important to note that the discrepancy between the rtaCFR_t and srtaCFR_t depends on the changing age distribution of confirmed cases, as well as the pattern of the underlying fatality rates for each age group. Specifically, the greater the shift in the age structure of the confirmed cases, the higher the age-specific fatalities, and the greater the difference between these two measures. Due to the relatively weak virulence of the mutant virus in mid-2021, the difference between the two real-time estimates we observed is not very large. However, the difference between these two estimators would be more pronounced for infectious diseases with a high fatality. As we observed in scenario IV of the simulation study, they may even give the opposite conclusion for changes in disease fatality. Therefore, it is better to consider srtaCFR which accounts for the changing age distribution of confirmed cases to provide a more accurate and reliable picture of the disease's virulence for tracking the disease's impact on the population.

Discussion

Age is a crucial factor contributing to the likelihood of severe illness and death of a patient with COVID-19. This phenomenon is evident when we refer to the estimated age-specific fatality rates in Germany. The fatality for the elderly group is considerably higher than that of the middle and young age groups, with the latter ones being relatively flat and almost zero throughout the pandemic. Hence, the age distribution among confirmed cases can have a substantial impact on the estimation of disease fatality. It is important to account for the changing age distribution of cases, particularly when the government monitors disease fatality to inform public health policy. As demonstrated in the simulation study, a shift in the age structure of cases can potentially result in a spurious decline in the overall disease fatality, even if the age-specific fatalities are unchanged or increasing. The inaccurate evaluation of the fatality trend may lead to delayed or even misguided control measures. To address this issue, we propose a novel standardized real-time fatality rate estimator that corrects bias caused by reporting delay in the traditional CFR and, on the other hand, considers both the changing age structure of confirmed cases and the corresponding age-specific fatalities to provide an accurate estimate of the true disease fatality during an ongoing epidemic.

The application to German data demonstrates the satisfactory performance of the proposed method for its capability to capture the moving trends of the disease fatality, accurately revealing peaks of fatality during each of the four waves of infections. By accounting for the age distribution of confirmed cases, the proposed srtaCFR estimator is capable of reflecting the true increase in fatality rate after the implementation of the "live with COVID" policy in the summer of 2021, while the traditional CFR and the rtaCFR estimator suggest a decreasing or flat trend of disease fatality. It is noteworthy that the pattern of age-specific fatality of COVID-19 cases is rather simple in the application as the estimated trend of overall fatality is dominated by the estimated trend of fatality in the elderly group (aged above 60) since the

changes in the fatality rates for both young and middle age groups are negligible. The proposed method can also be applied to other diseases with more complicated patterns of age-specific fatality rates. A typical example is the Ebola virus disease, where the overall case fatality ranges from 30% to 90%, depending on the species of Ebola virus and the quality of medical care provided [37]. Unlike the SARS-CoV-2 coronavirus, all age-specific fatality rates are well above zero for the Ebola virus, with particularly high risk in children aged below 5 and adults aged above 45 [38]. Therefore, the age distribution of the confirmed cases can have a substantial influence on the interpretation of the overall fatality over time. In addition, there is a delay ranging from 6 to 16 days from symptom onset to death for Ebola infections [39]. Hence, the proposed srtaCFR is expected to work well for the estimation of the real-time fatality rate of the Ebola virus disease, adjusting for both reporting delay and structural change in age over time.

To better understand the impact of infectious diseases, it is important to recognize that the fatality rate is a multifaceted measure that can be influenced by factors beyond age. For instance, sex, race, and vaccination status can play important roles, as certain diseases can disproportionately affect patients with different combinations of these variables. Therefore, it is crucial to take these factors into account when assessing the impact of infectious diseases over time. Depending on the characteristics of the infectious disease of interest and the availability of information on relevant factors, the proposed method can be easily extended to accommodate the adjustment for various time-varying nuisance factors based on the multi-level stratified epidemiological data.

A potential avenue for future research is to account for the possible ascertainment bias in estimating the fatality rate. Specifically, the proposed srtaCFR estimator may encounter such bias if the fatality rate of individuals diagnosed with COVID-19 differs from that of individuals who are infected but remain undiagnosed. The extent of this bias is contingent upon the level of test coverage specific to each country. In countries with widespread testing, such as China and Germany, this bias is negligible. However, in countries with less comprehensive testing, such as Italy and the United States, the proposed method may tend to reflect the situation for patients with moderate to severe symptoms, potentially leading to an overestimation of the fatality rates. In these circumstances, additional data about the testing volume and other relevant information is required to estimate the ratio of under-reporting rate [40]. Therefore, there is an urgent need for standardized data collection by national health authorities. Furthermore, in periods of high infection numbers, it is also important to consider the age-specific testing patterns and their potential influence on the estimated real-time fatality rate. If testing capacities are primarily directed towards older individuals who are at a higher risk of severe outcomes, the observed fatality rate may be inflated. This bias should be taken into account when estimating the fatality rate during such periods.

An alternative research approach involves employing indirect standardization, which entails standardizing with the age distribution from a reference population instead of the population of interest adopted by direct standardization. Direct standardization used in our method is known for its high precision, while the indirect alternative proves useful when there is limited availability of detailed age distribution data, as it still yields valuable insights despite potential precision limitations [41, 42].

In light of the federal structure of Germany, it is important to acknowledge the potential spatial variations in epidemiological reporting. Incorporating spatial information into our inference or standardization methods can provide a more comprehensive understanding of the disease dynamics across different regions [43]. As demonstrated in many previous studies [41, 44, 45], such approaches allow us to account for spatial heterogeneity and better capture the nuances of the epidemiological situation. By considering spatial factors, we can enhance

the precision and reliability of our analyses, leading to more informed decision-making processes.

In summary, the proposed method provides a more accurate and reliable depiction of the disease's virulence compared to other existing fatality rate estimators. By monitoring the latest trend in the fatality rate, the continuous decision-making process can be supported during the epidemic. As many countries have chosen to coexist with the coronavirus, frequent examination of the fatality rate is of paramount importance as policymakers can be well-informed when there is a rebound in the fatality of the disease, such that restrictive measures can be implemented promptly.

Author Contributions

Conceptualization: Yuanke Qu, Chun Yin Lee.

Formal analysis: Chun Yin Lee.

Investigation: Yuanke Qu, Chun Yin Lee.

Methodology: Yuanke Qu, Chun Yin Lee.

Software: Chun Yin Lee.

Supervision: Chun Yin Lee.

Visualization: Yuanke Qu, Chun Yin Lee.

Writing – original draft: Yuanke Qu, Chun Yin Lee.

Writing – review & editing: Yuanke Qu, Chun Yin Lee.

References

1. The Guardian. US experts question whether counting COVID cases is still the right approach; 2022. https://www.theguardian.com/us-news/2022/jan/01/us-covid-case-counts-omicron.

2. U.S. News. COVID case counts may be losing importance amid Omicron; 2022. https://www.usnews.com/news/health-news/articles/2022-01-04/covid-case-counts-may-be-losing-importance-amid-omicron.

3. Mizumoto K, Kagaya K, Chowell G. Early epidemiological assessment of the transmission potential and virulence of coronavirus disease 2019 (COVID-19) in Wuhan City, China, January–February, 2020. BMC Medicine. 2020; 18(1):1–9. https://doi.org/10.1186/s12916-020-01691-x PMID: 32664866

4. Mizumoto K, Chowell G. Estimating risk for death from coronavirus disease, China, January–February 2020. Emerging Infectious Diseases. 2020; 26(6):1251. https://doi.org/10.3201/eid2606.200233 PMID: 32168464

5. Garske T, Legrand J, Donnelly CA, Ward H, Cauchemez S, Fraser C, et al. Assessing the severity of the novel influenza A/H1N1 pandemic. BMJ. 2009; 339. https://doi.org/10.1136/bmj.b2840 PMID: 19602714

6. Qu Y, Lee CY, Lam KF. A novel method to monitor COVID-19 fatality rate in real-time, a key metric to guide public health policy. Scientific Reports. 2022; 12(1):18277. https://doi.org/10.1038/s41598-022-23138-4 PMID: 36316534

7. Khalili M, Karamouzian M, Nasiri N, Javadi S, Mirzazadeh A, Sharifi H. Epidemiological characteristics of COVID-19: a systematic review and meta-analysis. Epidemiology & Infection. 2020; 148. https://doi.org/10.1017/S0950268820001430 PMID: 32594937

8. Levin AT, Hanage WP, Owusu-Boaitey N, Cochran KB, Walsh SP, Meyerowitz-Katz G. Assessing the age specificity of infection fatality rates for COVID-19: systematic review, meta-analysis, and public policy implications. European Journal of Epidemiology. 2020; 35(12):1123–1138. https://doi.org/10.1007/s10654-020-00698-1 PMID: 33289900

9. Pezzullo AM, Axfors C, Contopoulos-Ioannidis DG, Apostolatos A, Ioannidis JP. Age-stratified infection fatality rate of COVID-19 in the non-elderly population. Environmental Research. 2023; 216:114655. https://doi.org/10.1016/j.envres.2022.114655 PMID: 36341800

10. Dowd JB, Andriano L, Brazel DM, Rotondi V, Block P, Ding X, et al. Demographic science aids in understanding the spread and fatality rates of COVID-19. Proceedings of the National Academy of Sciences. 2020; 117(18):9696–9698. https://doi.org/10.1073/pnas.2004911117 PMID: 32300018

11. Sudharsanan N, Didzun O, Bärnighausen T, Geldsetzer P. The contribution of the age distribution of cases to COVID-19 case fatality across countries: a nine-country demographic study. Annals of Internal Medicine. 2020; 173(9):714–720. https://doi.org/10.7326/M20-2973 PMID: 32698605

12. Zhou C, Wheelock ÅM, Zhang C, Dong K, Li Z, Liang W, et al. The role of booster vaccination in decreasing COVID-19 age-adjusted case fatality rate: Evidence from 32 countries. Frontiers in Public Health. 2023; 11:1150095. https://doi.org/10.3389/fpubh.2023.1150095 PMID: 37143970

13. Samir KC, Moradhvaj D. Impact of the COVID-19 pandemic on the age-sex pattern of COVID-19 deaths in India. Asian Population Studies. 2023; 19(2):148–167. https://doi.org/10.1080/17441730.2023.2193077

14. Nguyen NTH, Ou TY, Huy LD, Chang YM, Huang CC. A global analysis of COVID-19 infection fatality rate and its associated factors during the Delta and Omicron variant periods: an ecological study. Frontiers in Public Health. 2023; 11:1145138. https://doi.org/10.3389/fpubh.2023.1145138

15. Jones RP, Ponomarenko A. COVID-19-related age profiles for SARS-CoV-2 variants in England and Wales and States of the USA (2020 to 2022): Impact on all-cause mortality. Infectious Disease Reports. 2023; 15(5):600–634. https://doi.org/10.3390/idr15050058 PMID: 37888139

16. Samaddar A, Gadepalli R, Nag VL, Misra S. The enigma of low COVID-19 fatality rate in India. Frontiers in Genetics. 2020; 11:854. https://doi.org/10.3389/fgene.2020.00854 PMID: 32849833

17. Staerk C, Wistuba T, Mayr A. Estimating effective infection fatality rates during the course of the COVID-19 pandemic in Germany. BMC Public Health. 2021; 21(1):1–9. https://doi.org/10.1186/s12889-021-11127-7 PMID: 34090392

18. Nishiura H, Klinkenberg D, Roberts M, Heesterbeek JA. Early epidemiological assessment of the virulence of emerging infectious diseases: a case study of an influenza pandemic. PLoS One. 2009; 4(8): e6852. https://doi.org/10.1371/journal.pone.0006852 PMID: 19718434

19. Tibshirani R, Saunders M, Rosset S, Zhu J, Knight K. Sparsity and smoothness via the fused lasso. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2005; 67(1):91–108. https://doi.org/10.1111/j.1467-9868.2005.00490.x

20. Tibshirani RJ, Taylor J. The solution path of the generalized lasso. The Annals of Statistics. 2011; 39 (3):1335–1371. https://doi.org/10.1214/11-AOS878

21. Arnold TB, Tibshirani R. Introduction to the genlasso package; 2017. https://cran.r-project.org/web/packages/genlasso/vignettes/article.pdf.

22. Diaz H, España G, Castañeda N, Rodriguez L, de la Hoz-Restrepo F. Dynamical characteristics of the COVID-19 epidemic: Estimation from cases in Colombia. International Journal of Infectious Diseases. 2021; 105:26–31. https://doi.org/10.1016/j.ijid.2021.01.053 PMID: 33529705

23. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel coronavirus from patients with pneumonia in China, 2019. New England Journal of Medicine. 2020; 382(8):727–733. https://doi.org/10.1056/NEJMoa2001017 PMID: 31978945

24. World Health Organization. WHO Director-General's opening remarks at the media briefing on COVID-19—11 March 2020; 2020. https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19—11-march-2020.

25. John Hopkins University Coronavirus Resource Center. Baltimore, MD: Bloomberg School of Public Health; 2021. https://coronavirus.jhu.edu.

26. Riffe T, Acosta E. Data resource profile: COVerAGE-DB: a global demographic database of COVID-19 cases and deaths. International Journal of Epidemiology. 2021; 50(2):390–390f. https://doi.org/10.1093/ije/dyab027

27. Deutscheland. The timeline regarding coronavirus in Germany; 2021. https://www.deutschland.de/zh-hans/the-timeline-corona-virus-germany.

28. Wieler LH, Rexroth U, Gottschalk R. Emerging COVID-19 success story: Germany's push to maintain progress. Our World in Data. 2021.

29. Robert Koch-Institut. COVID-19 cases and deaths in Germany as of February 2022, by state.; 2023. https://www.statista.com/statistics/1100823/coronavirus-cases-development-germany.

30. The Federal Ministry of Health. The fight against the coronavirus; 2021. https://www.bundesgesundheitsministerium.de/en/press/2020/coronavirus.html.

31. Robert Koch Institute. Recommendation of the Standing Committee on Vaccination (STIKO) at the Robert Koch Institute—4th update: COVID-19 vaccination with Comirnaty® of BioNTech / Pfizer.; 2020. https://www.rki.de/DE/Content/Kommissionen/STIKO/Empfehlungen/COVID-19/index.html.

32. BBC News. Germany eases Covid restrictions as infection rate declines; 2021. https://www.bbc.com/news/world-europe-58377538.

33. The Guardian. Germany to ease COVID restrictions as infections fall; 2021. https://www.theguardian.com/world/2021/jun/06/germany-to-ease-covid-restrictions-as-infections-fall.

34. Wikipedia contributors. COVID-19 pandemic in Germany; 2021. https://en.wikipedia.org/wiki/COVID-19_pandemic_in_Germany#Timeline_by_state.

35. Robert Koch Institute. COVID-19 vaccination coverage in Germany: September 2022. https://www.rki.de/DE/Content/Infografiken/COVID-19/COVID19_Verlauf_2.png?__blob=publicationFile.

36. The Hankyoreh. Among nations 'living with COVID-19,' only S. Korea has seen upswing in fatality rate; 2021. https://english.hani.co.kr/arti/english_edition/e_national/1022339.html

37. Beeching NJ, Fenech M, Houlihan CF. Ebola virus disease. BMJ. 2014; 349. https://doi.org/10.1136/bmj.g7348

38. Agua-Agum J, Ariyarajah A, Blake IM, Cori A, Donnelly CA, Dorigatti I, et al. Ebola virus disease among children in West Africa. New England Journal of Medicine. 2015; 372(13):1274–1277. https://doi.org/10.1056/NEJMc1415318 PMID: 25806936

39. Goeijenbier M, Van Kampen J, Reusken C, Koopmans M, Van Gorp E. Ebola virus disease: a review on epidemiology, symptoms, treatment and pathogenesis. The Netherlands Journal of Medicine. 2014; 72 (9):442–448. PMID: 25387613

40. Lipsitch M, Donnelly CA, Fraser C, Blake IM, Cori A, Dorigatti I, et al. Potential biases in estimating absolute and relative case-fatality risks during outbreaks. PLoS Neglected Tropical Diseases. 2015; 9 (7):e0003846. https://doi.org/10.1371/journal.pntd.0003846 PMID: 26181387

41. Souris M, Gonzalez JP. COVID-19: Spatial analysis of hospital case-fatality rate in France. PLoS One. 2020; 15(12):e0243606. https://doi.org/10.1371/journal.pone.0243606 PMID: 33320895

42. Souris M, Gonzalez JP. COVID-19: Spatial analysis of lethality in France during the first epidemic inception. Revue francophone sur la santé et les territoires. 2021;.

43. Qu Y, Lee CY, Lam KF. A sequential test to compare the real-time fatality rates of a disease among multiple groups with an application to COVID-19 data. Statistical Methods in Medical Research, 31(2), 348–360. https://doi.org/10.1177/09622802211061927 PMID: 34878362

44. Gao F, Kihal W, Le Meur N, Souris M, Deguen S. Does the edge effect impact on the measure of spatial accessibility to healthcare providers? International Journal of Health Geographics. 2017; 16:1–16. https://doi.org/10.1186/s12942-017-0119-3 PMID: 29228961

45. Souris M, Marcombe S, Laforet J, Brey PT, Corbel V, Overgaard HJ. Modeling spatial variation in risk of presence and insecticide resistance for malaria vectors in Laos. PLoS One. 2017; 12(5):e0177274. https://doi.org/10.1371/journal.pone.0177274 PMID: 28494013