Adjusted mortality by age associated with SARS-CoV-2 infection during the early stages of the epidemics in Hubei, China and northern Italy

Anthony Hauser

Michel J. Counotte

Ben Bales

Charles C. Margossian

Garyfallos Konstantinoudis

Nicola Low

Christian L. Althaus

Julien Riou

The epidemic of SARS-CoV-2 that originated in Wuhan, China is now pandemic. We estimated the age-specific mortality by fitting a transmission model to data from China, accounting for preferential ascertainment of severe cases and right censoring of deaths. Overall mortality among all infections was 3.0% (2.6-3.4) and 3.3% (95%CrI: 2.0-4.7) in Hubei and northern Italy, respectively, and increased considerably for the elderly, highlighting the expected burden for populations with further expansion of the SARS-CoV-2 epidemic around the globe.

# Introduction

As of 16 March 2020, the novel coronavirus (SARS-CoV-2) pandemic that originated in Wuhan, Hubei, China, has resulted in more than 150,000 confirmed cases and 5,735 deaths globally (World Health Organization 2020b). The transmission characteristics of SARS-CoV-2 appear to be similar to those of pandemic influenza and will likely facilitate further global spread (Riou and Althaus 2020a). Large-scale autonomous transmission already exist in South Korea, Japan, Iran, Italy and several European countries (World Health Organization 2020c). During this early phase of the pandemic, it is critically important to obtain reliable estimates of the overall mortality, i.e. the proportion of all (asymptomatic and symptomatic) infected cases that will die as a result of the disease. Such estimates will help anticipate the expected morbidity and mortality due to SARS-CoV-2 infection and provide critical information for the planning of health care systems in countries that face an epidemic. Estimates of mortality among symptomatic patients seeking care is also of prime importance for clinicians to assess the prognosis and, in dire situations of overwhelmed health system, prioritize patients with better expected outcomes.

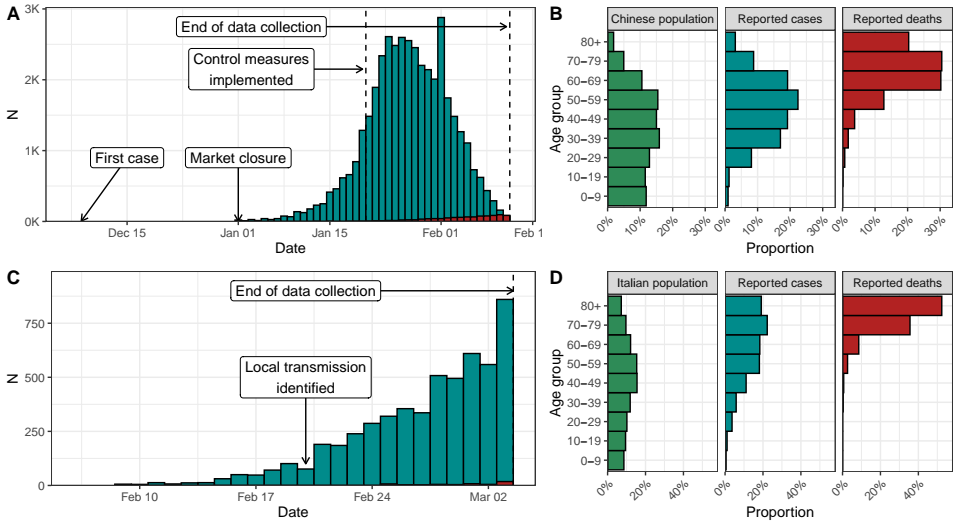
Obtaining reliable estimates of mortality can be challenging during the early phase of an epidemic. The crude case fatality ratio (CFR), defined as the number of reported deaths divided by the number of reported cases until a specific time point, is a widely used indicator (World Health Organization 2020d; C. Wang et al. 2020). Other reported measures such as confirmed case fatality ratio or symptomatic case fatality ratio, based on different denominators, are sometimes used, adding confusion. While the crude CFR can be a useful measure of the general scale of an ongoing epidemic, it can be misleading if used to assess the overall mortality because of two opposing biases (Lipsitch et al. 2015; Battegay et al. 2020). First, because of the delay between disease onset and death, the number of reported deaths until a certain time point does not consider the total number of deaths that will occur among already infected individuals (right-censoring). Second, the number of reported cases underestimates the total number of patients affected by the disease, because surveillance systems focus on symptomatic individuals, and among symptomatic cases on patients with more severe manifestations (preferential ascertainment). While it is often assumed that the crude CFR overestimates true mortality (World Health Organization 2020a), there is no indication that this will be the case, as it depends on the relative influence of the two opposing biases.

Without a longitudinal study of a representative sample of individuals with SARS-CoV-2 infection, estimates of mortality will rely on surveillance data and require statistical adjustment. In this analysis, we focus on two hotspots of the ongoing pandemic: the Hubei province of China and northern Italy. During the early stages of these epidemics, the crude CFR was 2.4% in Hubei (979 deaths over 41,092 cases on 11 February 2020) and 1.3% in northern Italy (79 deaths over 6,117 cases on 3 March 2020). Using surveillance data and several assumptions regarding the natural history of SARS-CoV-2, we simulate the dynamics of transmission, care and mortality of SARS-CoV-2 by age group with the objective of obtaining corrected estimates of (1) mortality among SARS-CoV-2 infections with symptoms and (2) mortality among all SARS-CoV-2 infections. -

# Methods

## Data from Hubei, China

The SARS-CoV-2 epidemic in the Hubei province of China appears to originate from multiple zoonotic transmission events in the main city Wuhan in early December 2019, with the animal source remaining unknown (World Health Organization-China Joint Mission on Coronavirus Disease 2019 Group 2020). Early January 2020, a novel coronavirus (subsequently named SARS-CoV-2) was identified as the causal agent of the epidemic (Zhou et al. 2020). In the first phase of the epidemic, human-to-human transmission occurred at a high rate in Wuhan and other areas of Hubei, leading to exponential growth of incidence (Figure [fig:desc]A). On 20 January, Chinese authorities implemented strict control measures in Hubei, including early identification and isolation of cases, contact tracing, extension of holidays, temperature checks before accessing public areas, cancellation of mass gatherings and the promotion of extreme social distancing (World Health Organization-China Joint Mission on Coronavirus Disease 2019 Group 2020). Three days later, a cordon sanitaire was imposed, with strict traffic restrictions. From 27 January, the daily incidence of cases by disease onset started plateauing, then decreased. The number of reported deaths was delayed compared to the incidence (Figure [fig:desc]C). We extracted data from the Chinese Center for Disease Control (CDC) report on the epidemiological characteristics of cases and deaths reported until 11 February 2020, which provided the number of reported cases by date of disease onset and the age distribution of cases and deaths reported up to this point in China, that we applied to Hubei (Figure [fig:desc]B). We also used data about the age distribution of the Chinese population, and about the daily number of potentially infectious contacts by age group in Shanghai (Zhang et al. 2019), which we assumed were similar to the population of Hubei. Age-specific prevalences of diabetes, chronic respiratory diseases, cardiovascular diseases and hypertension in China were extracted from the IHME website (Institute for Health Metrics and Evaluation 2020).



(A) Reported confirmed cases of SARS-CoV-2 infection in Hubei by date of disease onset (blue) until 11 February 2020. (B) Age distribution of the Chinese population compared to that of reported cases of and deaths due to SARS-CoV-2 infection. (C) Reported number of deaths in Hubei until 11 February 2020. (D) Daily number of potentially infectious contacts across age classes in Shanghai.

## Data from northern Italy

Italy reported its first case of SARS-CoV-2 infection on 30 January, but the first case of local transmission was identified in 20 February in Lombardy. In the following days, new cases were reported in Lombardy and the local authorities put the whole area under strict quarantine. Despite this, Italy faced an exponential growth of new SARS-CoV-2 infection cases. This led the government to place the whole country on lockdown from 9 March, restricting movement of the population and banning public events. As of 16 March, Italy has reported 25,000 cases and 1,800 deaths due to SARS-CoV-2 infection. We collected data from reports from the Instituto Superiore di Sanita and the Dipartimento della Protezione Civile, which reported the number of new cases and deaths per day by date of disease onset from 8 February to 3 March, and their distribution by age group (Civile 2020; Istituto Superiore di Sanita 2020). We also collected data about the age distribution of the Italian population, as well as its contact patterns by age group using results from the POLYMOD study (Mossong et al. 2008).

# An age-structured model of SARS-CoV-2 transmission and mortality

We used an age-stratified susceptible-exposed-infected-removed (SEIR) compartmental model, with a distinction between incubating, asymptomatic and symptomatic infections. We stratified the population by 10-year ranges, leading to 9 age classes (0-9 years old, ..., 80 years old and more). We assumed that the risk of transmission per contact is identical across all age classes. We used age-specific contact matrices for each country to model the contact patterns across the 9 age classes within the SEIR model as defined in the previous section. In addition, we modelled the decrease in the transmission of SARS-CoV-2 due to the progressive implementation of control measures from 20 February by using a logistic function for the transmission rate.

After an average incubation period of 5.9 days (Bi et al. 2020), 82.1% (95%CrI: 79.8-84.5) of infected people develop mild or severe symptoms and become infectious, while the remaining remain asymptomatic and do not transmit the disease further. Data from passengers of the “Diamond Princess” was used to inform the proportion of symptomatic infections, which was implemented as a beta distribution to propagate uncertainty (Mizumoto et al. 2020). This data did not provide any conclusive evidence of an age trend in the proportion of true asymptomatics, so we assumed it constant across age groups (Japanese National Institute of Infectious Diseases 2020). The mean time from disease onset to isolation was fixed to 2.4 days (Bi et al. 2020).

An SEIR model was used to compute the number of symptomatic and asymptomatic infections by day of disease onset in each age class. We then used an age-specific ascertainment proportion to estimate the reported cumulative number of symptomatic infections. In order to identify the parameters, we assumed that all infected patients aged 80 and more were reported. We assumed that mortality only occurred in symptomatic people, and that the time from disease onset to death followed a log-normal distribution with mean 20.2 days and standard deviation 11.6 (Linton et al. 2020). This allowed us to account for the deaths occurring after the date of data collection (11 February for Hubei, 3 March for northern Italy).

We simultaneously fitted our model to four data sets: (1) the number of confirmed cases by day of disease onset, (2) the number of deaths by day of occurrence, (3) the age distribution of all confirmed cases and (4) the age distribution of all reported deaths, separately for Hubei and northern Italy. All parameters were estimated from data except for the incubation period, the proportion of symptoms, the time from disease onset to isolation and the time from disease onset to death. The fitted model was used to produce estimates of the total number of symptomatic and asymptomatic infections (corrected for preferential ascertainment) and of the total number of deaths (corrected for right-censoring), which were then transformed into adjusted estimates of mortality. Besides parameter values and model structure, these estimates rely on the following assumptions:

1. The risk of infection upon contact is the same across age groups;
2. The probability of symptoms upon infection is the same across age groups, and asymptomatic infected individuals don’t transmit;
3. The severity of symptoms differs by age group and influences the probability of reporting;
4. All deaths due to SARS-CoV-2 infection have been identified and reported;
5. All patients aged 80 and older that have symptoms are identified and reported;
6. The average standard of care is stable between the period of interest (until 11 February for Hubei, until 3 March for northern Italy) and the following period of two months during which a proportion of the infected people will eventually die.

We implemented the model in a Bayesian framework using Stan (Carpenter et al. 2017). All code and data are available from . Further details about the method are available in the Supplementary Appendix.

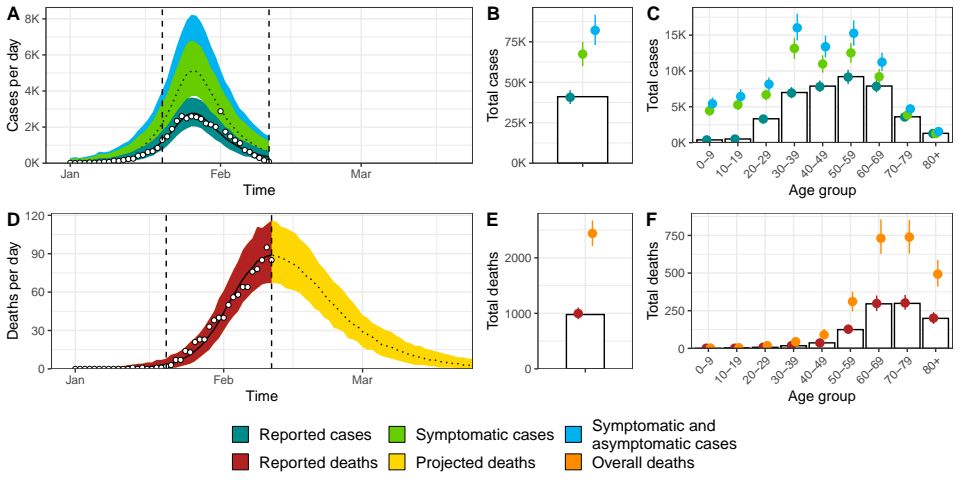
# Results

## Hubei, China

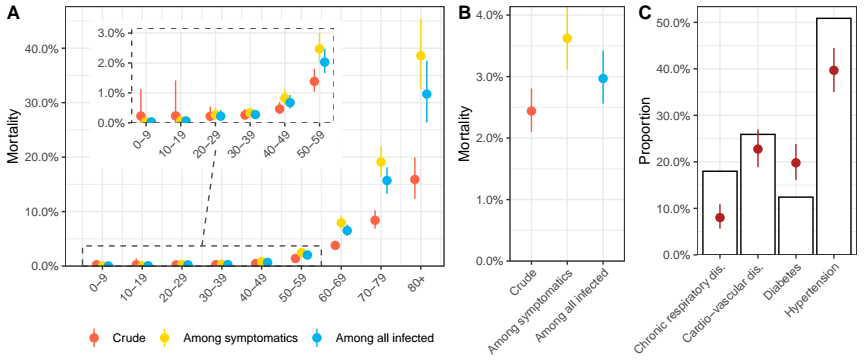
Our model accurately describes the dynamics of transmission and mortality by age group during the SARS-CoV-2 epidemic in Hubei from 1 January to 11 February 2020 (Figure [fig:fit]). While mortality data reported after 11 February includes deaths among people infected after this date, contrary to the model predictions, later reports regarding the number of deaths align with our prediction, providing external validation. Control measures implemented from 20 January led to a reduction of transmissibility by 99% (95% credible interval [CrI]: 97-100), with a steep diminution in case incidence after six days. Under the assumption that the risk of transmission of SARS-CoV-2 upon contact is homogeneous by age, so that the age distribution of reported cases can be entirely explained by the age structure in the population and the patterns of contacts by age, the total number of symptomatic cases during that period is estimated to 67,500 (95%CrI: 59,900-74,800), 1.6 times (95%CrI: 1.5-1.8) more than the 41,092 reported cases during that period. The ascertainment proportion showed a clear age trend, from less than 10% under 20 years old to 93% (95%CrI: 87-99) in the age group 70-79 (it was assumed that the ascertainment proportion was 100% in the age group 80+, supplementary table 5). Under the assumption that 82% (95%CrI: 79-85) of infections lead to symptoms, a total of 82,300 individuals (95%CrI: 73,000-91,800) were infected in Hubei between 1 January and 11 February 2020.

As of 11 February 2020, 979 deaths had been reported in Hubei. Under our assumption regarding the distribution of the delay between disease onset and death, the model predicts a total of 2,439 deaths (95%CrI: 2,210-2,673) among all people infected until 11 February. This results in an adjusted mortality of 3.6% (95%CrI: 3.1-4.1) among infected individuals with symptoms, and of 3.0% (95%CrI: 2.6-3.4) among all infected individuals (Figure [fig:mortality]B). This adjustment leads to sensible modifications of the age-specific mortality (Figure [fig:mortality]A). Compared to the crude CFR, the adjusted mortality among all individuals infected with SARS-CoV-2 is lower or similar in the younger age classes (0-49 years old) but higher in people aged 50 and more, reaching 32.0% (95%CrI: 25.3-40.1) among individuals aged 80 years and more.

Estimates of mortality among infected individuals with symptoms, used by clinicians to assess the prognosis of patients seeking care, follow a similar pattern. Under 20 years of age, mortality among symptomatics is estimated below 1 in 1,000 and rises to between 3 to 8 per 1,000 symptomatic infections for individuals aged 20 to 49. Mortality among symptomatics is estimated to 2.5% (95%CrI: 1.9-3.1) among individuals aged 50-59, 8.0% (95%CrI: 6.6-9.5) among individuals aged 60-69, 19.2% (95%CrI: 15.8-22.9) among individuals aged 70-79 and reaches 39.0% (95%CrI: 31.1-48.9) among individuals aged 80 and more. We also examined the links between mortality and four comorbidities associated with older age (Figure [fig:mortality]C). Compared to the expected prevalence given the age-distribution of deaths associated with SARS-CoV-2 infection, data shows an excess of diabetes and a lack of chronic respiratory disease or hypertension among deaths recorded in Hubei.



Model fit for Hubei, China of (A) incident cases of SARS-CoV-2 infection by date of disease onset, (B) total cases, (C) age distribution of cases, (D) incidence of deaths, (E) total deaths and (F) age distribution of deaths. White circles and bars represent data. Lines and shaded areas or points and ranges show the posterior median and 95% credible intervals for six types of model output: reported cases, symptomatic cases, overall cases (i.e. symptomatic and asymptomatic cases), reported deaths until 11 February 2020, projected deaths after 11 February 2020 and overall deaths.

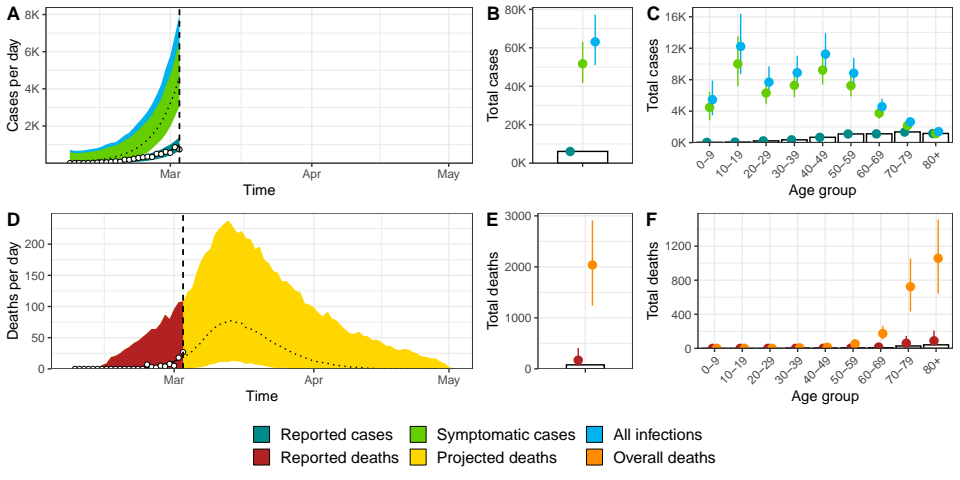


(A-B) Estimates of mortality during the first wave of the SARS-CoV-2 epidemic in Hubei, China by age group and overall. (C) Observed prevalence four comorbidities among deaths associated with SARS-CoV-2 infection in Hubei (purple) compared to the expected prevalence given the age distribution of deaths and the age-specific prevalence of each comorbidity in the Chinese population (white bar). Points and ranges show the posterior median and 95% credible intervals.

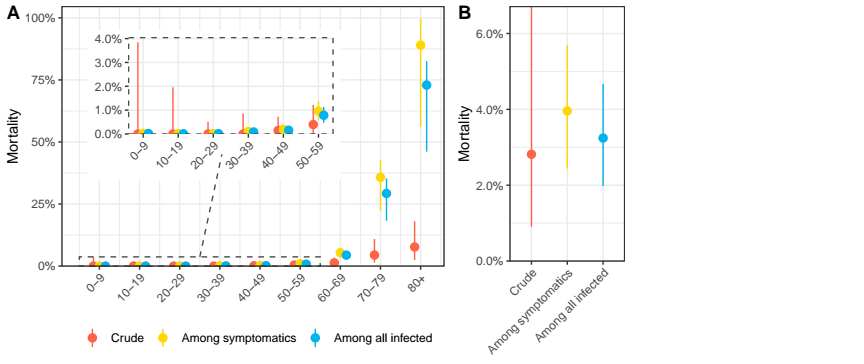
## Northern Italy

The model is also able to describe the early dynamics of transmission and mortality by age group during the SARS-CoV-2 epidemic in northern Italy from to 2 February to 3 March 2020 (Figure [fig:fitit]). Given the age distribution of cases, and assuming that all symptomatic cases among individuals aged 80 and more have been detected, we predict that 51,900 (95%CrI: 41,800-63,100) symptomatic cases occurred during that period, 8.5 (95%CrI: 6.8-10.3) times more than the 6,117 reported cases. The ascertainment proportion was estimated to 1% under 20 years old, and rose up to 62% (95%CrI: 58-67) in the 70-79 age group (it was assumed that the ascertainment proportion was 100% in the 80+ age group, supplementary table 8). Under our assumptions regarding the proportion of symptomatic infections, this implies a total number of 63,300 (95%CrI: 51,000-77,200) people infected with SARS-CoV-2 in the area until 3 March 2020.

Among these people, we estimate that 2,053 (95%CrI: 1,238-2,910) will die, of which only 79 had been reported on March 3. This translates into an adjusted mortality of 4.0% (95%CrI: 2.4-5.7) among infected individuals with symptoms and of 3.3% (95%CrI: 2.0-4.7) among all people infected with SARS-CoV-2 (Figure [fig:mortalityit]B). We observe an even sharper age trend in mortality in northern Italy compared to Hubei, China, with an estimated mortality of 1.0% (95%CrI: 0.6-1.4) in symptomatic individuals aged 50-59, 5.4% (95%CrI: 3.4-6.9) in symptomatic individuals aged 60-69, 35.7% (95%CrI: 22.4-42.7) in symptomatic individuals aged 70-79 and 89.0% (95%CrI: 56.2-99.6) in symptomatic individuals aged 80 and older.



Model fit for northern Italy of (A) incident cases of SARS-CoV-2 infection by date of disease onset, (B) total cases, (C) age distribution of cases, (D) incidence of deaths, (E) total deaths and (F) age distribution of deaths. Coloured dots and white bars represent data. Lines and shaded areas or points and ranges show the posterior median and 95% credible intervals for five types of model output: reported cases, symptomatic cases, overall cases (i.e. symptomatic and asymptomatic cases), reported deaths until 3 March 2020, and overall deaths including these that will occur after this date.



(A-B) Estimates of mortality during the first wave of the SARS-CoV-2 epidemic in northern Italy by age group and overall. Points and ranges show the posterior median and 95% credible intervals.

# Discussion

In this work, we propose a comprehensive solution to the estimation of mortality from surveillance data during outbreaks (Lipsitch et al. 2015), and apply it to data from the SARS-CoV-2 epidemic in Hubei, China until 11 February 2020 and from the epidemic in northern Italy until 3 March 2020. After correcting for right-censoring and preferential ascertainment, we estimate that 3.0% (2.6-3.4) and 3.3% (95%CrI: 2.0-4.7) of all individuals infected during that period died in Hubei and northern Italy, respectively. This compares to crude CFR values of 2.4% in Hubei and 1.3% in northern Italy over the same periods. This illustrate that crude CFR values are not necessarily overestimates of true mortality, as the effect of right-censoring can be massive, especially at early stages of the epidemic.

These estimates among all infections are of interest to epidemiologists when assessing the potential consequences of the SARS-CoV-2 pandemic (Hethcote 2000). The direct application of these estimates to other countries must be done with caution, as the standard of care and, as a result, mortality are certainly time- and setting-dependent. However, observing such similar estimates in very different settings may be an indication of some degree of generalizability. As a counter-example, the epidemic in South Korea seems very different, with very few deaths reported so far (KCDC 2020). In this country, the disproportionate importance of the Shincheonji Church of Jesus cluster, associated with a large excess of young people among cases (Shim et al. 2020), prevented us from applying the same model based on the age distributions in the general population. The higher number of younger people among cases in South Korea, associated with wide testing and a better management of the epidemic might explain the fewer numbers of reported fatalities. Data from other countries, in particular the number of cases by date of disease onset and the age distribution of cases and deaths, are necessary to better understand the variability in mortality across settings.

We confirm a strong age trend in mortality, with a sharp increase in mortality from 50 years old, reaching very high values in people aged 80 and older: 39.0% (95%CrI: 31.1-48.9) and 89.5% (95%CrI: 62.0-99.6) in northern Italy. This difference can be intuitively understood by comparing the age distributions of cases and deaths in both settings, with an even more marked age shift in Italy (Figure [fig:desc]B and D). Estimates of mortality by age among individual infected with SARS-CoV-2 are central to clinicians to assess prognosis and prioritize their efforts in dire situations of overwhelmed healthcare systems as have been reported at some stages of the SARS-CoV-2 epidemics in Hubei and northern Italy. The specific causes of this age trend are unknown, but early discussions have focused on the associations between SARS-CoV-2 and comorbidities such as diabetes and hypertension and the role of ACE inhibitors (Fang, Karakiulakis, and Roth 2020). By comparing the prevalence of four comorbidities (diabetes, chronic respiratory disease, cardio-vascular disease and hypertension) among deaths associated with SARS-CoV-2 infection in China and the expected prevalence of these comorbidities according to the age distribution of deaths, we find that only diabetes are in excess among SARS-CoV-2 deaths, and that the prevalence of the other comorbidities is similar or lower than expected in a Chinese population with that age distribution. This ecological observation does not refute any causality between these comorbidities and SARS-CoV-2-related mortality, as the age trend itself may be related to a higher prevalence of ageing-associated diseases, but highlights that the very specific age pattern of mortality associated with SARS-CoV-2 infection must be accounted for when discussing association with comorbidities.

Our estimate of mortality among all infected in Hubei is higher than the one found in a first version of this work (3.0% against 1.6%) (Riou and Althaus 2020b). The implementation of age-specific risk of transmission through a contact matrix partly explains this difference, as the contact patterns lead to a lower number of infections among the elderly (for the same number of death), thus increasing mortality. The use of a higher proportion of symptomatic people of 82%, coming from two new separate studies (Mizumoto et al. 2020; Bi et al. 2020), rather than 49% as previous data suggested, also led to higher estimates of mortality among all infected. Our estimate of mortality is higher than 1.38% in mainland China reported in a recent preprint (Verity et al. 2020). In this paper, Verity et al. used a very similar approach to ours, but different modelling choices can explain the lower results: they considered all mainland China, rather than limiting to Hubei (mortality appears to be lower outside Hubei (Team 2020)); they assumed a homogeneous attack rate across age groups rather than a contact matrix; and they used a lower reporting rate for the elderly (70% against 100%). Other analyses attempted to correct for right-censoring of deaths, focusing on data from China or on international travellers. A study using a competing risk model reported an estimate of mortality among cases of 7.2% (95% confidence interval: 6.6%-8.0%) for Hubei province using a competing risk model (X. Wang et al. 2020). Using data on exported cases and correcting for right-censoring of deaths occurred in China, another team reported an estimate of mortality of 5.3% (95% confidence interval: 3.5%, 7.5%) among confirmed cases (Jung et al. 2020). Another team reported a CFR of 18% (95% credible interval: 11-81%) among cases detected in Hubei, accounting for the delay in mortality (Dorigatti et al. 2020). The same study provided adjusted estimates of the overall CFR based on data from the early epidemic in Hubei and from cases reported outside China at 1% (95% CI: 0.5%-4%).

Our work has three important strengths. (1) We use a mechanistic model for the transmission of and the mortality associated with SARS-CoV-2 infection that is a direct translation of the data-generating mechanisms leading to the biased observations of the number of deaths (because of right-censoring) and of cases (because of preferential ascertainment). Our model also accounts for the effect of control measures on disease transmission. We implemented the model in a Bayesian framework and attempted to propagate all sources of uncertainty. (2) Our model is stratified by age group, which has been shown as a crucial feature for modelling emerging respiratory infections (Pellis et al. 2020). (3) The estimates rely on routinely collected surveillance data such as incident cases by disease onset, incidence deaths, and the age distribution of cases and deaths, and does not require individual-level data nor studies in the general population. As more data about the SARS-CoV-2 pandemic becomes available, our model could be applied to different settings.

Our work also has several limitations. First, our results depend on the central assumption that the cause of the deficit of reported cases among younger age groups is due to preferential ascertainment and does not reflect a lower risk of infection in younger individuals. The cause of this age shift in reported cases remains unclear. Uneven age distributions in the risk of infection can be attributed to immunological features, such as the lower circulation of H1N1 influenza in older individuals due to residual immunity (Pérez-Trallero et al. 2009). An immunological explanation of the opposite phenomenon, with a lower susceptibility of younger individuals, seems unlikely, and there is no indication of pre-existing immunity to SARS-CoV-2 in humans (World Health Organization-China Joint Mission on Coronavirus Disease 2019 Group 2020). Different contact patterns could play a role in creating different attack rates by age group, but we included age-specific contact patterns in the model. The last explanation that we assume here is that younger individuals, when symptomatic, have milder symptoms that decrease the probability of seeking care and being identified.

Second, our results depend on the assumption that older individuals that have more severe symptoms are very likely to be identified (ascertainment proportion of 100% for people aged 80 and more with symptoms). In the absence of an outside reference point, the reporting rate cannot be estimated from surveillance data only. The information that older people are much more at risk of dying from SARS-CoV-2 infection circulated very early on, and we believe this can justify the assumption that most old people with symptoms sought care. If further data, coming from a study in the general population, shows that this assumption is wrong, this would imply that mortality in our study is overestimated. Sensitivity analyses with a lower value of ascertainment proportion for the people aged 80 and more show the linear relation between and the final results (supplementary appendix, section 5).

Third, there is uncertainty around the proportion of asymptomatic infections. Currently, the detection of asymptomatic patients is limited by the focus on symptomatic patients seeking care and the lack of population seroprevalence data (Carrat et al. 2008). Two independent studies have used alternative designs to obtain similar estimates. During the outbreak on the ship “Diamond Princess”, nearly all individuals were tested regardless of symptoms, leading to an average proportion of symptomatic infections of 82.1% (95%CrI: 79.8-84.5) of infected people develop symptoms and become infectious (Mizumoto et al. 2020). Another study focused on 87 contacts of infected cases in Shenzhen, China, and estimated the proportion of symptomatics to 80.4% (95%CrI: 70.9-87.4) (Bi et al. 2020). Still, uncertainty about the proportion of symptomatic infections will remain until a large retrospective seroprevalence study is conducted in the general population, and our results are dependent on this estimate. Additionally, the dichotomization of infection into asymptomatic and symptomatic is a simplification of reality; the infection with SARS-CoV-2, will likely cause a gradient of symptoms in different individuals depending on age, sex and comorbidities. The proportion of asymptomatic infections might also show an age-dependent structure.

Fourth, our findings regarding mortality associated with SARS-CoV-2 infection are specific to the context, and should be interpreted in that light. The findings describe the situation in Hubei before 11 February and in Northern Italy before 3 March 2020. It was demonstrated there, that mortality rates have changed over time as a result of an improvement of the standard of care (World Health Organization-China Joint Mission on Coronavirus Disease 2019 Group 2020). The standard of care and, as a result, the mortality is time- and setting-dependent and cannot be directly applied to other contexts.

# Conclusions

We developed a mechanistic approach to correct the crude CFR for bias due to right-censoring and preferential ascertainment and provide adjusted estimates of mortality due to SARS-CoV-2 infection by age group. In Hubei, China, we estimate that 3.0% (2.6-3.4) of individuals infected until 11 February with or without symptoms died or will die before or after this date, with even more important differences by age group than suggested by the raw data. In northern Italy between 8 February and 3 March, we estimate mortality to 3.3% (95%CrI: 2.0-4.7) of all individuals infected with SARS-CoV-2, with an ever sharper age effect. In both settings, the probability of death among infected individuals with symptoms shows a steep increase for people aged 60 years and older, reaching extremely high values in people aged 80 years and older. While specific to the situation in Hubei, China and northern Italy during these periods, these findings will help the mitigation efforts and planning of resources as other regions prepare for SARS-CoV-2 epidemics.

# Acknowledgements

We

# Funding

JR is funded by the Swiss National Science Foundation (grant 174281). MC is funded by the Swiss National Science Foundation (grant 176233).

# Conflict of interest

None.

# Authors’ contributions

AH and JR designed the study. AH, MC, BB, CM and JR implemented the model and performed the statistical analyses. AH, MC, BB, CM, GK, NL, CA and JR interpreted the results and wrote the manuscript.

Battegay, Manuel, Richard Kuehl, Sarah Tschudin-Sutter, Hans H Hirsch, Andreas F Widmer, and Richard A Neher. 2020. “2019-novel Coronavirus (2019-nCoV): estimating the case fatality rate–a word of caution.” *Swiss Medical Weekly* 150 (0506).

Bi, Qifang, Yongsheng Wu, Shujiang Mei, Chenfei Ye, Xuan Zou, Zhen Zhang, Xiaojian Liu, et al. 2020. “Epidemiology and Transmission of COVID-19 in Shenzhen China: Analysis of 391 cases and 1,286 of their close contacts.” *medRxiv*, March. Cold Spring Harbor Laboratory Press, 2020.03.03.20028423. doi:[10.1101/2020.03.03.20028423](https://doi.org/10.1101/2020.03.03.20028423).

Carpenter, Bob, Andrew Gelman, Matthew D Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. 2017. “Stan: A probabilistic programming language.” *Journal of Statistical Software* 76 (1).

Carrat, Fabrice, Elisabeta Vergu, Neil M Ferguson, Magali Lemaitre, Simon Cauchemez, Steve Leach, and Alain-Jacques Valleron. 2008. “Time Lines of Infection and Disease in Human Influenza: A Review of Volunteer Challenge Studies.” *American Journal of Epidemiology* 167 (7). Oxford University Press: 775–85.

Civile, Presidenza del Consiglio dei Ministri - Dipartimento della Protezione. 2020. “GitHub repository on COVID-19.” Accessed March 16. <https://github.com/pcm-dpc>.

Dorigatti, Ilaria, Lucy Okell, Anne Cori, Natsuko Imai, Marc Baguelin, and Sangeeta Bhatia. 2020. “Severity of 2019 novel coronavirus (nCoV).” <https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial- College-2019-nCoV-severity-10-02-2020.pdf>.

Fang, Lei, George Karakiulakis, and Michael Roth. 2020. “Are Patients with Hypertension and Diabetes Mellitus at Increased Risk for Covid-19 Infection?” *The Lancet Respiratory Medicine*. Elsevier.

Hethcote, Herbert W. 2000. “The Mathematics of Infectious Diseases.” *SIAM Review* 42 (4). SIAM: 599–653.

Institute for Health Metrics and Evaluation. 2020. “Global Health Data Exchange.” <http://ghdx.healthdata.org/ihme_data>.

Istituto Superiore di Sanita. 2020. “COVID-19 integrated surveillance: the main national data.” Accessed March 16. <https://www.epicentro.iss.it/coronavirus/sars-cov-2-sorveglianza-dati>.

Japanese National Institute of Infectious Diseases. 2020. “Field Briefing: Diamond Princess COVID-19 Cases.” <https://www.niid.go.jp/niid/en/2019-ncov-e/9407-covid-dp-fe-01.html>.

Jung, Sung-mok, Andrei R Akhmetzhanov, Katsuma Hayashi, Natalie M Linton, Yichi Yang, Baoyin Yuan, Tetsuro Kobayashi, Ryo Kinoshita, and Hiroshi Nishiura. 2020. “Real-Time Estimation of the Risk of Death from Novel Coronavirus (Covid-19) Infection: Inference Using Exported Cases.” *Journal of Clinical Medicine* 9 (2). Multidisciplinary Digital Publishing Institute: 523.

KCDC. 2020. “Press Release KCDC: 2020-3-19.” Accessed March 19. <https://www.cdc.go.kr/board/board.es?mid=a30402000000{\&}bid=0030>.

Linton, Natalie M, Tetsuro Kobayashi, Yichi Yang, Katsuma Hayashi, Andrei R Akhmetzhanov, Sung-mok Jung, Baoyin Yuan, Ryo Kinoshita, and Hiroshi Nishiura. 2020. “Incubation Period and Other Epidemiological Characteristics of 2019 Novel Coronavirus Infections with Right Truncation: A Statistical Analysis of Publicly Available Case Data.” *Journal of Clinical Medicine* 9 (2). Multidisciplinary Digital Publishing Institute: 538.

Lipsitch, Marc, Christl A Donnelly, Christophe Fraser, Isobel M Blake, Anne Cori, Ilaria Dorigatti, Neil M Ferguson, Tini Garske, Harriet L Mills, and Steven Riley. 2015. “Potential biases in estimating absolute and relative case-fatality risks during outbreaks.” *PLoS Neglected Tropical Diseases* 9 (7).

Mizumoto, Kenji, Katsushi Kagaya, Alexander Zarebski, and Gerardo Chowell. 2020. “Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020.” *Eurosurveillance* 25 (10). doi:[https://doi.org/10.2807/1560-7917.ES.2020.25.10.2000180](https://doi.org/https://doi.org/10.2807/1560-7917.ES.2020.25.10.2000180).

Mossong, Joël, Niel Hens, Mark Jit, Philippe Beutels, Kari Auranen, Rafael Mikolajczyk, Marco Massari, et al. 2008. “Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases.” *PLoS Medicine* 5 (3). Public Library of Science.

Pellis, Lorenzo, Simon Cauchemez, Neil M Ferguson, and Christophe Fraser. 2020. “Systematic Selection Between Age and Household Structure for Models Aimed at Emerging Epidemic Predictions.” *Nature Communications* 11 (1). Nature Publishing Group: 1–11.

Pérez-Trallero, E, L Piñeiro, D Vicente, M Montes, and G Cilla. 2009. “Residual Immunity in Older People Against the Influenza a (H1n1)–Recent Experience in Northern Spain.” *Eurosurveillance* 14 (39). European Centre for Disease Prevention; Control: 19344.

Riou, Julien, and Christian L Althaus. 2020a. “Pattern of Early Human-to-Human Transmission of Wuhan 2019 Novel Coronavirus (2019-nCoV), December 2019 to January 2020.” *Eurosurveillance* 25 (4). European Centre for Disease Prevention; Control.

———. 2020b. “Pattern of early human-to-human transmission of Wuhan 2019 novel coronavirus (2019-nCoV), December 2019 to January 2020.” *Eurosurveillance* 25 (4).

Shim, Eunha, Amna Tariq, Wongyeong Choi, Yiseul Lee, and Gerardo Chowell. 2020. “Transmission Potential of Covid-19 in South Korea.” *medRxiv*. Cold Spring Harbor Laboratory Press.

Team, The Novel Coronavirus Pneumonia Emergency Response Epidemiology. 2020. “Vital Surveillances: The Epidemiological Characteristics of an Outbreak of 2019 Novel Coronavirus Diseases (COVID-19) — China, 2020.” *China CDC Weekly* 2 (8): 113–22. <http://weekly.chinacdc.cn/en/article/id/e53946e2-c6c4-41e9-9a9b-fea8db1a8f51>.

Verity, Robert, Lucy C Okell, Ilaria Dorigatti, Peter Winskill, Charles Whittaker, Natsuko Imai, Gina Cuomo-Dannenburg, et al. 2020. “Estimates of the severity of COVID-19 disease.” *medRxiv*, March. Cold Spring Harbor Laboratory Press, 2020.03.09.20033357. doi:[10.1101/2020.03.09.20033357](https://doi.org/10.1101/2020.03.09.20033357).

Wang, Chen, Peter W Horby, Frederick G Hayden, and George F Gao. 2020. “A Novel Coronavirus Outbreak of Global Health Concern.” *The Lancet* 395 (10223). Elsevier: 470–73.

Wang, Xing, Zihui Ma, Yi Ning, Chen Chen, Rujin Chen, Qiwen Chen, Heng Zhang, et al. 2020. “Estimating the Case Fatality Ratio of the Covid-19 Epidemic in China.” *medRxiv*. Cold Spring Harbor Laboratory Press.

World Health Organization. 2020a. “Coronavirus disease 2019 (COVID-19) Situation Report 46.” <https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200306-sitrep-46-covid-19.pdf?sfvrsn=96b04adf_2>.

———. 2020b. “Coronavirus disease 2019 (COVID-19) Situation Report 55.” <https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200315-sitrep-55-covid-19.pdf?sfvrsn=33daa5cb_6>.

———. 2020c. “Coronavirus disease 2019 (COVID-19) Situation Report 55.” <https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200315-sitrep-55-covid-19.pdf?sfvrsn=33daa5cb_8>.

———. 2020d. “WHO Director-General’s opening remarks at the media briefing on COVID-19 - 3 March 2020.” <https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---3-march-2020>.

World Health Organization-China Joint Mission on Coronavirus Disease 2019 Group. 2020. “Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19).” <https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf>.

Zhang, Juanjuan, Petra Klepac, Jonathan M Read, Alicia Rosello, Xiling Wang, Shengjie Lai, Meng Li, et al. 2019. “Patterns of Human Social Contact and Contact with Animals in Shanghai, China.” *Scientific Reports* 9 (1). Nature Publishing Group: 1–11.

Zhou, Peng, Xing-Lou Yang, Xian-Guang Wang, Ben Hu, Lei Zhang, Wei Zhang, Hao-Rui Si, et al. 2020. “A Pneumonia Outbreak Associated with a New Coronavirus of Probable Bat Origin.” *Nature*. Nature Publishing Group, 1–4.