Global perspective of COVID-19 epidemiology for a full-cycle pandemic

John P. A. Ioannidis

Departments of Medicine, of Epidemiology and Population Health, of Biomedical Data Science, and of Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, CA, USA

Correspondence
John P.A. Ioannidis, Departments of Medicine, of Epidemiology and Population Health, of Biomedical Data Science, and of Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, CA, USA. Email: jioannid@stanford.edu

Funding information
Laura and John Arnold Foundation

Abstract

As of October 2020, there are >1 million documented deaths with COVID-19. Excess deaths can be caused by both COVID-19 and the measures taken. COVID-19 shows extremely strong risk stratification across age, socioeconomic factors, and clinical factors. Calculation of years-of-life-lost from COVID-19 is methodologically challenging and can yield misleading over-estimates. Many early deaths may have been due to suboptimal management, malfunctioning health systems, hydroxychloroquine, sending COVID-19 patients to nursing homes, and nosocomial infections; such deaths are partially avoidable moving forward. About 10% of the global population may be infected by October 2020. Global infection fatality rate is 0.15-0.20% (0.03-0.04% in those <70 years), with large variability across locations with different age-structure, institutionalization rates, socioeconomic inequalities, population-level clinical risk profile, public health measures, and health care. There is debate on whether at least 60% of the global population must be infected for herd immunity, or, conversely, mixing heterogeneity and pre-existing cross-immunity may allow substantially lower thresholds. Simulations are presented with a total of 1.58-8.76 million COVID-19 deaths over 5-years (1/2020-12/2024) globally (0.5-2.9% of total global deaths). The most favorable figures in that range would be feasible if high risk groups can be preferentially protected with lower infection rates than the remaining population. Death toll may also be further affected by potential availability of effective vaccines and treatments, optimal management and measures taken, COVID-19 interplay with influenza and other health problems, reinfection potential, and any chronic COVID-19 consequences. Targeted, precise management of the pandemic and avoiding past mistakes would help minimize mortality.

KEYWORDS
COVID-19, epidemiology, infection fatality rate, mortality, risk factors

1 | INTRODUCTION

Almost a year since the first known cases in November 2019 in Wuhan, COVID-19 has been linked to over 1 million deaths and active epidemic waves continue to spread. It is important to overview the emerging epidemiological footprint and understand the current situation and its implications for the future of the pandemic. It is unknown exactly how long a full cycle for the pandemic spreading worldwide may take, and this is likely to vary across different locations. Regardless,
insights from the first year may help optimize responses to this ongoing crisis.

2 NUMBER OF DEATHS

Official COVID-19 deaths reached 1 million in late September 2020. Marked differences in overall mortality rates exist across countries and locations. As of early October 2020, 66 countries recorded <1 death per 100 000 population (including 21 mostly small countries without any deaths), while 17 countries exceeded 50 deaths per 100 000. These stark differences are mostly genuine, reflecting massive variability in viral spread, substantial variability in infection fatality rate (IFR), and both under- and overcounting of deaths across locations. Limited testing still leaves some COVID-19 deaths undocumented. Conversely, many countries may count some spurious COVID-19 deaths. Death certificates are notoriously error-prone in general\(^1\) and may be even more error-prone with COVID-19. Adherence to stringent clinical case definitions plus imaging/pathology documentation for SARS-CoV-2 causal impact is often lacking.\(^2\) In high-income countries, almost all the deceased have known comorbidities, raising causality debates on whether some deaths are with rather than by COVID-19.\(^3\) Deaths in people without documented comorbidities are more frequent in low- and mid-income countries,\(^4\) but perhaps comorbidities remained undetected in resource-poor settings.

<table>
<thead>
<tr>
<th>Cause of excess death</th>
<th>Reason/comments</th>
<th>Possible time horizon for excess deaths</th>
</tr>
</thead>
<tbody>
<tr>
<td>People with AMI and other acute disease not given proper hospital care</td>
<td>Patients afraid to go to hospital and hospitals reducing admissions afraid of overload</td>
<td>Acute, during pandemic</td>
</tr>
<tr>
<td>People with cancer having delayed treatment</td>
<td>Postponement of cancer treatment in anticipation of COVID-19 overload</td>
<td>Next 5 y</td>
</tr>
<tr>
<td>Disrupted cancer prevention</td>
<td>Inability to offer cancer prevention services under aggressive measures</td>
<td>Next 20 y</td>
</tr>
<tr>
<td>Other healthcare disruption</td>
<td>Postponement or cancellation of elective procedures and regular care</td>
<td>Variable for different medical conditions</td>
</tr>
<tr>
<td>Suicides</td>
<td>Mental health disruption</td>
<td>Both acute and long-term</td>
</tr>
<tr>
<td>Violence (domestic, homicide)</td>
<td>Mental health disruption</td>
<td>Acute, possibly long-term</td>
</tr>
<tr>
<td>Starvation</td>
<td>Disruption in food production and transport</td>
<td>Acute, and possibly worse over next several years</td>
</tr>
<tr>
<td>Tuberculosis</td>
<td>Disruption of tuberculosis management programmes</td>
<td>Next 5 y</td>
</tr>
<tr>
<td>Childhood diseases</td>
<td>Disruption of vaccination programmes</td>
<td>Next 5 y</td>
</tr>
<tr>
<td>Alcoholism and other diseases of despair</td>
<td>Mental health disruption, unemployment</td>
<td>Next 10 y</td>
</tr>
<tr>
<td>Multiple chronic diseases</td>
<td>Unemployment, lack of health insurance and poverty</td>
<td>Next 20 y</td>
</tr>
<tr>
<td>Lack of proper medical care</td>
<td>Disruption of healthcare, as hospitals and health programmes get financially disrupted, furlough personnel or even shut down services</td>
<td>Next 20 y</td>
</tr>
</tbody>
</table>

Abbreviation: AMI, acute myocardial infarction.
COMMENTARY

3 | EXCESS DEATHS FROM COVID-19 VERSUS FROM MEASURES TAKEN

Counting population-wide excess deaths offers complementary perspectives, but exhibits considerable year-to-year variation. More importantly, differentiating between COVID-19 deaths and those due to harmful response measures is challenging.5,6 Some of the deaths due to the measures taken happen acutely (e.g., due to people with acute myocardial infarction not coming to the hospital for care),7,8 but the majority may accrue over longer periods of time (Table 1). There is strong evidence on the adverse effects of unemployment, financial crises, depression, and social isolation on long-term morbidity and mortality,9-14 but caution is needed to extrapolate this evidence to the current situation which is unprecedented in terms of the acuteness and massive impact of the measures taken. Some projections have been made for these excess deaths, and evidence is already accumulating for some of these excess death causes.7,8,15-25 Putting projections together, the excess deaths from the measures taken is likely to be much larger than the COVID-19 deaths, for example, disruption of tuberculosis programmes alone is expected to cause 1.4 million extra deaths over the next 5 years and the death toll from famine can be even more staggering. However, the exact impact of these major problems has very large uncertainty, and some projections may be exaggerated (as was the cause also for COVID-19 projections).26 Their excess death toll will likely depend on our ability to address these problems early on and to avoid recurrent lockdowns and other draconian measures.

4 | AGE AND RISK STRATIFICATION

COVID-19 death risk shows tremendous risk stratification with over 1000-fold variability between children and elderly nursing home residents.4,27,28 Median age of death with COVID-19 typically tracks average life expectancy in high-income countries. Life expectancy (median age of death with COVID-19) is 81 (82) in Germany, 84 (82) in Italy, 81 (85) in the UK and 79 (77) in the USA. Divergence may be larger in some low-income countries, for example, India,29 perhaps because many extremely frail individuals survive to old age in high-income countries (and are candidates for succumbing to COVID-19) but not in low-income countries.

Within several countries, disadvantaged minorities have a greater toll.30,31 For example, in the USA, median age of COVID-19 death among Hispanic and nonwhite decedents (71 and 72 years, respectively) was 9-10 years lower than that of white decedents (81 years).32 The difference of median age of COVID-19 death from life expectancy is 11 years less for Hispanics, 3 years less for nonwhites, but 2 years more for white non-Hispanics.33 Similarly, UK has almost 5-fold higher COVID-19 death rate in blacks and Bangladeshi/Pakistani than in whites.34 Disadvantaged minorities tend to have lower income, worse health care (or even no health care), and unfavourable circumstances where they cannot be protected as easily. The extent to which lifestyle, nutrition, genetics, and adverse social environment may interact needs better study. Regardless, COVID-19 is a disease of inequality and it also creates even more inequality.

Besides age, socioeconomic factors and doubling of risk in men versus women, several clinical risk factors predispose for unfavourable outcome.27 Substantial increases in death risk (1.5- to 5-fold) are conferred by organ transplantation, severe obesity, uncontrolled diabetes, severe chronic pulmonary obstructive disease, liver failure, kidney failure, haematological malignancy and recent cancer. There is no increased risk with hypertension or remote history of cancer and only small increases (<1.5-fold) with asthma, chronic heart disease, mild obesity and cancer 1-5 years ago.27

Further study is needed on possible effects of genetic and epigenetic factors, history of other vaccinations, air pollution, lifestyle choices and previous infection with other coronaviruses on the susceptibility to SARS-COV-2 and the severity of the infection.

5 | YEARS OF LIFE LOST

Assuming that those dying with COVID-19 have the same profile of comorbidities as those of similar age in the general USA population, Goldstein and Lee estimated that on average, a person dying with COVID-19 loses 11.7 years of life.35 However, this estimate is probably highly upward biased. Those who die with COVID-19 may have more comorbidities (and thus shorter life expectancy) than the general population at same age. Hanlon et al adjusted for comorbidities36 and found that this adjustment decreases the estimated average years of life lost (YLLs) only by 1 year. However, their correction is inadequate because they considered only 11 comorbidities and lacked information on comorbidity severity (which markedly affects life expectancy). Moreover, their model that considers correlated comorbidities did not even converge, apparently due to sparse data and dense correlation structure. Even then, they observed that YLLs markedly depend on the number of comorbidities, for example, those ≥80 years without comorbidities have over 10 YLLs while those with many comorbidities have only 2-4 YLLs.36

Consideration of additional comorbid conditions and careful modelling of their correlation may further shrink YLLs estimates. Separate modelling is also needed for...
institutionalized and non-institutionalized individuals, given their markedly different life expectancy. For example, it is known already from the pre-COVID era that average length of stay in nursing homes is slightly over 2 years and those who died in nursing homes had spent there on median only 5 months. Moreover, it has long been known and pointed again recently that traditional YLLs calculations are by default inflated because inherently they count remaining life (based on life tables with or without risk adjustments) even for people dying at their expected time without any actual life loss. Finally, quality-adjusted YLLs and disability-adjusted life years (DALYs) would add valuable information, if computed carefully for COVID-19. The quality of life of many deceased patients is limited and many have major disabilities in their pre-existing situation.

6 | AVOIDABLE DEATHS

Some/many of the first 1 million recorded deaths were potentially due to errors and mismanagement that might be avoidable moving forward. For example, some health care systems were caught unprepared; widely used hydroxychloroquine may have increased mortality; and suboptimal mechanical ventilation management may have worsened outcomes. Some strategic choices, for example, sending COVID-19 infected patients to nursing homes (in anticipation of predicted acute care bed shortages) probably caused many excess deaths and nosocomial infections contributed many deaths in some hard-hit locations like Lombardy. Hopefully, many of these problems can be avoided in the future. Some are more intractable than others, for example, some health care systems may remain malfunctional and lack resources. Conversely, some deaths may be averted with the wider future use of dexamethasone that decreases the risk of death in severe illness.

7 | CURRENT EXTENT OF VIRAL SPREAD

Population seroprevalence studies published to-date show tremendous variability in evolving spread of the infection across countries, in locations within countries, and within locations according to socioeconomic and other exposure risk features. All studies, however, agree that infections far exceed the documented PCR-positive numbers. Many infections (~40%) are entirely asymptomatic, and many more have limited symptoms and/or do not lead to testing. Underestimation may have been 50-100-fold or more in the early days of the pandemic, especially in locations with limited testing. The ratio total/documentated infections has probably decreased as more testing is done. However, as of summer 2020 underestimation was apparently still 11-fold in the USA and about 30-fold in India. With 36 million documented infections worldwide as of early October 2020, the true total number of infections is probably >20 times larger: about 10% of the global population is probably already infected. This estimate is in agreement also with a recent WHO statement. Rates of further current increase may vary markedly across locations, with some locations maintaining suppressed epidemic activity, several others showing clear decline of infections with sigmoid (Gompertz) epidemic waves, and some other experiencing continued waves or resurgence after suppressed first waves.

8 | INFECTION FATALITY RATE

Infection fatality rate in different locations can be inferred from seroprevalence studies. While these studies have caveats, they show IFR ranging from 0.00% to 1.54% across 82 study estimates. Median IFR across 51 locations is 0.23% for the overall population and 0.05% for people <70 years old. IFR is larger in locations with higher overall fatalities. Given that these 82 studies are predominantly from hard-hit epicentres, IFR on a global level may be modestly lower. Average values of 0.15%-0.20% for the whole global population and 0.03%-0.04% for people <70 years old as of October 2020 are plausible. These values agree also with the WHO estimate of 10% global infection rate (hence, IFR ~ 0.15%) as of early October 2020. Earlier higher quotes of average IFR that were irresponsibly circulated widely in media and social media were probably extremely flawed, as they depended on erroneous modelling assumptions, and/or were done by inexperienced authors who used overtly wrong meta-analysis methods in a situation where there is extreme between-study heterogeneity. For discussion of analytical issues, see ref. 43.

The sharp age dependence of risk means that IFR is expected to vary substantially, other things being equal, across different countries. Median population age is 15-20 years for most African countries versus 43 years in the European Union. Globally, the median age is 30; 9% of the 7.7 billion people are ≥65 years old, 50% are 25-64 and 41% are younger than 25. IFR estimates across different locations are expected (and observed) to vary many-fold based on differences in population age structure, presence of elderly institutionalized populations, socioeconomic inequalities, population-level clinical risk profile, measures taken and healthcare. It is unclear whether differences in host genetic susceptibility, viral clades and other unknown factors may also diversify IFR.
TABLE 2 Estimated COVID-19 deaths during the full cycle of the pandemic under different scenarios of population infection rate (PIR) that is the same across all risk strata or differs in high-risk (PIRH) and low-risk (PIRL) strata

<table>
<thead>
<tr>
<th>Global population (millions)</th>
<th>Infection fatality rate</th>
<th>Estimated COVID-19 deaths during the full cycle of the pandemic (millions)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PIR = 60%</td>
<td>PIR = 30%</td>
</tr>
<tr>
<td>Institutionalized frail elderly</td>
<td>1.5</td>
<td>0.75</td>
</tr>
<tr>
<td>Other &gt;75 y</td>
<td>3</td>
<td>1.5</td>
</tr>
<tr>
<td>Other 65-74 y</td>
<td>2.7</td>
<td>1.35</td>
</tr>
<tr>
<td>Upper-risk &lt;65 y</td>
<td>1.2</td>
<td>0.6</td>
</tr>
<tr>
<td>Low-risk &lt;65 y</td>
<td>0.36</td>
<td>0.18</td>
</tr>
<tr>
<td>All</td>
<td>8.76</td>
<td>4.38</td>
</tr>
<tr>
<td>COVID-19/total 5-y global deaths</td>
<td>2.9%</td>
<td>1.5%</td>
</tr>
</tbody>
</table>

*Simulations are given for illustrative purposes and need to be seen with great caution. They should not be interpreted by any means that a ‘herd immunity’ strategy is proposed where people are encouraged to become infected. It is also unknown whether a full cycle would last 5 y, or less or more, and what the long-term behaviour of SARS-CoV-2 would be (eg, whether it may behave like the other four coronaviruses that cause sporadic outbreaks). Infection fatality rate is classified here in 5 bins for parsimony, but of course risk functions in reality are continuous. The presented simulations correspond to a global infection fatality rate (IFR) of 0.19% if people in all risk strata have an equal chance of infection, but this would vary across locations and countries, for example, the same assumptions translate to IFR = 0.37% in the USA (0.25% in non-institutionalized people) versus approximately 0.1% in India. IFR can be modulated to decrease sharply if high-groups are selectively protected, while it may increase sharply if high-risk groups are infected more frequently than low-risk groups.

9 | FUTURE POTENTIAL PANDEMIC SPREAD

Per standard epidemic modelling, a basic reproductive number of 2.5 translates to 1-(1/2.5) = 60% of the population required to be infected to reach ‘herd immunity’. However, these estimates assume equal mixing within populations, while real-world heterogeneity is the norm. Seroprevalence values approaching 60% have been documented in overcrowded urban areas in India\(^50\) or South America\(^51\) and highly congested settings, for example, aircraft carriers.\(^52\) With mixing heterogeneity, lower values, for example, 43%\(^53\) or even 10%-20%\(^54\) have been proposed as required thresholds to stop epidemic propagation. Moreover, multiple studies have identified pre-existing cellular immunity that may be effective against SARS-CoV-2 in 20%-50% of participant samples.\(^55\)-\(^57\) If so, the proportion of people who need to be infected to reach herd immunity may be much lower than originally estimated. Thresholds for herd immunity remain a contested, but crucial issue as they determine the projected potential total fatalities.

10 | TOTAL FATALITIES IN A FULL-CYCLE PANDEMIC

Table 2 shows illustrative projections for total global COVID-19 deaths for a full cycle of the pandemic without considering modifications due to currently unavailable or unknown factors (eg, vaccines, see next section). The time it takes for the pandemic to unfold may vary across locations, depending on original seeding load, timing of re-seedings and real-world effectiveness of employed non-pharmaceutical interventions—a hotly debated topic beyond the scope of the current article. It is argued that 2-5 years may be needed for full cycling.\(^58\) However, some locations around the world may have already completed a largely full cycle, while others may remain mostly unscathed by the virus (but thus also continuously susceptible) for long even without effective vaccines, for example if they continue to block seeding from external sources.

Table 2 simulations show that if eventually 60% of the global population is infected and there is the same risk of infection across all risk strata, the total number of deaths is expected to be 8.76 million for the full cycle. If one assumes a 5-year horizon, this represents 2.9% of all deaths globally in the period 2020-2024. If only 30% of the global population is infected (a more plausible expectation) without differentiation across risk strata, the total number of deaths (4.38 million) is 1.5% of all deaths globally in 2020-2024. Further major reductions in total deaths can be achieved, if measures succeed to keep infection rates in high-risk groups at half or one-third of the rate in remaining populations: 2.28 and 1.58 million deaths, respectively, would represent only 0.8% and 0.5% of all deaths globally. If the infection rate among high-risk groups can be kept at 10%, then even if 60% of the remaining population is infected, total COVID-19 death count would remain 1.76 million. Given that >1 million deaths are already
documented as of October 2020, if the minority of high-risk individuals can be preferentially protected with modest effectiveness, the remaining deaths would be fewer than those already accrued. The proportion of global quality-adjusted or disability-adjusted life years lost due to COVID-19 may be even less than the proportion contributed in terms of death counts, as discussed above.

In the first half of 2020, high-risk groups were not strongly preferentially protected in many locations. In fact, in some countries, some high-risk groups were probably infected at higher rates than low-risk groups. Horizontal lockdown protected several low-risk groups (eg, wealthy healthy professionals working from home) more than high-risk groups who could not shelter effectively. This applies both to people at high risk because of socioeconomic inequalities (eg, homeless, low-wage essential workers and minorities in the USA, poor urban dwellers and manual workers in Latin America), as well as age group and debilitation (eg, with massive infections in nursing homes in USA and Europe). Large seroprevalence studies with sufficient participants in different age strata to allow meaningful comparisons suggest that, compared with younger people, non-institutionalized people >65 years were equally likely to be infected in Spain, slightly less likely to be infected in the USA, and substantially less likely to be infected in England. Moreover, as discussed above, minorities and poor people were often disproportionately infected.

Box 1 Additional factors that may affect the toll from the COVID-19 pandemic

- Vaccines—successful development, availability, effectiveness, safety, uptake, coverage of high-risk populations, impact on transmission, duration of protection.
- Development and use of effective treatments and management options and avoidance of detrimental ones
- Impact of economic and social disruption on the course and management of the pandemic
- Public health and personal hygiene measures
- Interplay with other emerging health problems—respiratory infectious (eg, influenza), other infectious (eg, tuberculosis), other diseases induced/worsened by the epidemic response (eg, competing for resources)
- Re-infection potential—loss of immunity and/or mutating virus
- Chronic COVID-19 disease consequences and long-term morbidity leading to late mortality
- Catastrophic chaotic events (eg, wars, riots, revolutions and other social meltdown)

11 | ADDITIONAL FACTORS THAT MAY SHAPE THE PANDEMIC FOOTPRINT

Box 1 summarizes several other factors that may affect the total pandemic toll. Those that have the highest likelihood of occurring may have positive impact, further reducing the pandemic impact. Emergence of effective and safe vaccines and additional effective treatments, and avoidance of ineffective and detrimental management options are all highly desirable. As of October 2020, it is precarious to speculate about their exact impact, which may vary from very modest to paradigm-changing. Conversely, highly disruptive measures (eg, lockdown) may drain resources and hinder responding to the pandemic, besides whatever other major adverse effects they may have on other health problems and society at large.

The co-existence of COVID-19 and seasonal influenza remains a major unknown as of October 2020. Preliminarily, there is some evidence that influenza seems suppressed while the COVID-19 pandemic is active. If true, this may reflect effectiveness of hygiene, masks and non-pharmaceutical social distancing measures against influenza as well. In addition, one perspective is that there is a pool of frail, susceptible individuals who are at high risk of succumbing to respiratory viruses. Thus, a less severe season may be followed by a more severe one, and vice versa. Moreover, influenza and SARS-CoV-2 would be competing for the same pool of susceptible individuals. In the absence of COVID-19, influenza would be expected to kill 2.5 million people or more in 5 years, including approximately 150 000 children <5 years old. It would be very interesting to note whether its death toll in 2020-2024 is actually smaller, given the advent of COVID-19. An optimistic scenario would be that influenza recedes during COVID-19 waves, and that total number of deaths during 2021-2024 due to respiratory pathogens is cumulatively not much different from pre-COVID-19 4-year periods. Conversely, the pessimistic scenario is that influenza and COVID-19 both strike heavily and concurrently with multiplicative adverse impact.

Another unknown feature is the exact frequency, timing and clinical severity of re-infections from SARS-CoV2. Data to-date do not suggest that this is a significant contributor to mortality, but the impact of re-infections needs long-term tracking. Long-term morbidity and mortality among COVID-19-infected patients are also poorly understood and systematic study is needed.
Finally, both COVID-19 and the response measures (especially if they are too aggressive) can disrupt life, economy, civilization and society at large. A catastrophic impact on mental health is already well documented.\(^6\) Catastrophic social meltdown and chaotic events such as riots, wars and revolutions have unpredictable dynamics but, if they happen, can be devastating. Many measures taken to halt the pandemic may be seriously destabilizing, adding hundreds of millions of people at the brink of starvation, skyrocketing unemployment and resulting in recrudescence of other infectious diseases such as tuberculosis and childhood diseases from disrupted vaccination schedules.\(^6\) Learning to live with COVID-19 and using effective, precise, least disruptive measures is essential to avoid such disasters and to help minimize the adverse impact of the pandemic.

**CONFLICT OF INTEREST**
None.

**FUNDING INFORMATION**
No funding was received for this work. METRICS has been supported by grants from the Laura and John Arnold Foundation.

**ORCID**
John P. A. Ioannidis https://orcid.org/0000-0003-3118-6859

**REFERENCES**
24. Fu SJ, George EL, Maggio PM, Hawn M, Nazarelli R. The consequences of delaying elective surgery: surgical perspective. Ann...


56. Doshi P. Covid-19: Do many people have pre-existing immunity? *BMJ.* 2020;370:m3563. https://doi.org/10.1136/bmj.m3563


