Estimating the Impact of COVID-19 on the Individual Lifespan:

A Conceptual Detour and an Empirical Shortcut

First Version, 7/14/2020

Revised, 8/7/2020

Patrick Heuveline, Ph.D.
California Center for Population Research (CCPR)
University of California, Los Angeles (UCLA)

Contact author: Patrick Heuveline <heuveline@soc.ucla.edu>

Abstract

The impact of COVID-19 on the individual lifespan can be measured by the difference in period life expectancy at birth (PLEB), an intuitive indicator of mortality conditions during a reference period. When mortality conditions are changing rapidly, however, that intuitive interpretation of the PLEB for short reference periods and of its change conflict with the assumptions under which the PLEB is derived.

To avoid assumptions about future mortality, I propose measuring instead the Mean Unfulfilled Lifespan (MUL), defined as the average difference between the actual and otherwise expected ages at death in a recent death cohort. For fine-grained tracking of the pandemic, I also provide an empirical shortcut to MUL estimation for small areas or short periods.

I estimate quarterly MUL values for the first half of 2020 in 142 national populations and 91 sub-national populations in Italy, Spain and the US. Across countries, the highest quarterly
values were reached in the second quarter in Peru (3.90 years) and in Ecuador (4.59 years). Higher quarterly values still were found in New York and New Jersey, where individuals died respectively 5.41 and 5.56 years younger on average than their expected age at death.

Using a seven-day rolling window, I estimate the MUL peaked at 7.32 years in Lombardy, 8.96 years in Madrid, and 8.93 years in New York, but reached 12.86 years for the entire month of April in Guayas (Ecuador). These results illustrate how the MUL provides an intuitive metric to track the pandemic without requiring assumptions about future mortality.

Introduction

In the past few months, the numbers of deaths from the novel coronavirus disease 2019 (COVID-19) have become part of the daily news cycle the world over. Impressively though these numbers are, they may not convey a clear sense of the pace and scale of the pandemic. By contrast, declines in life expectancies induced by COVID-19 mortality provide a simple and intuitive metric.

The period life expectancy at birth (PLEB) is the inverse of a “stationary” death rate: a weighted average of the period age-specific death rates with weights derived from these death rates through life table construction. Using these internally-derived weights rather than an external, standard age-distribution, the stationary death rate still achieves the main goal of age-standardization, namely, to yield a mortality measure independent of the actual age composition of the population. The inverse of that weighted average, the PLEB thus allows for mortality comparisons that account for the known variations of COVID-19 mortality with age, and is also much more intuitively measured in years lived per person rather than in deaths per thousand person-years lived.

An aggregate indicator of all-age mortality conditions during a given period, however, the period life expectancy at birth (PLEB) is relatively insensitive to mortality changes at older ages. In high-income countries, where mortality at young ages is already low, annual changes in PLEB have been in the order of +.2 years annually. With the notable exception of periods of armed conflict, declines in PLEB have become rare and similarly modest. In the US,
instance, the most recent reversals in the PLEB upward trend are a .3 years decline during the opioid-overdose crisis, from 78.9 years in 2014 to 78.6 years in 2017, and a previous .3 years decline in just one year, from 75.8 years in 1992 to 75.5 years in 1993, at the peak of the HIV epidemic.5

The impact of COVID-19 mortality on 2020 PLEB can be expected to be substantially larger in the US and a number of other countries. The nearly 300,000 COVID-19 deaths projected by December 1st for the US by the University of Washington’s Institute for Health Metrics and Evaluation (IHME) would translate into a 2020-PLEB reduction of over one year. Based on this set of projections, reductions in 2020-PLEB would exceed two years in Peru and three years in Ecuador.6

Estimates of PLEB reductions are sensitive to the scale of the population and the length of the period they refer to. For most purposes, a national focus might not be ideal to gain insights on the pandemic, especially in countries spread on large territories like the US, Brazil or Mexico, not to mention China or India, where national trends may be the sum of different local epidemics. By averaging out COVID-19 death rates in the least and most affected states and countries, for instance, the impact of COVID-19 on the US 2020 PLEB conceals large differences across the country. The 2020 PLEB similarly averages out death rates before the first COVID-19 deaths and during the most severe months of the pandemic.

Consequently, estimating years of PLEB reductions for smaller, particularly affected areas during shorter, particularly intense periods yields more impressive results, some in double digits.7·8·9 As is de rigueur in the limitations section of the accompanying papers, we may be warned in passing against any literal interpretation of these PLEB estimates, but with such numbers all but meant to attract attention beyond a small circle of demographers, their
underlying assumptions may warrant a somewhat longer discussion. Since the PLEB is the expected age at death of a newborn experiencing the mortality conditions of the reference period during her entire lifetime, the PLEB for 2020 assumes an annual re-occurrence of the 2020 swings in mortality induced by COVID-19 or another pandemic with a similar mortality impact. One may hope that such an annual occurrence will not become a “new normal,” but this possibility cannot be entirely ruled out on principle either. PLEB estimates for shorter periods and smaller areas estimate the age at death of a newborn experiencing the conditions in that short period and small area for her entire lifetime. The highest PLEB reductions are thus estimated on the rather sinister assumption of a Groundhog-Day-like time loop repeating the worst week of the COVID-19 crisis in some of the worst affected areas, with those born there unable to ever leave.

Meanwhile, tracking the pandemic at a finer-grained geographical and temporal scale than what reasonable PLEB estimates can provide might indeed provide useful insights. Age-standardized death rates are perfectly adequate for this, purpose but expressed in deaths per thousand, ten thousand or a hundred thousand of person-years, the significance of changes in those rates remains difficult to grasp. In the rest of this paper, I suggest an alternative measure of the impact of a cause of death on the individual lifespan, that is expressed in years and can be estimated for populations of any size and for periods of any length. I estimate its value for COVID-19-induced changes in mortality in 142 countries and 91 first-level sub-national administrative entities in Italy, Spain and the US, for each sex and each of the first two quarters of 2020. Finally, I show that in a given population, its value remains nearly proportional to the ratio of COVID-19 deaths to total deaths, providing an easy short-cut to track the pandemic over time in terms of years of lost length of life.
**Conceptual Detour**

The impact of COVID-19 on life expectancies is estimated by comparing two period life tables, one representing the prevailing mortality conditions and another one representing the counterfactual mortality conditions expected in the absence of COVID-19. One approach is to start with the latter using a previous projection of period mortality conditions and to derive the former by adding estimates of COVID-19 deaths by sex and age group in the period, reversing the steps of the standard demographic technique that estimates the impact of eliminating a cause of death on life table functions. This approach is sensitive to biases in the estimates of COVID-19 deaths and ignores any potential “indirect” effect of COVID-19 on death rates from other causes.

To include these indirect effects and by-pass the estimation of COVID-19 deaths, an alternative approach is to derive one life table from past “benchmark” mortality conditions and one from the prevailing conditions, treating deaths derived with the first table as expected deaths and additional deaths derived with the second table as “excess” deaths directly or indirectly attributable to COVID-19. The estimation of excess deaths is thus quite sensitive to the choice of a benchmark period to represent past mortality conditions. In France, for instance, the count of roughly 30,000 COVID-19 deaths for the first half of 2020 is substantially higher than the number of excess deaths for the same period when the year 2019 is used as a mortality benchmark. The estimate of excess deaths would increase to over 40,000 deaths, however, were years 2000 through 2019 used to estimate benchmark mortality. By contrast, counterfactual mortality is largely netted out in the previous approach and the choice of particular mortality conditions has little impact on the estimates of PLEB reductions.
While the two approaches have slightly different aims and neither one has a clear empirical edge, they both amount to assessing the difference in PLEB between life tables with and without deaths attributed or attributable to COVID-19. Many steps are involved in this assessment, but as Nathan Keyfitz pointed out when considering the effect of eliminating a cause of death, the difference “depends on the average time that elapses before the persons rescued will die of some other cause.”\(^{15}\) In the case at hands, the effect of a new cause of death on longevity depends on the average time that would have elapsed before the persons who died from that cause would have died from other causes. In life table notations, that average is:

\[
\frac{1}{l_0} \int_0^\omega (d^C(a) \cdot e_a^{o-C}) da
\]

where \(l_0\) and \(d^C(a)\) are the radix and the number of decrement from the new cause of death at age \(a\) in the multi-decrement life table representing the prevailing mortality conditions, whereas \(e_a^{o-C}\) is the life expectancy at age \(a\) in the single-decrement (or all decrements combined) period life table representing the counterfactual mortality conditions in the absence of that new cause.

This difference relates to the concept of “potential years of life lost,”\(^{16}\) which has been popularized by burden of disease assessments as Years of Life Lost (YLL):

\[
YLL = \int_0^\omega (D^C(a) \cdot e_a^{o*}) da
\]

Here \(D^C(a)\) is the number of deaths from a certain cause at age \(a\) in the population during a certain period and \(e_a^{o*}\) is life expectancy at age \(a\) in a counterfactual period life table. (Various options exist for selecting that life table in burden of disease assessments, including a universal table representing optimal survival conditions). Several assessments of the impact of COVID-19 using this approach already exist, measuring this impact in total number of years lost to the pandemic and allowing to compare it to the impact of other causes of death.\(^{17, 18, 19}\)
Combining these two traditions, an alternative to the decline in PLEB as a measure of the impact of an additional cause of death on the individual lifespan is the Mean Unfulfilled Lifespan (MUL):

\[ MUL = \frac{1}{D} \int_{0}^{\omega} (D^c(a) \cdot e_{a}^{\omega-c}) \, da \]

where \( D \) is now the total number of deaths (from all causes at all ages) in the population during a reference period and \( e_{a}^{\omega-c} \) is the life expectancy at age \( a \) in the period life table representing mortality conditions in the population without that particular cause of death. Following Keyfitz’ insight, the MUL can be interpreted as the difference between the actual and expected age at death, averaged over a death cohort, that is, members of the population dying in a certain period.

Using the actual distribution of deaths in the population rather than life table decrements implies that, unlike reductions in PLEB, MUL values depend on population composition. All else equal, a younger population composition yields a younger distribution of deaths and a higher MUL value. However, as discussed above, using the life table decrements amount to assuming stationarity, which is precisely what makes interpreting short-term changes in PLEB problematic. Making no assumption about future mortality conditions, the MUL shares with the PLEB the appeal of measuring the impact of a cause of death on longevity in years.

*Empirical Shortcut*

Since MUL values average differences between actual and expected ages at death across all deaths occurring during a certain period and in a certain population, they are sensitive to the length of the period and the size of the population to which they refer to. Further disaggregation to smaller geographical areas and to shorter durations thus produces both smaller and larger than average values. Studying these variations over time and across areas might also be of interest, and, contrary to reductions in PLEB, calculating MUL values for small geographical areas and
short periods is not conceptually problematic as they refer to lived lives, not to future expectations. The demand on data (including a separate counterfactual life table for each population of interest) is substantial, however, and the computations are not particularly straightforward. To circumvent those, the definition of the MUL can be reconsidered as follows:

\[
MUL = \frac{1}{D} \cdot \int_0^{\omega} (D^c(a) \cdot e_a^{o-c}) da = \frac{D^c}{D} \cdot \int_0^{\omega} \left( \frac{D^c(a)}{D^c} \cdot e_a^{o-c} \right) da \\
= \frac{D^c}{D} \cdot \int_0^{\omega} \left( \frac{M^c(a) \cdot N(a)}{M^c \cdot N} \cdot e_a^{o-c} \right) da = \frac{D^c}{D} \cdot \int_0^{\omega} \frac{M^c(a)}{M^c} \cdot (C(a) \cdot e_a^{o-c}) da
\]

where \(M^c\) and \(M^c(a)\) are the all-age death rate and the death rate at age \(a\) from a specific cause, and \(C(a)\) is the proportion of individuals aged \(a\) in the population.

In the case at hands, the last terms in the integral sum, life expectancies in the absence of COVID-19 are fixed. Population composition may also be treated as relatively invariant within each of the three-month periods. Finally, the age pattern of COVID-19 death rates can also be expected to vary little within short periods. This suggests that MUL values for any period of time within a given quarter can be approximated as the product of a constant value and the ratio of COVID-19 deaths to all deaths during that period. Assuming only modest variation between sub-populations in pre-COVID-19 life expectancies and population compositions, MUL values for sub-populations can similarly be estimated as the product of a common value (for all sub-populations) and the ratio of COVID-19 deaths to all deaths during that period in each of the sub-populations.

**Results**

Figure 1 shows first-quarter MUL values (darker bars) and the inter-quarter increases in MUL values (lighter bars) for both sexes in countries or sub-national entities with the largest second-
quarter values (represented by the total length of the two bars) and at least 1,000 COVID-19 deaths by July 1, 2020.

Figure 1: Mean Unfulfilled Lifespan (MUL) for both sexes, by quarter and populations with 1,000 or more COVID-19 deaths by July 1st, 2020, in years.

Figure 1 shows COVID-19 impacting the individual lifespan first in parts of Italy and Spain, with limited impact elsewhere except in Ecuador. In the first quarter, individuals died in Madrid 3.16 years younger on average than their otherwise expected age at death (in the absence of COVID-19). The corresponding figure for Lombardy was 2.75 years. The average for Spain (.98 years) and Italy (.88 years) were lower than for Ecuador, however, the only country to experience a MUL in excess of one year during that quarter (1.29 years, Figure 1).
MUL values further increased from the first to the second quarter across Europe, reaching 3.66 years in Madrid and 3.04 years in Lombardy in particular, but increased sharply as well in Belgium (2.81 years, Figure 1). By the second quarter, however, the largest impacts of COVID-19 on the individual lifespan had shifted to the Western Hemisphere. In that quarter, individuals died in Ecuador 4.59 years younger on average than their otherwise expected age at death. The corresponding figure was 3.90 years in Peru. In the US, the second-quarter MUL averaged 1.76 years for the nation as a whole, but reached 5.41 in New York and 5.56 years in New Jersey (Figure 1). The above values refer to both sexes combined, but MUL values for men only are even higher. During the second quarter, men died in New Jersey 6.05 years younger on average than their expected age at death before COVID-19.

To apply the suggested empirical short-cut, Appendix Table S1 provides quarterly values of the ratio of the MUL to the ratio of COVID-19 deaths to all deaths. These values range from a low of 9.62 years in Bulgaria to a high of 26.48 years in Qatar. These differences can be explained by different age compositions, with younger compositions giving more weight to remaining life expectancies at younger ages, which are obviously higher. Values for each of the two quarters in the Appendix Table S1 are relatively close, providing support for the assumption that they are nearly invariant within a quarter. Our values for the USA, 12.64 years in the first quarter and 12.71 years in the second are also consistent with an earlier estimate that the country would lose 12.3 million years of remaining life if the COVID-19 death toll was to reach 1 million.20

To illustrate the approximation using values in Appendix Table S1, we approximated MUL values for a rolling seven-day period in Lombardy, Madrid and New York. Figure 2 shows
these MUL values from mid-March to mid-May, peaking at 7.32 years in Lombardy, 8.96 years in Madrid and 8.93 years in New York.

![Figure 2: Mean Unfulfilled Lifespan (MUL) for both sexes, seven-day averages, in years.](image)

The approximation for sub-populations is illustrated by focusing on the province of Guayas in Ecuador. Data on the monthly number of deaths by province show the marked increase in Guayas in March, April and May from a baseline of 1,700-2,000 per month in January, February and again in June. In April, the number of deaths reached 12,004, of which 84.5% could be assumed to be COVID-19 using the January-February-June average to estimate deaths from other causes (without adjustment for competing causes). Whereas no specific life table is available to estimate MUL values for the province directly, a value for the month of April can be estimated, assuming the same proportionality of the MUL to the ratio of COVID-19 to all deaths as in Ecuador during the second quarter. The resulting approximation suggests that
in April individuals died in the province of Guayas 12.86 years younger, on average, that their expected age at death, possibly the largest impact of COVID-19 on the individual lifespan to date.

**Discussion**

The Mean Unfulfilled Lifespan provides an alternative to the reduction in PLEB to express the impact of COVID-19 on the individual lifespan in an easily interpretable metric (years of life lived), especially for tracking the impact of the pandemic over short periods of time, for which the assumptions of PLEB are untenable. This mean value applies to a death cohort, that is individuals who die in a certain area during a certain period. Based on the expected mortality conditions before the onset of COVID-19, the MUL indicates how much younger members of this death cohort actually died on average.

MUL values were estimated for a total of 233 populations for each sex and each of the first two quarters of 2020. To estimate this large number values, the simplifying assumption of a common age-and-sex pattern of COVID-19 mortality was used (i.e., the same age-and-sex-specific death rates relative to all-age, both-sex death rate in all populations). The distribution of COVID-19 deaths by sex and age-group derived so can obviously be replaced with the actual distribution in populations for which that distribution is available. A second approximation was to estimate the difference between the actual and expected ages at death for individuals dying on an age interval (ages $x$ to $x+n$) using the average age at death on that age interval and life expectancy prior to COVID-19 at beginning of that age interval (exact age $x$). This may entail a slight underestimation of life expectancy at the exact age at death and thus of the MUL overall. MUL values should also be sensitive to the age groups for which deaths are reported, with higher cut-offs for the open age-interval likely resulting in higher MUL values.
To estimate MUL values for smaller-area populations for which a separate life table might not be available or for short-duration periods, another simplifying assumption was introduced, namely, the proportionality of the MUL value to the ratio of COVID-19 to total deaths in the population during the period. This proportionality seems to be a good assumption within relatively short-duration periods as the proportion was shown to hardly change between the first and the second quarter of 2020. Within administrative units, the proportionality assumption is only valid if the age composition of sub-unit populations is not too dissimilar to the composition of the entire-unit population.

MUL values are not sex or age standardized and do not substitute for sex and age standardized COVID-19-specific death rates for comparisons of COVID-19 mortality across populations that differ markedly in age composition. While MUL comparisons across such populations will be biased, the MUL provides an easily interpretable indicator for comparisons across smaller areas within a larger administrative unit for which the necessary data are provided or for tracking trends over short periods of time, without making any unrealistic assumption about future mortality trends.

**Materials and Methods**

The equations defining YLL or MUL appear deceivingly simple. To implement them, one has to match numbers of deaths or decrements that are available or can be estimated on age intervals, with life expectancies that refer to exact ages. When age intervals are sufficiently narrow, interpolations of life expectancies on these intervals would likely provide good approximations of the integral sum in the MUL equation. This might not be the case for COVID-19 deaths. In the US for instance, about 60% of COVID-19 deaths are reported in just two age groups: 75 to 84 years, ages between which mortality changes rapidly, and 85 years and over, for which the YLL
and MUL equations do not provide an estimate of the difference between the actual and expected ages at death.

For deaths between ages $x$ and $x+n$ in a given period, $nD_x$, the difference between the actual and expected ages at death can be estimated as:

\[
\left( nD_x - nD_x^- \right) \cdot (e_x^{o-c} - n\alpha_x) + nD_x^- \cdot (n\alpha_x^- - n\alpha_x)
\]

where $nD_x^-C$ is the number of deaths between ages $x$ and $x+n$ that would have been expected during the same period without COVID-19, while $n\alpha_x$ and $n\alpha_x^-C$ are the average number of years lived after age $x$ for individuals dying between ages $x$ and $x+n$ in the period life tables with and without COVID-19. The first term above corresponds to excess deaths, directly or indirectly attributable to COVID-19, for which the difference between the actual and expected ages at death is estimated at age $x$, as the difference between life expectancy at that age and the average number of years lived after that age for those dying in the age interval. This entails a small underestimation of life expectancy at age $x+n\alpha_x$. The second term, relatively minor numerically, accounts for the fact that even individuals who were expected to die in the age interval do so a little earlier on average.22 For the open-ended age interval, the estimated difference between actual and expected age at death is the same for expected and for excess deaths, and averaged across all age groups, the MUL value can be estimated as:

\[
MUL = \frac{1}{D} \left( \sum_{x=0}^{N-n} \left( \left( nD_x - nD_x^- \right) \cdot (e_x^{o-c} - n\alpha_x) + nD_x^- \cdot (n\alpha_x^- - n\alpha_x) \right) + \left( D_{N+} \cdot (e_N^{o-c} - e_N) \right) \right)
\]

Using this approximation to estimate the impact of COVID-19 on the individual lifespan in each of the first two quarters of 2020 first requires a life table representing survival conditions
in the first half of 2020 in the absence of COVID-19 from which the values of $e_{x}^{C}$ and $n_{d_{x}}^{C}$ can be used. Combined with the number of individuals by sex and age-group, the life table values of $n_{m_{x}}^{C}$ then provide the expected numbers of deaths $n_{D_{x}}^{C}$ in the absence of COVID-19. Population data and life table functions for countries were obtained from the UN Population Division. Corresponding data for the first sub-national administrative units in Italy, Spain, and the US were obtained from national statistical agencies.

New life tables representing actual mortality conditions (with COVID-19) in each quarter must then be derived to calculate the corresponding values of $n_{d_{x}}$. The construction of these life tables require the quarterly numbers of deaths by sex and age group. In countries where vital statistics are unavailable or incomplete, but estimates of COVID-19 deaths are available, the total number of deaths, $n_{D_{x}}$, can be obtained by adding these estimates (through a multi-decrement life table to adjust for competing risks of deaths) to the expected numbers of deaths in the absence of Covid-19 ($n_{D_{x}}^{C}$). When COVID-19 estimates are not broken down by sex and age-group, an alternative is to use a reference set of age-and-sex death rates from COVID-19 from another population for which these rates are deemed reliable. Centers for Disease Control and Prevention (CDC) data provided the reference set of age-and-sex death rates from COVID-19. Estimates of COVID-19 deaths by country and sub-national units for April 1, and July 1, 2020 were taken from the IHME. All of these data were downloaded from institutional websites.

**Acknowledgments**

The author benefited from facilities and resources provided by the California Center for Population Research at UCLA (CCPR), which receives core support (P2C-HD041022) from the
Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD).

The author thanks Michel Guillot for comments on an earlier draft this manuscript.
References

https://bmjopen.bmj.com/content/10/7/e035932


4 Obermeyer et al. 2008. Fifty Years of Violent War Deaths from Vietnam to Bosnia: Analysis of Data from the World Health Survey Programme. BMJ 336(7659): 1482. doi: https://doi.org/10.1136/bmj.a137


https://www.medrxiv.org/content/10.1101/2020.04.29.20084335v3

https://www.medrxiv.org/content/10.1101/2020.06.03.20120972v2

https://www.medrxiv.org/content/10.1101/2020.07.16.20155077v1


23 United Nations, Department of Economic and Social Affairs, Population Division. *World Population 
used June 2, 2020.


27 Centers for Disease Control and Prevention (CDC). *CDC Wonder*. https://wonder.cdc.gov/cmft-

28 Guilmoto CZZ. 2020. COVID-19 Death Rates by Age and Sex and the Resulting Mortality Vulnerability 
of Countries and Regions in the World. MedRxiv
https://www.medrxiv.org/content/10.1101/2020.05.17.20097410v1

29 Centers for Disease Control and Prevention. *Provisional COVID-19 death counts by sex, age and state*. 
July 12, 2020.