

1 **Full Title**

2 **The Mean Unfulfilled Lifespan (MUL):**
3 **A new indicator of the impact of mortality shocks on the individual lifespan,**
4 **with application to global 2020 quarterly mortality from COVID-19**

5
6 **Short Title**

7 **The Mean Unfulfilled Lifespan:**
8 **A new indicator of the impact of mortality shocks on the individual lifespan**

9
10 *Updated Version, 2/18/2020*

11 *Original Manuscript Submitted 8/9/2020, Published 8/13/2020 at:*

12 <https://www.medrxiv.org/content/10.1101/2020.08.09.20171264v1>

13
14 Patrick Heuveline, Ph.D.

15 California Center for Population Research (CCPR)

16 University of California, Los Angeles (UCLA)

17
18 Corresponding author: Patrick Heuveline <heuveline@soc.ucla.edu>

19
20

21 **Abstract**

22 Declines in period life expectancy at birth (PLEB) provide intuitive indicators of the impact of a
23 cause of death on the individual lifespan. Derived under the assumption that future mortality
24 conditions will remain indefinitely those observed during a reference period, however, the
25 intuitive interpretation of a PLEB becomes problematic when that period conditions reflect a
26 temporary mortality “shock”, resulting from a natural disaster or the diffusion of a new epidemic
27 in the population for instance.

28 Rather than to make assumptions about future mortality, I propose measuring the
29 difference between a period average age at death and the average expected age at death of the
30 same individuals (death cohort): the Mean Unfulfilled Lifespan (MUL). For fine-grained
31 tracking of the mortality impact of an epidemic, I also provide an empirical shortcut to MUL
32 estimation for small areas or short periods.

33 For illustration, quarterly MUL values in 2020 are derived from estimates of COVID-19
34 deaths in 159 national populations and 122 sub-national populations in Italy, Mexico, Spain and
35 the US. The highest quarterly values in national populations are obtained for Ecuador (5.12
36 years, second quarter) and Peru (4.56 years, third quarter) and, in sub-national populations, for
37 New York (5.52 years), New Jersey (5.56 years, second quarter) and Baja California (5.19 years,
38 fourth quarter). Using a seven-day rolling window, the empirical shortcut suggests the MUL
39 peaked at 9.12 years in Madrid, 9.20 years in New York, and 9.15 years in Baja California, and
40 in Guayas (Ecuador) it even reached 12.6 years for the entire month of April.

41 Based on reported COVID-19 deaths that might substantially underestimate overall
42 mortality change in affected populations, these results nonetheless illustrate how the MUL tracks

43 the mortality impact of the pandemic, or any mortality shock, retaining the intuitive metric of
44 differences in PLEB, without their problematic underlying assumptions.

45

46

47 **Introduction**

48 For months, the numbers of deaths from the novel coronavirus disease 2019 (COVID-19) have become
49 part of the daily news cycle the world over. Even when related to the population size in deaths per capita
50 ratio, however, these numbers do not really provide any intuition for the magnitude nor the dynamics of
51 the pandemic. Quite useful for between-population comparisons, the age standardization of these ratios
52 does not make them more easily interpretable.

53 The period life expectancy at birth (PLEB) is probably the most readily interpretable of the period
54 indicators mortality. Translating a number of deaths from a given cause into its impact on PLEB involves
55 multiple steps but is fairly straightforward.^{1, 2} Unfortunately, the intuitive appeal of the PLEB, its
56 interpretation as a measure of the individual lifespan, derives from the assumption that period mortality
57 conditions will continue to prevail indefinitely. Declines in PLEB induced by a relatively rapid and likely
58 temporary increase in mortality, such as currently experienced with the COVID-19 pandemic, can thus
59 provide misleading indicators of changes in the individual lifespan.³

60 The assumption is even more problematic when PLEB or changes therein are estimated for
61 smaller populations and shorter periods of time. Tracking the pandemic at a finer-grained geographical
62 and temporal scale undoubtedly provides better insights on the pandemic than annual, national averages.^{4,}
63 ^{5,6} But while the assumption underlying an annual PLEB estimate—that mortality conditions of a given
64 year will be repeated year after year in the future—may seem unlikely, the seasonality of mortality makes
65 the indefinite repetition of the mortality conditions in any fraction of a year not just unlikely but
66 impossible.⁷ When referring to mortality conditions not only in a short period but also in a small area, the
67 assumption on which PLEB estimates are built resort to a Groundhog-Day-like⁸ time loop repeating itself
68 in a small area from which individuals are unable to leave.

69 For a both interpretable and scalable, over space and time, measure of the impact of changing
70 mortality conditions on the individual lifespan, I suggest an alternative to the reduction in PLEB, the
71 Mean Unfulfilled Lifespan (MUL). Making no assumption about future mortality, the MUL translates

72 past changes in mortality into average difference in the length of lived lives. That difference is obtained
73 by comparing the actual average age at death during a given period and the expected ages at death of the
74 same individuals in the absence of mortality changes, whether induced by a specific cause of death or by
75 an event affecting multiple causes of death. The MUL can thus be estimated for populations of any size
76 and for periods of any length, from data on excess deaths or cause-specific deaths, as illustrated here with
77 COVID-19 mortality data from 159 national populations and 122 sub-national populations in Italy,
78 Mexico, Spain and the USA, for each sex and each quarter in 2020. The MUL also equals the product of
79 (1) the proportion of deaths in the population during a reference period that are due to a specific cause or
80 event and (2) the average reduction in length of life among individuals who died from that specific cause
81 or due to that specific event in the population during the reference period. In the case of COVID-19, I
82 show that for a given population the value of that average reduction only changes very slowly over time,
83 providing an easy short-cut for fine-grained tracking of the pandemic.

84 *Conceptual Detour*

85 Assessing PLEB reductions induced by a specific cause or event requires two period life tables, one
86 representing the prevailing mortality conditions and another one representing the counterfactual mortality
87 conditions expected in the absence of that cause or event. The assessment involves a relatively copious
88 amount of life table manipulations, but decades ago Nathan Keyfitz provided most useful insights as to
89 what these manipulations boil down to. Considering the related issue of estimating the increase in PLEB
90 brought by the permanent elimination of a cause of death, he summarized that the increase “depends on
91 the average time that elapses before the persons rescued will die of some other cause.”⁹ Conversely, the
92 decrease induced by a new cause of death depends on the average time that would have elapsed before the
93 persons who died from the new cause would have died from other causes.

94 This average time can be derived from the synthetic cohort approach modelled in the period life
95 table where each death at age a from a cause C , $d^C(a)$, reduces the number of person-years lived by the
96 life expectancy at age a in the absence of that new cause, $e^{o-C}(a)$. This assumes that persons dying from
97 the new cause would have had the same life expectancy in the absence of that cause as same-age persons

98 who survived that cause. This common assumption may appear unlikely, and interactions between causes
99 of death can be incorporated instead, but the data requirements are substantial. Under the common
100 assumption, the difference in PLEB is thus the average over all members of the synthetic cohort, l_0 (the
101 radix of the life table), of the difference in person-years lived by cohort members:

$$\Delta PLEB = \frac{1}{l_0} \cdot \int_0^{\omega} (d^C(a) \cdot e^{o-C}(a)) da$$

102 Keyfitz' insight relates to the concept of "potential years of life lost"¹⁰ developed a couple of
103 decades earlier still. The initial approach, designed to measure premature mortality, compared ages at
104 death to a fixed value (70 or 75 years).¹¹ This approach is not suited to study cause of deaths at older ages
105 since deaths at ages above the fixed value are not considered.¹² In burden-of-disease assessments, it has
106 become customary to estimate Years of Life Lost (YLL) as:

$$YLL = \int_0^{\omega} (D^C(a) \cdot e^{o^*}(a)) da$$

107 where $D^C(a)$ is the number of deaths from a certain cause C at age a observed in the population during a
108 reference period and $e^{o^*}(a)$ is life expectancy at age a in a counterfactual life table. YLL to COVID-19
109 have been estimated using this approach.^{13, 14, 15}

110 Three differences between the two above equations can be observed. First, YLL are estimated
111 from actual numbers of deaths by age rather than from numbers of life table decrements. This implies that
112 the YLL estimate is sensitive to the age distribution of the population, as in turn it affects the distribution
113 of deaths by age. While the PLEB is not an age-standardized measure,¹⁶ it equals the inverse of the
114 "stationary" death rate, that is, a weighted average of the period age-specific death rates with weights
115 derived from these death rates through life table construction. Using these internally-derived weights
116 rather than an external, standard age distribution, the stationary death rate and PLEB are independent of
117 the actual age composition of the population. This relative advantage of the difference in PLEB comes at
118 the cost of using a "stationary" age distribution of deaths, however, represented by the life table
119 decrements that result from indefinitely subjecting the population to the mortality conditions of the

120 period. As discussed in the introduction, this assumption is precisely what is problematic for studying the
121 impact of a mortality shock or an emerging disease such as COVID-19. Moreover, the actual distribution
122 of deaths can be used in a population of any size and periods of any length, allowing for the mortality
123 impact to be tracked on short temporal and small spatial scales for which interpreting differences in PLEB
124 hardly makes sense.

125 The second difference refers to the counterfactual life expectancies. In global burden-of-disease
126 assessments, a universal life table representing optimal survival conditions is typically used. This has the
127 advantage of making YLL for different populations additive allowing for the derivation of a global
128 estimate of YLL due to a cause by simple summation. However, using a universal life table may
129 misrepresent the actual gains from averting a death in a specific population.

130 The last difference concerns the denominator, or lack thereof in YLL. The use of a denominator
131 in the difference in PLEB allows for relating a total number of years in a population, measured by YLL,
132 to a number of persons, and thus for a more intuitive interpretation as an average difference in years lived
133 per person. Two ratios involving YLL can be found in the literature. First, the average YLL (AYLL)
134 relates YLL to all the deaths from that cause in the population during the reference period, D^C :

$$AYLL = \frac{\int_0^{\omega} (D^C(a) \cdot e^{o^*(a)}) da}{D^C} = \int_0^{\omega} \left(\frac{D^C(a)}{\int_0^{\omega} D^C(a) da} \right) \cdot e^{o^*(a)} da$$

135 The AYLL thus represents the average (universal) life expectancy left to population members who died
136 from the specific cause during a given period. On the one hand, it is a coherent measure as its
137 denominator includes all the deaths that contribute to lost years in the numerator. On the other hand, it is
138 only a function of the distribution of deaths by age, irrespective of the prevalence of that cause of death.
139 The AYLL thus cannot provide a measure of the intensity of a mortality shock. The second measure is
140 YLL per capita. While it does depend on the prevalence of the different causes of death, it is a less
141 coherent measure as it includes in the denominator all the individuals in the population, including many
142 that survived the mortality shock and do not contribute to YLL in the numerator. In turn, this complicates
143 providing a precise interpretation for the value of YLL per capita.

144 Considering the advantages and limitations of the extant measures, I propose to add one measure
 145 of the mortality impact of a cause of death or an event on the individual lifespan. This measure, the Mean
 146 Unfulfilled Lifespan (MUL), is intended for situations where the underlying assumptions of PLEB might
 147 be implausible, and thus based, as the estimate of YLL, on actual numbers of deaths in a population
 148 during a period rather than on life table decrements. To retain its intuitive interpretation, however, the
 149 MUL is structured like the difference in PLEB as summarized by Keyfitz and, using counterfactual life
 150 expectancies representing the mortality conditions in the population of interest, similarly expressed as an
 151 average difference in person-years lived per person. Since the life table radix, l_0 , equals the sum of all
 152 decrements at all ages, the structural equivalence is maintained by defining the MUL as:

$$MUL \stackrel{\text{def}}{=} \frac{1}{D} \cdot \int_0^{\omega} (D^C(a) \cdot e^{o-C}(a)) da$$

153 where D is the total number of deaths (from all causes at all ages) during the reference period.

154 This intuitive interpretation of the MUL can be derived by rewriting this defining equation as:

$$\begin{aligned} MUL &= \frac{\int_0^{\omega} \left((D^C(a) \cdot (a + e^{o-C}(a) - a)) + (D^{-C}(a) \cdot (a - a)) \right) da}{\int_0^{\omega} (D^C(a) + D^{-C}(a)) da} \\ &= \frac{\int_0^{\omega} \left((D^C(a) \cdot (a + e^{o-C}(a))) + (D^{-C}(a) \cdot a) \right) da}{\int_0^{\omega} (D^C(a) + D^{-C}(a)) da} - \frac{\int_0^{\omega} ((D^C(a) + D^{-C}(a)) \cdot a) da}{\int_0^{\omega} (D^C(a) + D^{-C}(a)) da} \end{aligned}$$

155 where $D^{-C}(a)$ is the number of deaths from all causes but C at age a observed in the population during the
 156 reference period. The second term represents the average age at death in a given period. If we assume that
 157 individuals who die of other causes than C die at the same age as they would have in the absence of cause
 158 C (no indirect effect of cause C on other causes of death), and that individuals who die of cause C at age a
 159 would have otherwise lived to age $a+e(a)$, the first term represents the average expected age at death of
 160 the same individuals in the absence of cause C . The MUL is thus the difference between the average age
 161 at death in the population during a given period and the average expected age at death of the same
 162 individuals in the absence of cause C .

163 The assumption that cause C has no indirect effect on other causes of death is actually not
164 required. As will be shown below, one can also derive the MUL from data on all deaths, and
165 distinguishing between deaths that, based on counterfactual “benchmark” mortality conditions, were
166 expected to occur in that period and those that were not, use the latter number of “excess” deaths instead
167 of deaths from a specific cause. Finally, note that the MUL differs from changes in average ages at death
168 across periods, which can be readily measured but may actually be positive with the emergence of a new
169 cause of death if that cause affects people who are older on average than those dying from other causes.
170 To sum up this conceptual detour, the MUL thus complements existing indicators of the impact on the
171 individual lifespan of a cause of death by providing a measure of the average potential years of life lost to
172 a specific cause of death, or to any mortality shock, for all population members dying in a certain period,
173 regardless of their cause of death.

174 *Empirical Shortcut*

175 Calculating the MUL for small geographical areas and short periods is not conceptually problematic since
176 it captures the actual length of lives that ended there and then, unlike differences in life expectancies that
177 require assumptions about future conditions. The demand on data (including a separate counterfactual life
178 table for each population of interest) is substantial, however, and the life table manipulations are not
179 particularly straightforward.

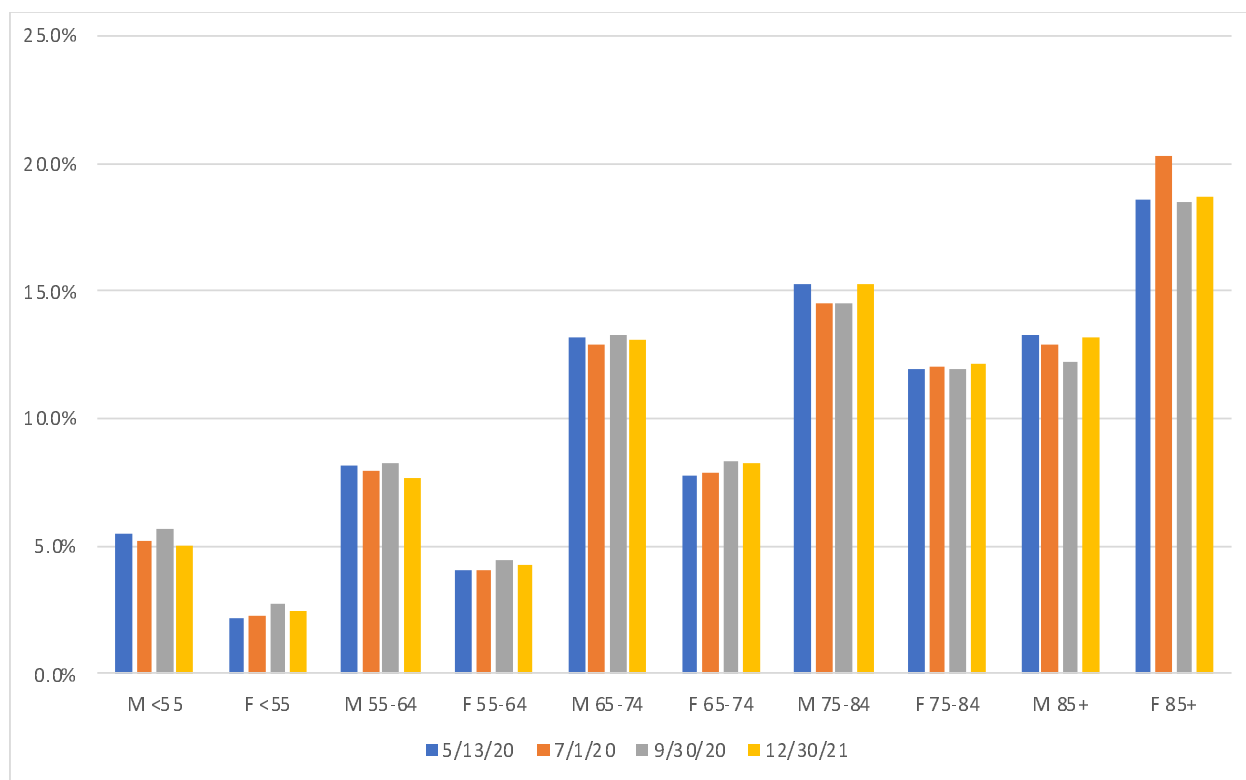
180 To simplify the estimation, the MUL can be rewritten as:

$$MUL = \frac{D^C}{D} \cdot \frac{\int_0^{\omega} D^C(a) \cdot e^{o-C}(a) da}{D^C} = \frac{D^C}{D} \cdot PAYLL$$

181 The second term is similar to the AYLL in burden-of-disease assessments, but again based on population-
182 specific, counterfactual life expectancies instead of universal ones, and to underscore the difference is
183 termed here the Population AYLL (PAYLL). In a given population, the PAYLL is a weighted average of
184 counterfactual life expectancies that are estimated from prior conditions and do not change over time.
185 Meanwhile, the weights are the ratios:

$$\frac{D^c(a)}{D^c} = \frac{M^c(a) \cdot N(a)}{M^c \cdot N} = \frac{M^c(a)}{M^c} \cdot \frac{N(a)}{N}$$

186 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 N and
 187 $N(a)$ are the total population size and number of individuals at age a in the population. These weights
 188 should also be expected to vary little within short periods because their values depend on the population
 189 composition, $N(a)/N$, and on the age pattern of cause-specific death rates, $M^c(a)/M^c$, both of which should
 190 vary little within short periods. To confirm this, Figure 1 compares the distribution of provisional
 191 COVID-19 death counts in the USA at four points in time: on May 13 (one of the earliest dates for which
 192 this distribution is available) and on July 1st, September 30th and December 30th.



193
 194 Fig 1: Distribution of Provisional COVID-19 Death Counts, by Sex- and Age-Groups, USA as of 5/13
 195 (54,860 deaths), 7/1 (112,223 deaths), 9/30 (194,087 deaths) and 12/30 (301,671 deaths)

196 Source: Centers for Disease Control and Prevention (CDC)

197 This suggests that the value of PAYLL can be expected to only change slowly over time and to be
 198 relatively close across populations with similar life expectancies and population compositions. In

199 particular, MUL values for a sub-population or during a sub-period can be approximated as the product of
200 the PAYLL for the whole population or the entire period and the variable all-age ratio of deaths from a
201 specific cause to all deaths, D^C/D , in the sub-population or during that sub-period. As noted above, this
202 can be readily extended to any mortality shock, with data on excess deaths and all-cause deaths.

203 **Materials and Methods**

204 The equations defining YLL and the MUL look deceptively simple. To implement them, one has to apply
205 estimates of life expectancies, which refer to exact ages, to numbers of deaths that are available or can be
206 estimated only for age intervals. The varying value of life expectancy on a closed age interval is typically
207 approximated by linear interpolation,¹⁷ in which case the contribution of the closed age interval between
208 ages x and $x+n$ to the MUL equals:

$$\frac{\int_0^n D^C(x+a) \cdot e^{o-c}(x+a) da}{D} = \frac{{}_n D_x^C}{D} \cdot e^{o-c}(x + {}_n a_x^C)$$

209 where ${}_n D_x^C$ is the number of individuals between ages x and $x+n$ who died of cause C during the reference
210 period and ${}_n a_x^C$ is the average number of years lived after age x by these individuals. In turn, life
211 expectancy at exact age $x+{}_n a_x^C$ can be derived by linear interpolation between the values of life
212 expectancy at ages x and $x+n$.

213 The linear interpolation is more problematic for wider the age intervals and for older age groups.
214 In the 2018 US life table for males for instance,¹⁸ life expectancy declines by 2.9 years between ages 75
215 and 80 and by 2.4 years between ages 80 and 85. In this case, the linear approximation over-estimate life
216 expectancy in the interval. The average number of years lived after age 75 by individuals dying between
217 ages 75 and 85 is 5.4 years, and linear interpolation would yield a life expectancy at age 80.4 years of 8.5
218 years, whereas life expectancy has already dropped to 8.4 years at age 80. This upward bias is particularly
219 undesirable in the case of COVID-19 deaths at those ages because COVID-19 victims are more likely to
220 suffer from other long-term conditions that more likely reduced their life expectancies compared to same-
221 age individuals. Even more problematic with linear interpolation is the open-ended interval, as it requires
222 that an arbitrary upper age limit be set. Unfortunately, these couple of issues may concern a large share of

223 COVID-19 deaths: as shown in Figure 1, close to 60% of US deaths are above age 75 years and reported
224 in just one ten-year closed interval (75 to 84 years) and one open age interval (over 85 years).

225 In this respect, working with all deaths in a given period rather than with deaths from a specific
226 cause of interest presents an important empirical advantage in addition to the benefit of assessing both the
227 direct and the indirect effects of that cause of death. Using all deaths in a closed age interval between ages
228 x and $x+n$, ${}_nD_x$, one may first calculate the contribution to the MUL of the deaths that were expected to
229 occur under the counterfactual mortality conditions, ${}_nD_x^C$. For these deaths, the difference in length of life
230 averages to the difference between the average number of years lived in the age interval under the
231 prevailing conditions, ${}_na_x$, and under the counterfactual conditions, ${}_na_x^C$. The advantage of this approach
232 is that setting an arbitrary upper value for the open-ended age interval becomes unnecessary because,
233 having already reached age N , individuals who contributed to the number of deaths in the open age
234 interval would all have been expected to die in the same open-ended age interval under the counterfactual
235 mortality conditions, albeit possibly not in that period. For all the open-interval deaths over age N , D_{N+} ,
236 the difference in length of life thus averages to the difference between life expectancy at age N under the
237 prevailing conditions, $e^o(N)$, and under the counterfactual conditions, $e^{o-C}(N)$.

238 This approach limits the issue of estimating the average reduction in length of lived lives on
239 closed age intervals from values of life expectancies that vary on these age intervals to “excess” deaths,
240 ${}_nD_x - {}_nD_x^C$. An alternative to linear interpolation derives from the fact that an individual’s expected length
241 of life ($e^o(x)+x$) gradually increases with age x . Life expectancy at any age is thus larger than the
242 difference between life expectancy at an earlier age and the difference between the two ages:

$$\int_0^n D(x+a) \cdot e^o(x+a) da > \int_0^n D(x+a) \cdot (e^o(x) - a) da = {}_nD_x \cdot (e^o(x) - {}_na_x)$$

243 Applying this approximation to excess deaths on any closed age interval will induce some
244 underestimation of the average life expectancy of individuals dying on the interval, which appears
245 preferable in this case to the overestimation induced by linear interpolation.

246 Adding the contributions of the different types of deaths yields:

$$\begin{aligned}
 MUL \cong \frac{1}{D} & \left(\sum_{x=0, n}^{N-n} {}_n D_x^{-c} \cdot ({}_n a_x^{-c} - {}_n a_x) + (D_{N+} \cdot (e^{o-c}(N) - e^o(N))) \right) \\
 & + \sum_{x=0, n}^{N-n} ({}_n D_x - {}_n D_x^{-c}) \cdot (e_x^{o-c} - {}_n a_x)
 \end{aligned}$$

247 Rearranging the sums corresponding to the closed age intervals and using the multiple-decrement life
 248 table relationship $e^o(N)/e^{o-c}(N) = D_{N+}^{-c}/D_{N+}$, this can be rewritten as:

$$MUL \cong \frac{1}{D} \left(\sum_{x=0, n}^{N-n} ({}_n D_x \cdot ({}_n a_x^{-c} - {}_n a_x) + ({}_n D_x - {}_n D_x^{-c}) \cdot (e_x^{o-c} - {}_n a_x^{-c})) + ((D_{N+} - D_{N+}^{-c}) \cdot e_N^{o-c}) \right)$$

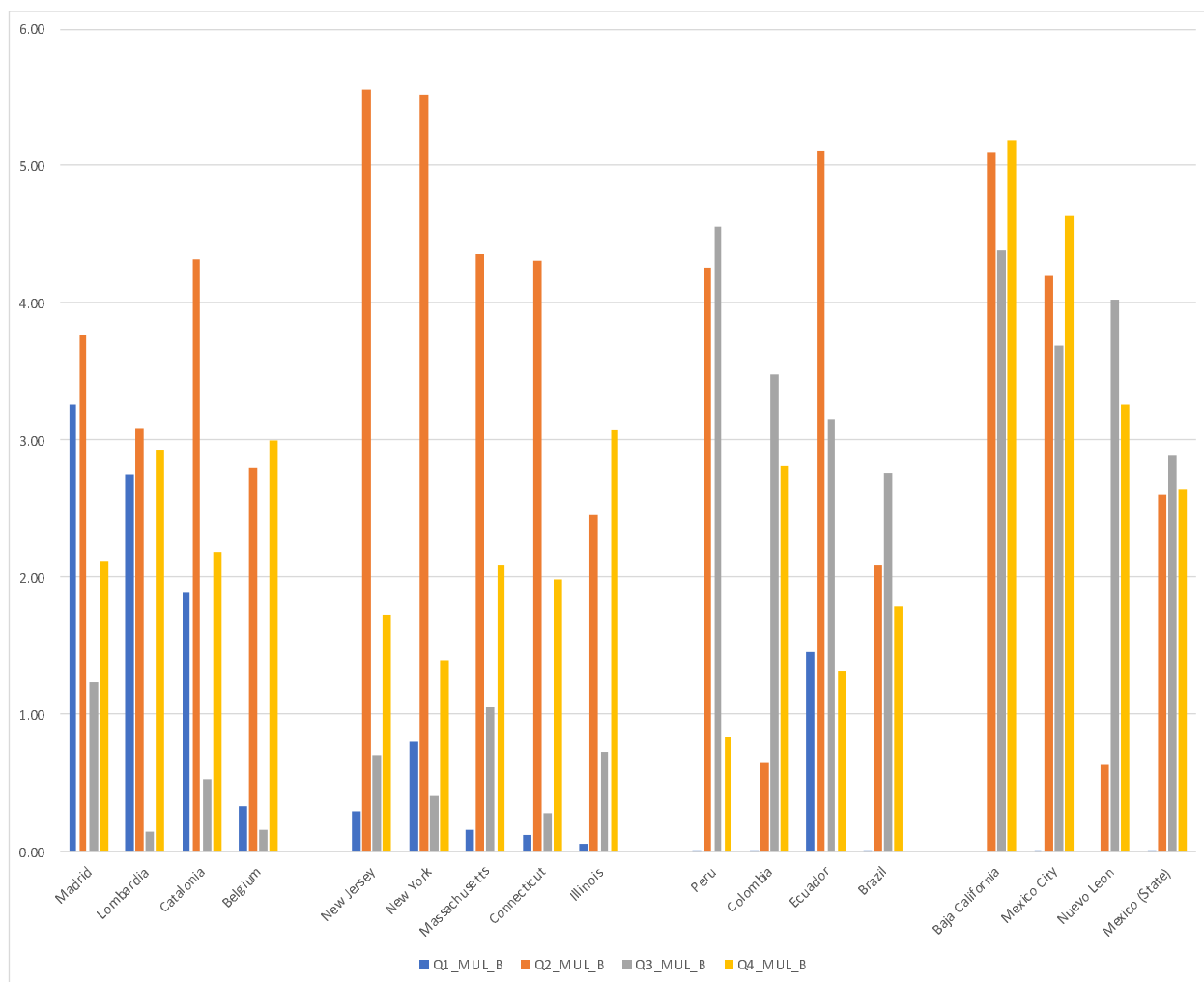
249 Using this approximation to estimate the impact of COVID-19 on the individual lifespan in each
 250 quarter of 2020 first requires a life table representing survival conditions in 2020 in the absence of
 251 COVID-19 whose values of e_x^{o-c} and ${}_n a_x^{-c}$ can be used. Combined with the number of individuals by sex
 252 and age-group, the life table values of age-specific death rates, ${}_n m_x^{-c}$, then provide the expected numbers
 253 of deaths ${}_n D_x^{-c}$ in the absence of COVID-19. Population data and life table functions for countries were
 254 obtained from the UN Population Division.¹⁹ Corresponding data for sub-populations in Italy, Spain, and
 255 the US were obtained from national statistical agencies.^{20, 21, 22, 23}

256 New life tables representing actual mortality conditions (with COVID-19) in each quarter must
 257 then be derived to calculate the corresponding values of ${}_n a_x$. The construction of these life tables requires
 258 the quarterly numbers of deaths by sex and age-group. In countries where vital statistics are incomplete
 259 not available yet, but estimates of COVID-19 deaths are available, the total number of deaths, ${}_n D_x$, can be
 260 obtained by adding these estimates (through a multi-decrement life table to adjust for competing risks of
 261 deaths, which typically assumes no indirect effects on other causes of death)²⁴ to the expected numbers of
 262 deaths in the absence of Covid-19 (${}_n D_x^{-c}$). When COVID-19 estimates are not broken down by sex and
 263 age-group, an alternative is to use a reference set of age-and-sex death rates from COVID-19 from
 264 another population for which these rates are deemed reliable.²⁵ Centers for Disease Control and
 265 Prevention (CDC) data provided the reference set of age-and-sex death rates from COVID-19.²⁶ Estimates

266 of COVID-19 deaths by March 31, June 30, September 30 and December 31, 2020 were taken from the
267 IHME.²⁷ All of these data were downloaded from institutional websites.

268 **Results**

269 Figure 2 shows MUL values for selected national and sub-national populations in each of the four
270 quarters of 2020 (for both sexes combined). First-quarter MUL values illustrate substantial impacts of
271 COVID-19 on longevity already in parts of Spain (Madrid, 3.26 years) and Italy (Lombardy 2.76 years).
272 In the second quarter, the impact continued to increase in these two nations, with the MUL reaching 4.32
273 years in Catalonia for instance, but also in other European nations such as Belgium (2.80 years).
274 However, the second-quarter MUL values reached their highest values in parts of the USA (New Jersey,
275 5.56 years, and New York, 5.52 years). Very high second-quarter values are also estimated for parts of
276 South America (Ecuador, 5.12 years) and Mexico (Baja California, 5.10 years).



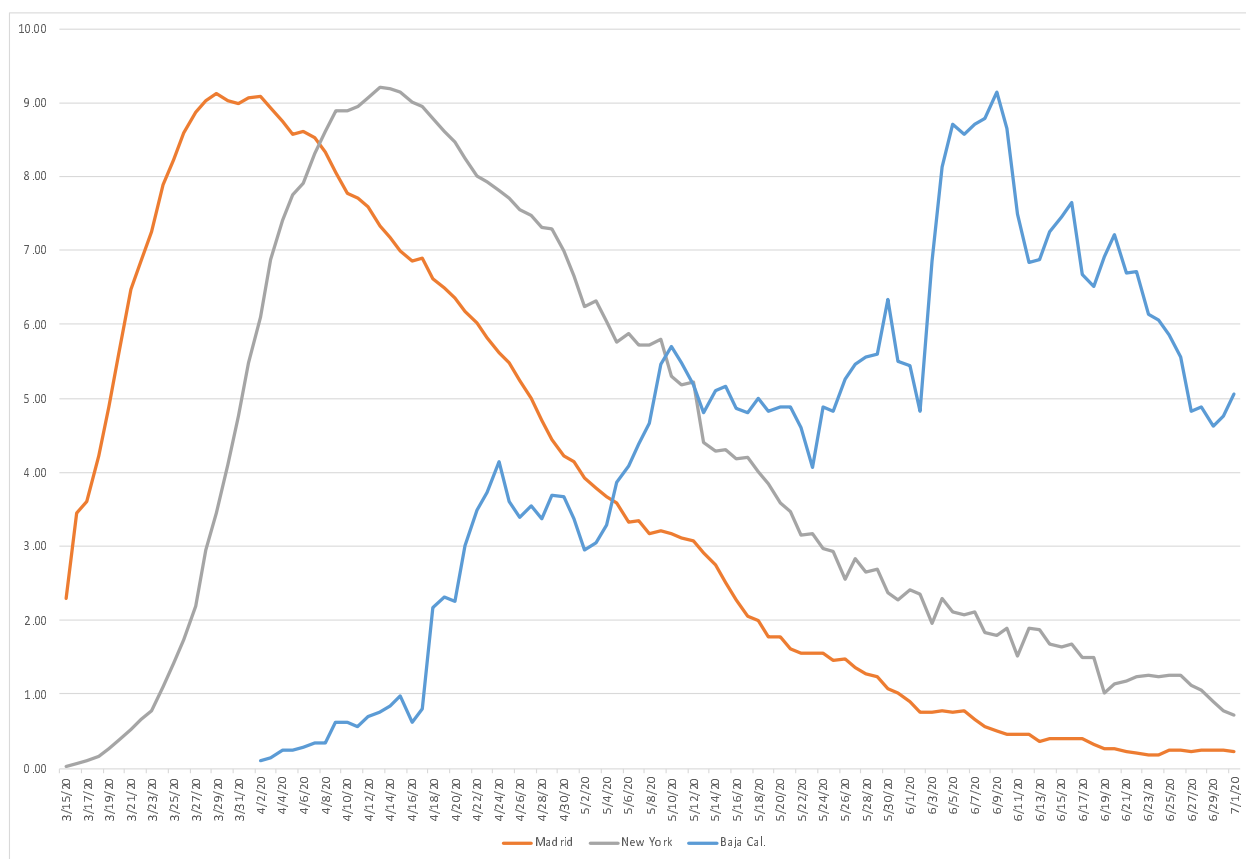
277

278 Fig 2: Quarterly Mean Unfulfilled Lifespan (MUL) for both sexes, in years, selected populations
 279 (populations with 5,000 or more COVID-19 deaths by January 1st, 2021, and the highest annual MUL
 280 values for 2020).

281 In the third quarter, MUL values continue to increase or remain high in parts of South America
 282 (Peru, 4.56 years) and Mexico (Baja California, 4.39 years, and Nuevo Leon, 4.03 years), whereas they
 283 were declining in Europe and the USA. In the last quarter of 2020, MUL values decline in South
 284 America, whereas the highest values remain for parts of Mexico (Baja California, 5.19 years, and Mexico
 285 City, 4.64 years). Fourth-quarter MUL values also rebound in parts of Europe (Belgium, 3.00 years) and
 286 the USA (Illinois, 3.07 years). MUL values for men are typically higher than for women, and the average
 287 age at death of men who died in New Jersey during the second quarter was nearly 6 years younger than

288 their average expected age at death without COVID-19 (5.97 years, compared to 5.14 years for women,
289 full results shown in supplementary file, Table S1).

290 The estimated quarterly PAYLL values (see supplementary file, Table S1) can be used to apply
291 the suggested empirical short-cut. These values range from a low of 9.7 years in Bulgaria to a high of 26.4
292 years in Qatar. These differences can be explained by different age compositions, with younger
293 compositions giving more weight to remaining life expectancies at younger ages, which are obviously
294 higher. As expected, however, these values change relatively little from one quarter to the next for a given
295 population. Figure 3 shows peak MUL values for a rolling seven-day period, derived from quarterly
296 PAYLL values, were reached between mid-March and the end of June in Madrid (9.12 years on 3/29), in
297 New York (9.20 years on 4/13) and in Baja California (9.15 years on 6/9).



298
299 Fig 3: Mean Unfulfilled Lifespan (MUL) for both sexes, by rolling seven-day periods, in years.

300 The same approximation but for sub-populations is illustrated by focusing on the province of
301 Guayas in Ecuador. Data on the monthly number of deaths by province show the marked increase in
302 March, April and May from a baseline of 1,700-2,000 per month in January, February and again in June.²⁸
303 In April, the number of deaths reached 12,004, of which the January-February-June average suggests only
304 17.2% might be estimated to be from causes other than COVID-19 (without adjustment for competing
305 causes). No specific life table is available to estimate MUL values for the province directly, but based on
306 the second-quarter PAYLL derived for Ecuador, individuals appeared to have died 12.6 years younger, on
307 average, than their expected age at death in April in the province of Guayas.

308 **Discussion**

309 The MUL is proposed here as an alternative to induced changes in PLEB to assess the impact of a cause
310 of death of the individual lifespan in situations where the assumptions underlying life table construction
311 are implausible, invalidating the usual interpretation of the PLEB. Complementing existing measures
312 such as the AYLL and the YLL per capita, the MUL is similarly based on estimating the number of
313 potential years of life lost corresponding to deaths from a specific cause or excess deaths from a specific
314 event in a given period. As shown above, by averaging estimated potential years of life lost over the total
315 number of deaths in the period, the MUL is structured like a difference in PLEB. The MUL retains the
316 intuitive interpretation of a difference in PLEB as a change in average length of life, but for an actual
317 cohort of individuals (those dying in the period) subjected to prevailing mortality conditions rather than
318 for a synthetic cohort subjected to indefinitely constant conditions.

319 To illustrate the derivation and interpretation of the MUL in the context of COVID-19 mortality,
320 MUL values were derived for a total of 281 populations for each sex and each quarter of 2020. As argued
321 above, the MUL is best estimated from data on numbers of excess deaths by age and sex in a given
322 period, estimated from data on all-cause mortality. At this writing, however, these data are not available
323 from the vast majority of the national and subnational populations around the world. Instead, total
324 numbers of quarterly COVID-19 deaths in each of these populations were used and their age-and-sex
325 breakdown derived with the assumption of a shared age-and-sex COVID-19 mortality “pattern” (i.e., the

326 same ratio of age-and-sex-specific death rates to the all-age, both-sex death rate in all populations).

327 Excess deaths by age and sex were then derived through life table adjustments for competing causes of

328 death (as COVID-19 deaths reduced exposure to other causes of deaths).

329 The adjustment for competing causes of death, as well as the estimate of potential years of life
330 lost corresponding to each excess death is based on the assumption that, in the absence of a certain cause
331 of death, individuals who died from that cause would have faced the same risk of death from other causes
332 as any individual of the same age and sex. As discussed above, this standard assumption of YLL-based
333 measures is problematic in the case of COVID-19 due to the higher proportion of several underlying long-
334 term conditions (e.g., obesity) observed among COVID-19 victims. Adjusting for differences in long-
335 term conditions prevalence is very data demanding, however, and the impact might not be as large as
336 expected, a study finding that performing this adjustment reduces the AYLL for COVID-19 in the United
337 Kingdom was from 13 to 12 years (average for both sex).²⁹ An alternative strategy was proposed here that
338 underestimates the value of potential years of life lost over an age interval by estimating its value at the
339 beginning of the age interval. As shown in the supplementary files, this yields quarterly PAYLL values
340 for the United Kingdom that vary between 11.9 and 12.0 years (Table S1), values thus quite close to what
341 the adjustment for underlying long-term conditions might have provided. While there is of course no
342 guarantee that this strategy would apply equally well to causes of death other than COVID-19, ignoring
343 comorbidities will likely lead to some overestimation of AYLL for most causes of death, and purposely
344 underestimating the number of potential years of life lost will likely be preferable in other situations as
345 well.

346 The MUL estimates derived here can also be improved by using actual age-and-sex breakdowns
347 of COVID-19 deaths in populations of interest, which are becoming available for an increasing number of
348 nations.³⁰ Here too, however, the actual impact on MUL estimates might not be as large as expected. The
349 age patterns of CoViD-19 death rates available so far similarly exhibit remarkable regularities, with some
350 modest variation in the slope of these age patterns at the oldest ages, probably due to the number of
351 fatalities in nursing homes across Europe and the USA.³¹ Considering the two countries with the largest

352 number of COVID-19 deaths in 2020, substituting the actual sex and age pattern in Brazil to the one
353 derived by assuming the same mortality pattern as in the USA only changes the estimated reduction in
354 2020 PLEB for Brazil from 1.72 to 1.67 years.³² The structural similarity between the MUL and
355 differences in PLEB suggest that using population-specific data on COVID-19 mortality by age and sex
356 might only improve MUL estimates by a similarly modest order of magnitude.

357 To date, by far the strongest reason to caution against taking these MUL estimates at face value
358 originates in lacking or incomplete data on the total number of deaths in 2020 for many populations and
359 how this number might eventually relate to the number of reported COVID-19 deaths. Even in countries
360 with reliable vital statistics, annual deaths data might not be finalized until several months into the
361 following year. Moreover, the estimation of excess deaths is performed by comparison with benchmark
362 mortality rates to represent counterfactual mortality conditions, and the results can be quite sensitive to
363 the choice of a benchmark period (e.g., the last five years) and the estimation methodology (e.g.,
364 assumptions about mortality trends).³³ In a study of 21 high-income nations,³⁴ the official number of
365 COVID-19 deaths from mid-February to end of May was found to be lower than the lower bound of the
366 95% credible interval for excess deaths during that period in 5 countries. The discrepancy was highest in
367 Spain, with the credible interval from 47% to 91% more excess deaths than reported COVID-19 deaths.
368 In the USA, between January 26th and October 3rd, excess deaths were also estimated to be 38% higher
369 than the COVID-19 death tally over the same period.³⁵ Unfortunately, these ratios exhibit insufficient
370 empirical regularities across populations (and possibly over time) to uniformly adjust numbers of reported
371 COVID-19 deaths across populations for which the number of excess deaths remains unknown at this
372 point.

373 More generally, another advantage of the MUL is that it remains interpretable regardless of its
374 temporal and geographical scale, pertaining to deaths in a given population in a given period, whereas we
375 argued the PLEB becomes difficult to interpret as a short-term, micro-level indicator. With the PAYLL
376 values provided in supplementary file (Table S1), MUL values can easily be approximated as the product
377 of the corresponding PAYLL value and the ratio of COVID-19 to total deaths in the population during a

378 given period. The approximation can be used for short-duration periods as long as PAYLL can be
379 assumed to be almost constant over time and quarterly PAYLL estimates are found indeed to change little
380 from one quarter to the next. This approximation can also be used for sub-populations for which the
381 necessary data are only provided of the entire population, but only when the age compositions of the sub-
382 populations can be held as relatively similar.

383 This relates more generally to the fact that, like YLL-based measures, the MUL is not sex- or
384 age-standardized and MUL comparisons across populations that differ markedly in age composition will
385 be biased. On the one hand, all else equal, a younger population composition yields a younger distribution
386 of deaths and a higher PAYLL value. On the other hand, known variations of COVID-19 mortality with
387 age³⁶ suggest that an older population distribution contributes to a higher proportion of COVID-19 deaths
388 relative to all deaths. Only by accident would the two, opposite age-composition effects cancel each other
389 and in general the MUL is not a standardized measure. This represents one disadvantage compared to
390 differences in PLEB. However, this advantage over the MUL again comes at a cost, since the internal
391 derivation assumes the period mortality conditions will become permanent.

392 When this assumption is not tenable, the MUL provides an unstandardized alternative to
393 differences in PLEB, similarly structured as an average difference in length of lives lived per person.
394 Related to other unstandardized measures such as the AYLL or YLL per capita, the MUL's interpretation
395 pertains to an actual death cohort, that is, population members who died during a certain period rather
396 than to a synthetic cohort as represented in the life table. To reiterate, the MUL indicates the difference
397 between their average age at death and their average expected age at death had a temporary mortality
398 shock not occurred, and its measurement requires no assumption that these temporary conditions will
399 either pass entirely or extend indefinitely into the future.

400

401 **Acknowledgments**

402 The author benefited from facilities and resources provided by the California Center for
403 Population Research at UCLA (CCPR), which receives core support (P2C-HD041022) from the
404 Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD).
405 The author thanks Michel Guillot, Philippe Bocquier, Tim Riffe, Michel Garenne, Christophe
406 Guilmoto and Sam Preston for comments on an earlier draft of this manuscript.
407

408 **References**

¹ Beltrán-Sánchez H, Preston SH & Canudas-Romo V. An integrated approach to cause-of-death analysis: Cause-deleted life tables and decompositions of life expectancy. *Demographic Research* 2008;19: 1323-1350. doi: 10.4054/DemRes.2008.19.35

² Ho J. The contribution of drug overdose to educational gradients in life expectancy in the United States, 1992–2011. *Demography* 2017;54: 1175-1202. doi: 10.1007/s13524-017-0565-3

³ Goldstein JR & Lee RD. Demographic perspectives on the mortality of COVID-19 and other epidemics. *Proc. Natl Acad. Sci.* 220; 117(36): 22035-22041. doi: 10.1073/pnas.2006392117

⁴ Ghislandi S et al. 2020. News from the Front: Estimation of excess mortality and life expectancy in the major epicenters of the COVID-19 pandemic in Italy. *MedRxiv [preprint]* 2020. doi: 10.1101/2020.04.29.20084335

⁵ Trias-Llimos S, Riffe T & Bilal U. Monitoring life expectancy levels during the COVID-19 pandemic: Example of the unequal impact of the first wave in Spanish regions. *PLOS One* 2020. doi: 10.1371/journal.pone.0241952

⁶ Aburto JM, Kashyap R, Schöley J et al. 2020. Estimating the burden of COVID-19 on mortality, life expectancy and lifespan inequality in England and Wales: A population-level study. *MedRxiv [preprint]* 2020. doi: 10.1101/2020.07.16.20155077

⁷ Ho T & Noymer A. Summertime, and the livin' is easy: Winter and summer pseudoseasonal life expectancy in the United States. *Demographic Research* 2017; 37(45): 1445-1476. doi: 10.4054/DemRes.2017.37.45

⁸ Ramis H (director and writer) & Rubin D (writer). *Groundhog day*. [Film] 1993. Washington, D.C., United States: National Film Registry.

⁹ Keyfitz N. What difference would it make if cancer were eradicated? An examination of the Taeuber paradox. *Demography* 1977; 14: 411-418. <https://link.springer.com/content/pdf/10.2307/2060587.pdf>

¹⁰ Dempsey M. Decline in tuberculosis - The death rate fails to tell the entire story. *Am Rev Tuberc.* 1947; 56: 157-164.

¹¹ OECD. Potential years of life lost (indicator). [Cited February 18, 2021]. doi: 10.1787/193a2829-en

¹² Mitra AK, Payton M, Kabir N, et al. Potential years of life lost due to COVID-19 in the United States, Italy, and Germany: An old formula with newer Ideas. *Int. J. Environ. Res. Public Health* 2020; 17: 4392. doi: 10.3390/ijerph17124392

¹³ Mohanty SK, Dubey M, Mishra US & Sahoo U. Impact of COVID-19 attributable deaths on longevity, premature mortality and DALY: Estimates of USA, Italy, Sweden and Germany. *MedRxiv [preprint]* 2020. doi: 10.1101/2020.07.06.20147009

¹⁴ Wetzler HP, Wetzler EA & Cobb HW. COVID-19: How many years of life lost? *MedRxiv [preprint]* 2020. doi: 10.1101/2020.06.08.20050559

¹⁵ Pifarré i Arolas H, Acosta E, López Casasnovas, G et al. Global years of life lost to COVID-19. *MedRxiv [preprint]* 2020. doi: 10.1101/2020.06.19.20136069

¹⁶ Modig K, Rau R & Ahlbom A. Life Expectancy: What Does it Measure? *BMJ Open* 2020; 10. <https://bmjopen.bmj.com/content/10/7/e035932>

¹⁷ Aragon TJ, Lichtensztajn DY, Katcher BS et al. Calculating expected years of life lost for assessing local ethnic disparities in causes of premature death. *BMC Public Health* 2020; 8: 116. doi: 10.1186/1471-2458-8-116

¹⁸ Arias E & Xu JQ. United States Life Tables, 2018. *National Vital Statistics Reports* 2020; 69(12). Hyattsville, MD: National Center for Health Statistics.

¹⁹ United Nations, Department of Economic and Social Affairs, Population Division. World population prospects 2019, online edition. rev. [Cited June 28, 2020]. <https://population.un.org/wpp/Download/Standard/CSV/>.

²⁰ Istituto Nazionale di Statistica (Istat). Life tables. [Cited June 28, 2020].

<http://dati.istat.it/Index.aspx?QueryId=19053&lang=en>.

²¹ Instituto Nacional Estadística (INE). Mortality tables. [Cited June 28, 2020].

https://www.ine.es/dyngs/INEbase/en/categoria.htm?c=Estadistica_P&cid=1254734710984.

²² United States Census Bureau. Explore census data. [Cited May 11, 2020]. <https://data.census.gov>.

²³ Centers for Disease Control and Prevention (CDC). CDC wonder. [Cited June 28, 2020].

<https://wonder.cdc.gov/cmfi-icd10.html>.

²⁴ Preston SH, Heuveline P & Guillot M. Demography: Measuring and modeling population processes.

Malden, MA: Blackwell; 2001.

²⁵ Guilamoto CZZ. COVID-19 death rates by age and sex and the resulting mortality vulnerability of countries and regions in the world. MedRxiv [preprint] 2020. doi: 10.1101/2020.05.17.20097410

²⁶ Centers for Disease Control and Prevention. Provisional COVID-19 death counts by sex, age and state. [Cited December 31, 2020]. <https://data.cdc.gov/NCHS/Provisional-COVID-19-Death-Counts-by-Sex-Age-and-S/9bhg-hcku>.

²⁷ University of Washington, Institute for Health Metrics and Evaluation. COVID-19 projections. [Cited January 2, 2021]. <https://covid19.healthdata.org/>.

²⁸ Dirección General de Registro Civil, Identificación y Cedulación. Defunciones, enero-junio 2020. [Cited July 28, 2020]. <https://www.registrocivil.gob.ec/cifras/>

²⁹ Hanlon P, Chadwick F, Shah A et al. COVID-19 – Exploring the implications of long-term condition type and extent of multimorbidity on years of life lost: A modelling study. *Wellcome Open Res* 2020; 5: 75. doi: 10.12688/wellcomeopenres.15849.1

³⁰ Riffe T & Acosta E. COVerAGE-DB: A database of COVID-19 cases and deaths by age. [Cited December 28, 2020]. <https://osf.io/mpwjq/>.

-
- ³¹ O'Driscoll M, Ribeiro Dos Santos G, Wang L et al. Age-specific mortality and immunity patterns of SARS-CoV-2. *Nature* 2021; 590: 140-145. doi: 10.1038/s41586-020-2918-0
- ³² Heuveline P & Tzen M. Beyond deaths per capita: Comparative CoVID-19 mortality indicators. MedRxiv [preprint] 2021. doi: 10.1101/2020.04.29.20085506
- ³³ Aron J & Muellbauer J. A pandemic primer on excess mortality statistics and their comparability across countries. Our world data. [Cited February 18, 2020]. <https://ourworldindata.org/covid-excess-mortality>.
- ³⁴ Kontis V, Bennett JE, Rashid T et al. Magnitude, demographics and dynamics of the effect of the first wave of the COVID-19 pandemic on all-cause mortality in 21 industrialized countries," *Nat. Med.* 2020; 26: 1919-1928. doi: 10.1038/s41591-020-1112-0
- ³⁵ Rossen LM, Branum AM, Ahmed FB et al. Excess deaths associated with COVID-19, by age and race and ethnicity – United States, January 26 – October 3, 2020," *Morb. Mortal. Wkly Rep.* 2020; 69: 1522-1527. doi: 10.15585/mmwr.mm6942e2external icon.
- ³⁶ Dowd JB, Andriano L, Brazel DM, et al. Demographic science aids in understanding the spread and fatality rates of COVID-19. *Proc. Natl Acad. Sci.* 2020; 117: 9696-9698. doi: 10.1073/pnas.2004911117