Beyond Deaths per Capita:

Comparative CoViD-19 Mortality Indicators

Updated Version, 9/11/2020

Original Manuscript Submitted 4/29/2020, Published 5/5/2020 at:

https://www.medrxiv.org/content/10.1101/2020.04.29.20085506v1

Patrick Heuveline, Ph.D. and Michael Tzen, M.S.

California Center for Population Research (CCPR)

University of California, Los Angeles (UCLA)

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

Acknowledgment

The authors 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 authors thank Patrick Gerland for clarifications regarding UN demographic data, Hiram Beltrán-Sánchez for additional data for Mexico, and Sam Preston, Jason Kerwin and Piedad Urdinola for comments on an earlier draft.

Beyond Deaths per Capita:

Comparative CoViD-19 Mortality Indicators

Abstract

The number of CoViD-19 deaths is a lagged but more reliable indicator of the disease progression across populations than the number of confirmed cases. With substantial age and sex differences in CoViD-19 mortality, that number should be adjusted for the age-and-sex composition of the population as well as its total size. Following well-established practices in demography, this article discusses several period measures based on cumulative numbers of CoViD-19 deaths up to a point in time and illustrate them with weekly-updated data from 386 national and subnational populations. First is an unstandardized occurrence-exposure rate comparable to the Crude Death Rate. The peak estimate for New York exceeded the state most recent annual Crude Death Rate and remains the highest on record. Second, an indirectly standardized rate is shown to perform quite like a directly standardized rate but without requiring an age-and-sex breakdown of CoViD-19 deaths. Relative to the US, standardization lowers death rates in European populations and increases them in most other populations, with the highest standardized rates currently found in subnational Mexican and Peruvian populations. Last, projected end-of-the-year CoViD-19 death tallies can be translated into reductions in 2020 life expectancies, which could exceed two and a half years in Peru and Ecuador, and in various subnational entities from Baja California to Madrid and New York. To put these in perspective, the 1.5-year projected reduction for the US would bring 2020 life expectancy at birth to its lowest level since 2003 and induce its largest annual decline since World War II.

Beyond Deaths per Capita

Background

As of June 1st, deaths from the novel coronavirus disease 2019 (CoViD-19) had been reported in 186 of the 235 countries and territories of the United Nations system (UN). As with previous viruses whose carriers are mostly asymptomatic,[1] the number of deaths due to CoViD-19 provides a lagged but more reliable indicator of its diffusion than the number of confirmed cases. The cumulative tallies of confirmed CoViD-19 deaths up to a given time depends on the determination of the cause of death, delays in reporting deaths to central reporting agencies—different for deaths at home, in hospitals and other institutions—and delays in verification, consolidation and publication at reporting agencies. These tallies might underestimate the actual number of CoViD-deaths even in countries with comparatively reliable reporting systems. In the US, for instance, the grim milestone of 100,000 cumulative CoViD-19 deaths was reached at the end of May, when data from the Center for Disease Control and Prevention (CDC) suggested that the number of deaths in the country exceeded expectations based on past trends by close to 130,000.[2]

While some "excess" mortality might be due to increase in other causes-of-death mortality, a proportion of CoViD-19 deaths is likely unreported, but that proportion is both easier to estimate and an order-of-magnitude smaller than the proportion of unreported cases. CDC data from large-scale seroprevalence surveys suggest than the number of actual cases might be as much as 10 times the number of confirmed cases[3]—a situation in no way unique to the US.[4] CoViD-19 mortality indicators are also more pertinent for assessing public-health measures that were intended less to reduce the eventual number of cases than to "flatten the curve" and

eventually limit the number of CoViD-19 deaths by keeping the need for emergency hospitalizations below local hospital capacity.

Cumulative number of deaths and their ratio to population size have been used, in particular, to assess how different countries have fared in their response to the pandemic. Population size is only one of several demographic characteristics that may affect the number of deaths in a particular country, however, and the deaths per capita ratio represents the first rather than the only adjustment that can be taken towards more meaningful CoViD-19 mortality comparisons. Following well-established practices in demography,[5] this article presents alternative indicators that can be derived with additional demographic data. These period measures are discussed using results for the 186 UN countries and territories with at least one death by June 1st. To illustrate issues of scale, the measures are also calculated at the subnational level (e.g., states or provinces) in selected countries, which were the largest countries in the successive "epicenters" of the pandemic over time: first China, then Italy and Spain, followed by the US, and now Brazil, Mexico and Peru. At this writing, at least one of the measures presented here is estimated for a total of 386 national and subnational populations.

Methods and Data

We first calculate an occurrence/exposure *rate*, the period Crude CoViD-19 Death Rate (*CCDR*):

$$CCDR[t_1, t] = \frac{D^C[t_1, t]}{N(t_m).(t - t_1)}$$

where t_1 is an initial time, $D^C[t_1,t]$ a cumulative CoViD-19 death count between times t_1 and t, and $N(t_m)$ an estimate of the total population size at time t_m between time t_1 and time t. The difference between this period rate and the deaths per capita ratio is easy to miss when the death count in the numerator, identical for both, is an annual number of deaths. In that case, the number of person-years in the denominator of the occurrence/exposure rate can indeed be

approximated by the population size at some point during the year. However, the two are no longer directly comparable, and the metric of the ratio difficult to interpret, when the death counts correspond to periods of different durations. On the contrary, the *CCDR* is expressed in deaths per person-*year* and remains directly comparable to the annual Crude Death Rate (*CDR*) available for most populations. We first calculate the *CCDR* for the period starting on the day of the first CoViD-19 death in the population, which was obtained from World Health Organization (WHO) daily situation reports,[6] and a variable end date (September 4 at this writing). The estimated deaths count on that end date was obtained from Johns Hopkins University's Center for Systems Science and Engineering (CSSE)[7] and total population size was obtained from the UN.[8] Additional sources used for sub-national units are referenced in the Technical Appendix (see online supplementary information). Using projections from the University of Washington's Institute for Health Metrics and Evaluation (IHME),[9] we also calculate a *CCDR* for the period ending on December 31, 2020.

As age and sex variations in CoViD-19 mortality have been clearly established,[10] the period rates should be adjusted to take into account differences in age and sex distributions. Direct age-and-sex standardization requires data on CoViD-19 deaths by age and sex, which are unavailable or unreliable for a majority of UN countries and territories and most sub-national populations. An alternative approach, known as indirect standardization, borrows an age-and-sex pattern of mortality from a well-documented population so that only the age-and-sex distribution of the populations of interest are required. Based on this approach, we calculate the Comparative CoViD-19 Mortality Ratio (*CCMR*):

$$CCMR[t_1, t] = \frac{D^C[t_1, t]}{\sum_j \sum_i {}^{US}M_{ij}^C . N_{ij}(t_m)}$$

where ${}^{US}M_{ij}^{C}$ is the CoViD-19 death rate specific to age group i and sex j in the US and $N_{ij}(t_m)$ is the size of the age group i for sex j in the population of interest. The reference age-and-sex death rates were obtained from the Centers for Disease Control and Prevention (CDC) weekly-updated distribution of CoViD-19 deaths by age and sex in the US, selected because this is to date the largest number of CoViD-19 deaths distributed by age and sex.[11] Unavailable only for the 13 countries/territories whose population size is less than 90,000, population age-and-sex distributions for other countries/territories were taken from the UN data and, for subnational populations, from national statistical office data (see Technical Appendix, online supplementary information).

Multiplying a population's *CCMR* by the US *CCDR* yields an Indirectly age-and-sex Standardized CoViD-19 Death Rate (*ISCDR*) for that population, with the US age-and-sex population distribution as the standard:

$$ISCDR[t_1, t] = \sum_j \sum_i \left({}^{US}M_{ij}^{C}.\, CCMR[t_1, t]\right).\, {}^{US}C_{ij}(t_m)$$

CCMR and *ISCDR* are again calculated for two periods, both starting with the first CoViD-19 death in the population and ending either on September 4, using CSSE estimates, or on December 31, using IHME projections.

Last, life expectancy at birth provides a summary indicator of mortality in a population in a more intuitive metric (years) than these rates. A standard demographic technique allows to estimate the impact that *eliminating* a cause of death would have on life expectancy at birth. When a prior period life table (i.e., not factoring CoViD-19 mortality) is available, applying this technique backward allows to translate a cumulative number of CoViD-19-deaths for a given period into a CoViD-19-induced *reduction* in male and female life expectancies at birth for that period. While this estimation could be performed for reference periods of any duration, the

shorter the reference period the more problematic the intuitive interpretation of life expectancies becomes in the context of a fast-evolving pandemic.[12] We used the IHME projections of the cumulative number of CoViD-19 deaths by December 31 to derive new male and female 2020 life expectancies at birth in 277 populations with extant life tables (155 countries, plus Italian regions, Spanish autonomous communities, Mexican and US states). This required previously projected male and female year-2020 life tables in these populations. For countries, these were again derived from UN data, by interpolation between the 2015-20 estimates and 2020-25 projections. For sub-national populations, life tables available from national statistical institutes were extrapolated to 2020. Additional details on these calculations are described in the Technical Appendix (see online supplementary information).

Results

To illustrate the properties of these indicators, we briefly describe results from the September-4 updates of the CCSE, IHME and CDC data. (Full results for that week, ranked on *CCMR* values, are also provided in the online supplementary information). Across countries (with at least 100,000 inhabitants), Peru has now surpassed Belgium as the country with the highest current-period *CCDR* value (1.91 deaths per thousand person-years). At the subnational level, however, the corresponding values for the states of New Jersey (3.75) and New York (3.49) still exceed the highest value for a Peruvian *departamento* (Ica, 3.31).

The main motivation for the *CCDR* is not to compare CoViD-19 mortality across populations, but rather to compare CoViD-19 and overall mortality. Of these 386 populations, New York has reached the highest *CCDR* value to date (9.44 per thousand on 4/25), above the state's most recent annual *CDR* (7.83 per thousand in 2017).[13] The period *CCDR* remained above the 2017 *CDR* until May 20 (Figure 1). Ignoring seasonality and period trends in overall

mortality, this indicates roughly equivalent mortality from CoViD-19 and from all other causes combined between March 14 (first CoViD-19 death) and May 20.

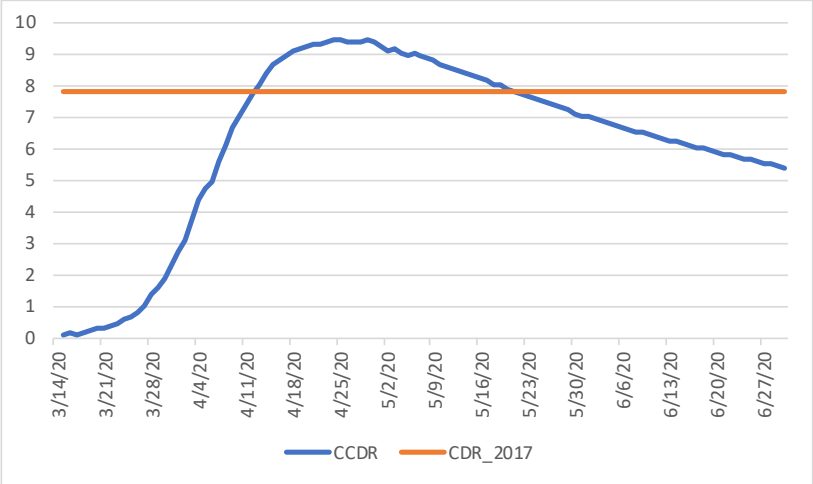

Figure 1: Estimated value of the period *CCDR*, New York (in deaths per 1,000 person-years, period starting on March 14 and ending on day shown on the horizontal axis)

Sources: CDC (*CDR*) and authors' calculations (*CCDR*, see technical appendix)

The effects of indirect age-standardization are illustrated in Figure 2, comparing current-period *CCDR* and *ISCDR* values for national and subnational populations of 10 million or more and the highest *ISCDR* values. By construction, the *CCMR* equals 1 and the *CCDR* and *ISCDR* are the same in the US, but the standardized *ISCDR* is lower than the unstandardized *CCDR* in Europe, whereas the standardized rate can be two to three times the unstandardized rate in other populations. With age standardization, several additional countries besides Peru now surpass Belgium: Mexico, Bolivia, Brazil, Chile, South Africa, Ecuador, Iraq and Columbia. At the subnational levels, Lima, Mexico City and Rio de Janeiro also have higher standardized rate than New York. Not appearing in Figure 2, 10 smaller subnational populations in Mexico and Peru have even higher ISCDR values than Lima, with the highest currently being in Quintana Roo (7.33 deaths per thousand person-years, Supplementary Materials).

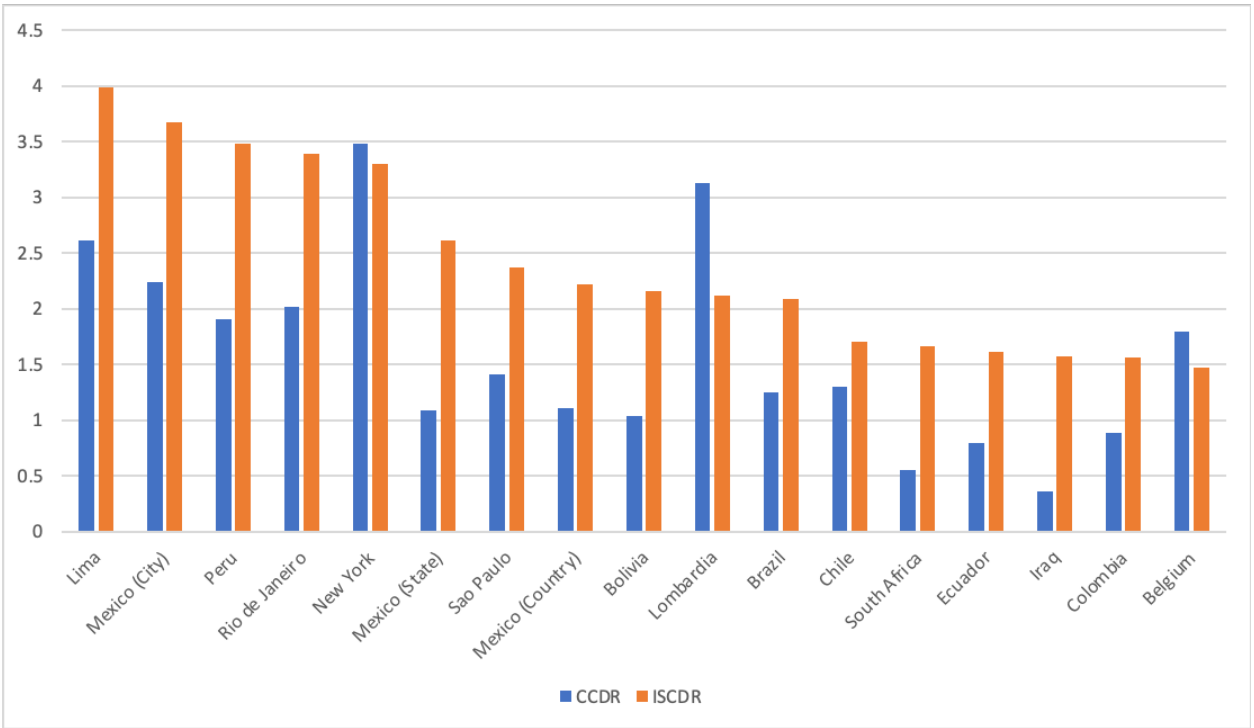

Figure 2: Estimated value of the *CCDR* and *ISCDR* (in deaths per 1,000 person-years), by country and subnational unit (countries and subnational units with the largest *ISCDR* values and a population size over 10 million)

Sources: Authors' calculations (see technical appendix)

Based on IHME projections, reductions in 2020 life expectancies at birth are currently estimated to exceed two and a half years in two countries, Ecuador (2.58 years) and Peru (2.51 years), and several subnational populations: Baja California (3.66 years), Madrid (3.26 years), Mexico City (2.57 years), and New York (2.54 years, Figure 3), as well as smaller ones not shown in Figure 3 (the US Virgin Islands; La Rioja in Spain; Quintana Roo, South Baja California, Tabasco, Sonora and Sinaloa in Mexico). Reductions are consistently larger for males than for females (e.g., 2.79 years for males and 2.36 years for females in Ecuador).

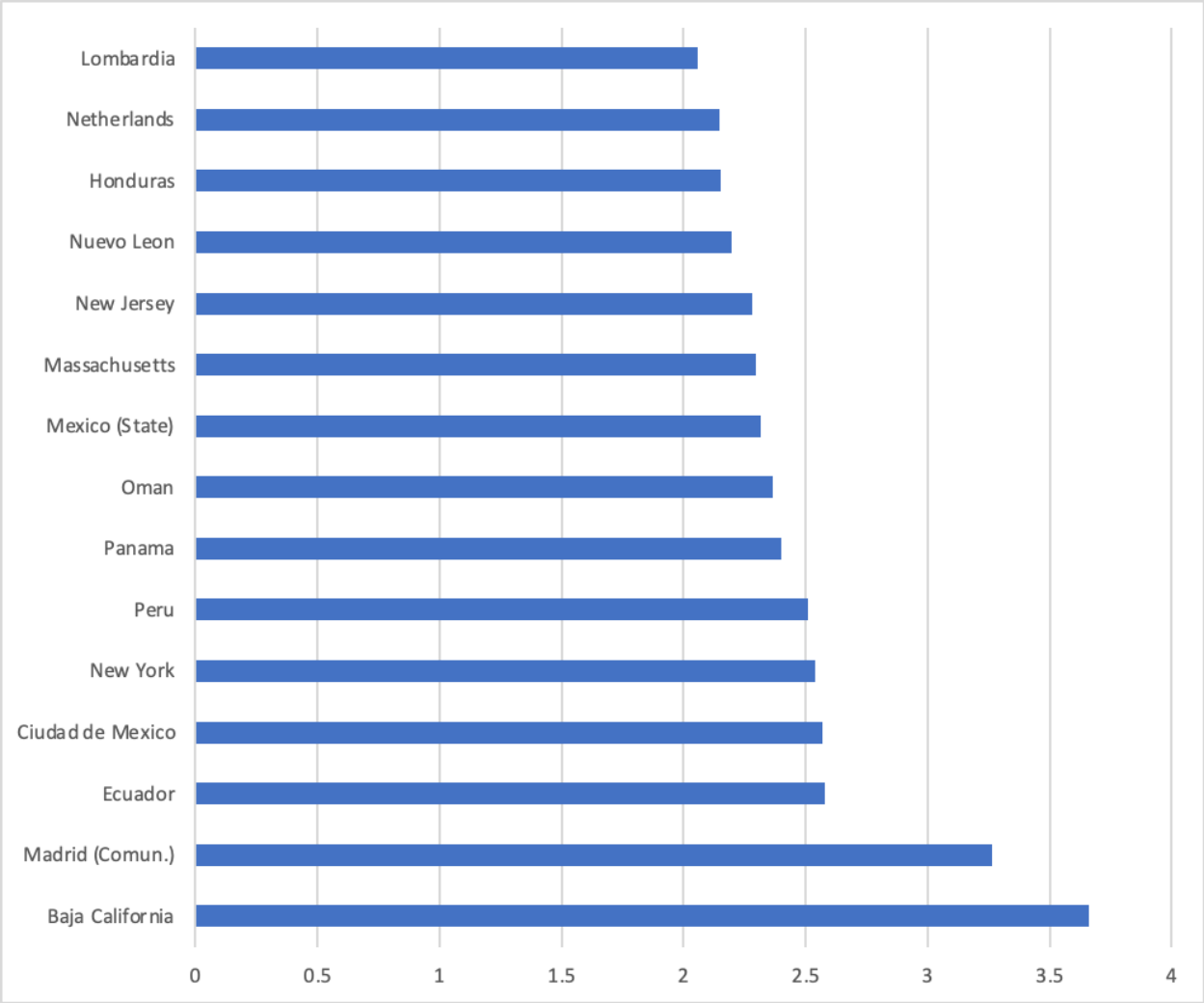

Figure 3: Estimated reduction in life expectancy at birth for year 2020, both sexes (in years), by country and subnational unit (countries and subnational units with the largest reductions and a population size over 3.5 million)

Sources: Authors' calculations (see technical appendix)

Projecting reductions in 2020 life expectancy at birth allows for comparison between the mortality impact of CoViD-19 and prior public health crises that might have interrupted the secular increase in life expectancies in most countries. While the 1.47-year reduction currently projected for the US is not among the largest, for instance, it would still be unprecedented since World War II, far exceeding the reversals induced by the HIV epidemic (from 75.8 years in 1992

to 75.5 years in 1993) and the opioid-overdose crisis (from 78.9 years in 2014 to 78.6 years in 2017).[14] Based on these projections, US life expectancy at birth would reach in 2020 its lowest level since 2003 (Figure 4).

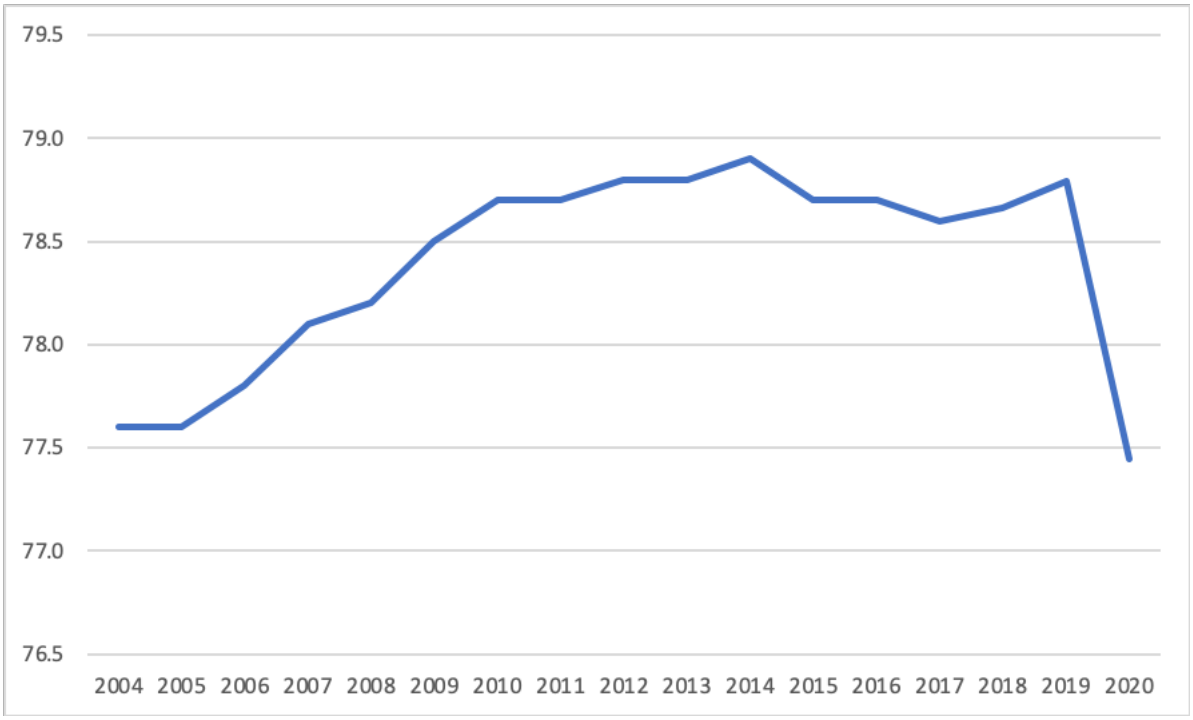

Figure 4: Estimated life expectancy at birth, U.S. population, both sexes, by year

Sources: CDC (2009-2017), UN and authors' calculations (2017-2020, see technical appendix)

Discussion

The results above illustrate the properties of different comparative indicators of CoViD-19 mortality. For comparisons across populations, the *ISCDR*, and *CCMR* on which it builds, control for three important factors that contribute to the cumulative count of CoViD-19 deaths in a population: the length of the period over which these deaths are cumulated, the size of the population, and its age-and-sex composition.

With respect to the first of these three factors, both the unstandardized and standardized rates are period indicators that increase and decrease as waves of the pandemic develop.

Contrary to the death per capita ratio, which can only increase over time, the period rates begin to decline when the daily number of additional deaths drops below its average for the period. This property of the period rates accurately reflects for CoViD-19 mortality a temporal dimension that can often be neglected for overall mortality. This also implies, however, that comparing *ISCDR* values across populations at too different durations of exposure to CoViD-19 would not be meaningful. As shown in Figure 1, this is more problematic earlier than later in the diffusion of the epidemic.

With respect to the second factor, comparing *ISCDR* values at the national or sub-national levels illustrate that dividing by population size does not make small and large populations fully comparable. National rates are but population-weighted averages of subnational rates. With person-to-person transmission and uneven population density, these rates can be expected to vary substantially across the territory of the largest countries, making it less likely that the national average will stand out in cross-national comparisons. While Brazil, Italy, Mexico, Spain and the US all have lower *CCDR* than Belgium, they each have a subnational entity of roughly similar population size with a higher *CCDR* than Belgium.

Disaggregation to smaller administrative entities with more comparable population sizes may provide better insights on the pandemic, but might be impeded by data availability. In this respect, indirect standardization has the advantage of not requiring data on CoViD-19 deaths by age and sex that may not be available or reliable for smaller areas. As a breakdown of CoViD-19 deaths *is* available from a number of European countries[15, 16] and US states, the *ISCDR* can actually be compared to a Directly age-and-sex Standardized CoViD-19 Death Rate (*DSCDR*) with the US age-and-sex population distribution as the standard. Doing so for 7 European countries and 6 states, Figure 5 shows that the indirectly standardized rate is much closer to the

directly standardized than to the unstandardized rate and typically very close to the former. Indirect standardization thus appears to be a valid alternative when direct standardization is not possible due to data limitations.

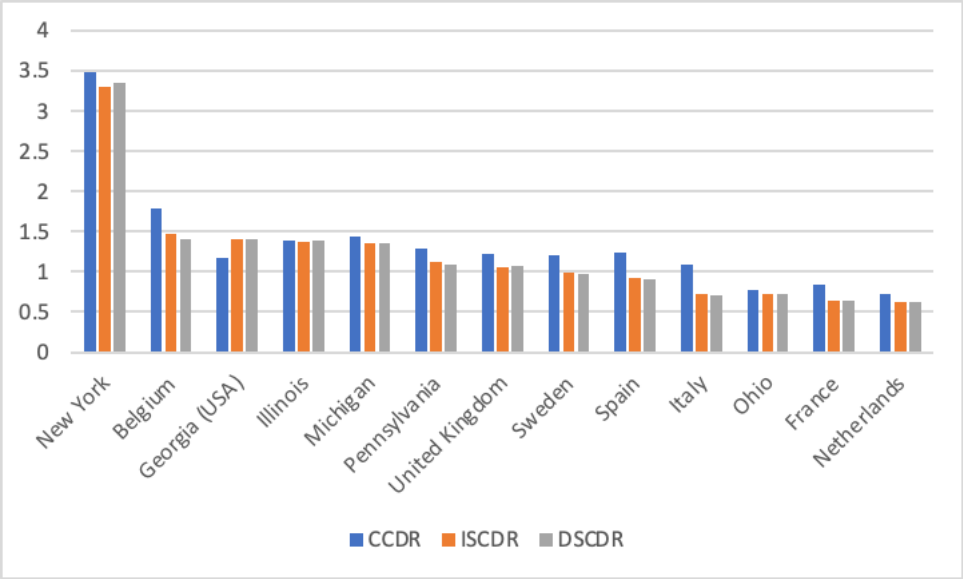

Figure 5: Estimated value of the *CCDR*, *ISCDR* and *DSCDR* (in deaths per 1,000 person-years), by selected countries and US states

Sources: CDC, Ined and authors' calculations (see technical appendix)

None of the illustrative results presented here make any adjustment for potential biases in the number of confirmed CoViD-19 deaths. Estimates of life expectancies based on these also assume no "indirect" effect of the pandemic on other-cause mortality. In populations with complete and timely registration of deaths, the reporting biases and indirect effects can be jointly assessed from the increase in overall mortality over past "benchmark" mortality levels. In this respect, CDC data continues to suggest that the cumulative number of CoViD-19 deaths to date does not fully account for the overall increase in US mortality.[17] However, the estimation of "excess" deaths directly or indirectly attributable to CoViD-19 can be quite sensitive to the choice of a benchmark period to represent past mortality conditions. In France, for instance, the

number of excess deaths for the first half of 2020 can be estimated to reach approximately 40,000 when years 2000 through 2019 are used to estimate benchmark mortality,[18] but using only the most recent year, 2019, for that mortality benchmar, the estimate drops below 30,000—the number of confirmed CoViD-19 deaths over the same period.[19]

Obviously estimates of year-2020 life expectancies also depend on the accuracy of current projections of the increases in the cumulative number of CoViD-19 deaths through the end of the year. At this writing, the CDC tracks no less than 39 forecasting models.[20] Our choice of the IHME projections among those to illustrate the properties of the different indicators was not based on a quality assessment, which would be beyond our expertise. The IHME projections global scope and end-of-the-year horizon simply makes them more suitable to illustrate the various indicators than other models.

The values of the indicators discussed above are the most recent of weekly updates of the CCSE, IHME and CDC data that we have calculated and shared on a Github repository since the end of April.[21] The rapidly evolving data and understanding of CoViD-19 mortality will likely continue to require frequent updates and flexibility. Calculations of the indicators presented here can easily be customized for different periods, different geographical scales, or to accommodate uncertainty across different sources of estimates and forecasts.

References

[1] Brookmeyer R & Gail MH. AIDS epidemiology: a quantitative approach. Oxford, England: Oxford University Press, 1994.

[2] Weinberger DW, Chen J, Cohen, T et al. 2020. Estimation of Excess Deaths Associated With the COVID-19 Pandemic in the United States, March to May 2020. *JAMA Intern Med*. Published online July 1, 2020. doi:10.1001/jamainternmed.2020.3391

[3] Centers for Disease Control and Prevention. Commercial Laboratory Seroprevalence Survey Data. https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/commercial-lab-surveys.html. Last used June 9, 2020.

[4] Modi C, Boehm V, Ferraro S, et al. How deadly is COVID-19? A rigorous analysis of excess mortality and age-dependent fatality rates in Italy. https://www.medrxiv.org/content/10.1101/2020.04.15.20067074v3. May 14, 2020.

[5] Preston SH, Heuveline P & Guillot M. Demography: measuring and modeling population processes. Blackwell: Malden, MA, Oxford, England & Carlton, Australia, 2001.

[6] World Health Organization. Daily situation reports. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/. Last used June 2, 2020.

[7] Dong E, Du H & Gardne L. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Inf Dis* (forthcoming). https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(20)30120-1/fulltext. February 19, 2000.

[8] United Nations, Department of Economic and Social Affairs, Population Division. World Population Prospects 2019, Online Edition. Rev. https://population.un.org/wpp/Download/Standard/CSV/. Last used June 2, 2020.

[9] University of Washington, Institute for Health Metrics and Evaluation. COVID-19 projections. https://covid19.healthdata.org/. Last used June 12,2020.

[10] Dowd JB, Andriano L, Brazel DM, et al. Demographic science aids in understanding the spread and fatality rates of COVID-19. PNAS 2020,117: 9696-8. https://doi.org/10.1073/pnas.2004911117

[11] Centers for Disease Control and Prevention. Provisional COVID-19 death counts by sex, age and state. https://data.cdc.gov/NCHS/Provisional-COVID-19-Death-Counts-by-Sex-Age-and-S/9bhg-hcku. Last used June 12, 2020.

[12] Heuveline P. 2020. Estimating the Impact of COVID-19 on the Individual Lifespan: A Conceptual Detour and an Empirical Shortcut. *MedRxiv* https://www.medrxiv.org/content/10.1101/2020.08.09.20171264v2

[13] Kochanek KD, Murphy SL, Xu JQ & Arias E. 2019. Deaths: final data for 2017. NVSR 2019,68(9). Hyattsville, MD: National Center for Health Statistics.

[14] Arias E, & Xu JQ. United States life tables, 2017. NVSR 2019,68(7). Hyattsville, MD: National Center for Vital Statistics.

[15] National Institute for Demographic Studies (INED). The demographics of COVID-19 deaths. https://dc-covid.site.ined.fr/en/. Last used June 9, 2020.

[16] 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

[17] Centers for Disease Control and Prevention. Excess deaths associated with COVID-19. https://www.cdc.gov/nchs/nvss/vsrr/covid19/excess_deaths.htm. Last used July 2, 2020.

[18] Human Mortality Database. 2020. *Short-term Mortality Fluctuations*. https://mpidr.shinyapps.io/stmortality/

[19] Guillot M & Khlat M. 2020. Epidémie de Covid-19: quel impact sur l'espérance de vie en France? *The Conversation.* https://theconversation.com/epidemie-de-covid-19-quel-impact-sur-lesperance-de-vie-en-france-141484. Published online June 28, 2020.

[20] Centers for Disease Control and Prevention. COVID-19 forecasts: Cumulative deaths. https://www.cdc.gov/coronavirus/2019-ncov/covid-data/forecasting-us.html. Last used September 11, 2020.

[21] https://github.com/statsccpr/ind-cov-mort