

Estimation of excess mortality due to COVID-19

Published online May 13, 2021

In the IHME estimation of COVID-19 infections, hospitalizations, and deaths to date, we have used officially reported COVID-19 deaths for nearly all locations. Since the release on May 6, 2021, we have incorporated an approach that relies on the estimation of excess mortality due to COVID-19. There are several reasons that have led us to include this approach. These reasons include the fact that testing capacity varies markedly across countries and within countries over time, which means that the reported COVID-19 deaths as a proportion of all deaths due to COVID-19 also vary markedly across countries and within countries over time. In addition, in many high-income countries, deaths from COVID-19 in older individuals, especially in long-term care facilities, went unrecorded in the first few months of the pandemic. In other countries, such as Ecuador, Peru, and the Russian Federation, the discrepancy between reported deaths and analyses of death rates compared to expected death rates, sometimes referred to as “excess mortality,” suggests that the total COVID-19 death rate is many multiples larger than official reports. Estimating the excess COVID-19 death rate is important both for modeling the transmission dynamics of the disease to make better forecasts, and also for understanding the drivers of larger and smaller epidemics across different countries.

Our approach to estimating the excess COVID-19 death rate is based on measurement of the excess death rate during the pandemic week by week compared to what would have been expected based on past trends and seasonality. However, the total excess death rate does not equal the excess COVID-19 death rate. Excess mortality is influenced by six drivers of all-cause mortality that relate to the pandemic and the social distancing mandates that came with the pandemic. These six drivers are: a) the excess COVID-19 death rate, that is, all deaths directly related to COVID-19 infection; b) the increase in mortality due to needed health care being delayed or deferred during the pandemic; c) the increase in mortality due to increases in mental health disorders including depression, increased alcohol use, and increased opioid use; d) the reduction in mortality due to decreases in injuries because of general reductions in mobility associated with social distancing mandates; e) the reductions in mortality due to reduced transmission of other viruses, most notably influenza, respiratory syncytial virus, and measles; and f) the reductions in mortality due to some chronic conditions, such as cardiovascular disease and chronic respiratory disease, that occur when frail individuals who would have died from these conditions died earlier from COVID-19 instead. To correctly estimate the excess COVID-19 mortality, we need to take into account all six of these drivers of change in mortality that have happened since the onset of the pandemic.

Our analysis follows four key steps. First, for all locations where weekly or monthly all-cause mortality has been reported since the start of the pandemic, we estimate how much mortality increased compared to the expected death rate. In other words, we estimate excess mortality in all locations with sufficient data. Second, based on a range of studies and consideration of other evidence, we estimate the fraction of excess mortality that is from excess COVID-19 deaths as opposed to the five other drivers that influence excess mortality. Third, we build a statistical model that predicts the ratio of excess COVID-19 deaths to reported COVID-19 deaths based on covariates and spatial effects. Fourth, we use this statistical relationship to predict the ratio of excess to reported COVID-19 deaths in places without data on excess COVID-19 deaths and then multiply the reported COVID-19 deaths by this ratio to generate estimates of excess COVID-19 deaths for all locations. More details on each of these analytical steps are presented below.

1. Estimating excess mortality compared to expected mortality for locations where all-cause mortality data have been reported during the pandemic

56 countries and 198 subnational units have reported either weekly or monthly deaths from all causes for parts of 2020 and for prior years. Our analysis of excess mortality follows three steps. First, we estimate expected mortality in the absence of COVID-19 based on the patterns of all-cause mortality reported in prior years; second, we subtract observed all-cause mortality from March 2020 onward from expected mortality; and third, we remove from the analysis known periods of excess mortality due to causes other than COVID-19, such as the August 2020 heat wave in many European countries. For locations where vital registration systems are not complete, we apply the adjustment to the reported death counts based on our estimated completeness from the Global Burden of Disease study.¹

To estimate expected mortality, we need to account for both seasonality and the secular trend in all-cause mortality. We developed a novel method to capture the expected mortality level and trend based on past data on all-cause mortality. In this method, we estimate the typical seasonal pattern of mortality and then estimate the secular trend of all-cause mortality after correcting for the seasonal pattern. Figure 1 below shows the application of this model to all-cause mortality data by week from Denmark. By grouping data by weeks, we are able to estimate an overall seasonality pattern by week (as shown in Panel A of Figure 1). Residuals between weekly observed data and the fitted seasonality pattern are shown in Panel B and represent the time trend. We fit a spline to the residuals to estimate the time trend and then use it to extend the time trend into 2020 and to the present day. By combining the seasonal trend and the secular trend in the residuals, for each location we generate a prediction of the expected level of mortality in 2020 and 2021, as shown in the red box in Panel C of Figure 1.

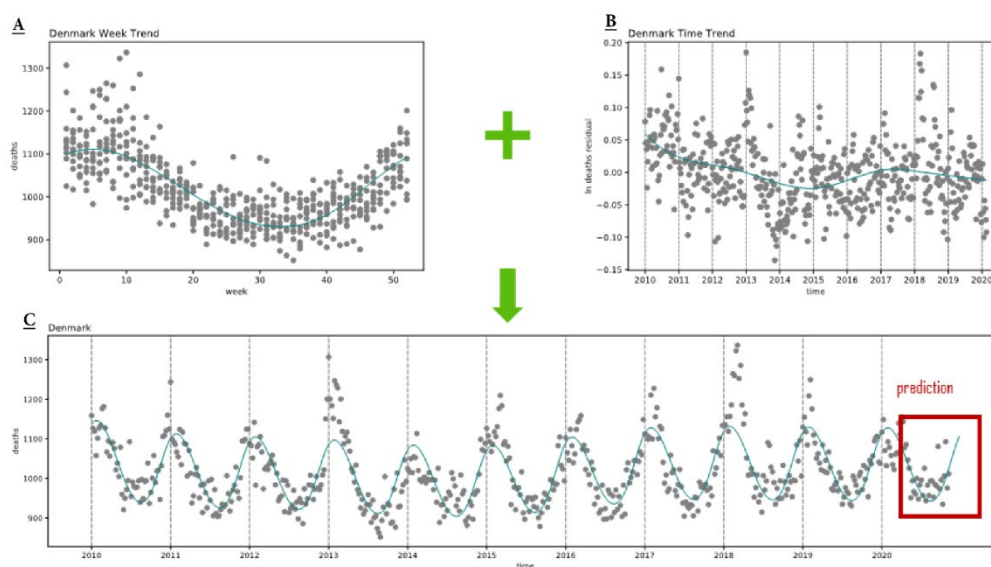


Figure 1. New model for the estimation of expected mortality

As expected, the model specification of the spline can have a sizeable impact on the estimated expected mortality. To make our results more robust to model specification, we create an ensemble of four different model specifications for the spline. In addition, we also include in the ensemble a Poisson model with fixed effects on week and year, as well as a model that assumes that the expected mortality rate for 2020 and 2021 is the same as the weekly mortality rate observed in 2019. To derive weights for the different models in the ensemble, we examined how each model performed out of sample. We fit the model to all data prior to 2019 and then evaluate how each model performed in predicting mortality levels in 2019 compared to observed mortality in 2019. We then use the root mean squared error (RMSE) of the predictions for 2019 to derive weights for each of the component

models in the ensemble. Figure 2A shows the distribution of RMSE by location for each of six models included in the model ensemble. Figure 2B shows the estimated excess mortality, which is the difference between the observed and estimated expected mortality, for each component model and for the ensemble for Spain.

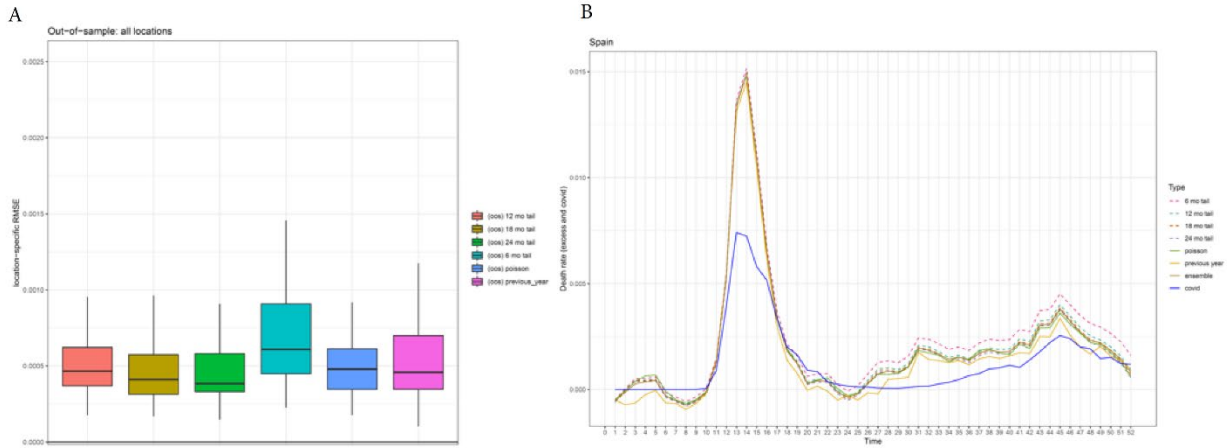


Figure 2. Estimation of excess mortality, an ensemble approach

For each location, we then compared the estimate of excess mortality by week (or month, depending on the data) with reported COVID-19 deaths. This revealed two types of data issues. First, in many European countries there was a spike of excess deaths in weeks 31–35 during a period when COVID-19 reported deaths were extremely low. This period coincided with a heat wave and national reports of deaths due to the heat wave. We excluded these weeks of data from subsequent analyses. Another type of data anomaly was related to lags in the reported all-cause mortality. As an example, Figure 3 shows the lags in the reported all-cause mortality from the vital registration system in Brazil. There is clear and significant late registration of deaths since June 2020. In this case, we have marked the 2020 vital registration data from Brazil as outliers and opted to use data from the civil registration system ([source link](#)). We systematically reviewed the input vital registration data and trimmed time periods that are likely to be subjected to late registration for all locations in the analysis.

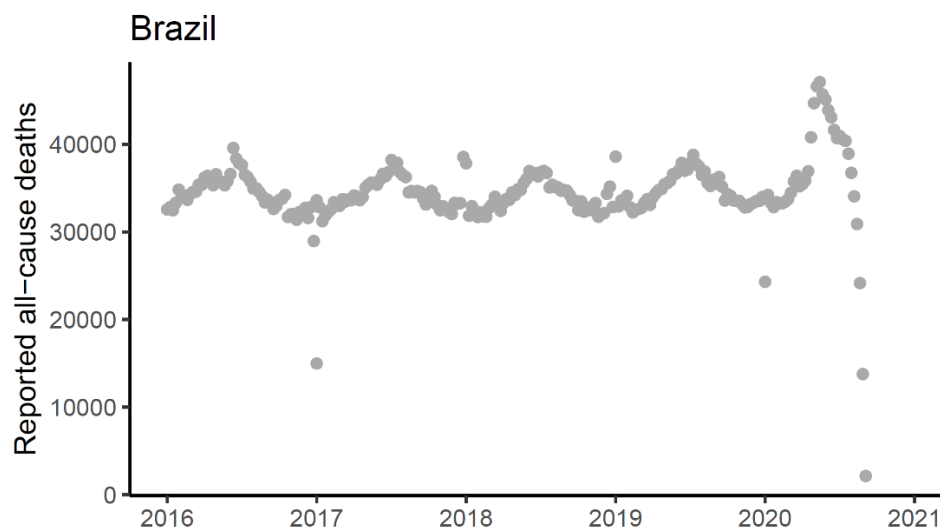


Figure 3. Reported all-cause deaths for Brazil

2. Estimating the fraction of excess mortality that is direct COVID-19 deaths

As noted above, excess mortality is a function of six potential drivers, the most important of which is the excess COVID-19 death rate. Deaths that are directly due to COVID-19 are likely underreported in many locations, particularly in settings where COVID-19 testing is in short supply. Most excess mortality is likely misclassified COVID-19 deaths. An analysis by the Netherlands statistical agency suggested that all excess deaths in the Netherlands were directly due to COVID-19.² In fact, their analysis actually suggested that direct COVID-19 deaths may be higher than estimated excess deaths because deaths due to some other causes have declined during the pandemic.

The second driver of excess mortality is reduced health care utilization for many causes;³ however, the impact of reduced health care use on health outcomes is harder to prove. Many mechanisms have been proposed, including reduced vaccination rates and reduced births in hospital.⁴ Demonstrated increases in cause-specific mortality related to these causes, however, have not yet been verified. The impact of changes in health care utilization on excess mortality may be observed in later years, rather than in 2020 or the first quarter of 2021.

Third, convincing evidence has been found that rates of anxiety and depression have increased, which might in turn lead to increases in deaths from suicide.⁵ To date, the evidence on increased suicide is very limited.⁶ Opioid deaths, on the other hand, have clearly increased⁷ in the United States. Compared to past trends, opioid deaths increased by around 15,000 since March 2020. Evidence on whether this has also occurred in other countries awaits further study.

Fourth, we reviewed the evidence on decreases in injuries as a result of reductions in mobility. We analyzed data from 12 countries that provide cause of death data by week or month, which allows us to test whether some causes decreased significantly during 2020 and whether that decrease was related to the decreases in mobility that have been reported. This analysis suggests that globally, injury mortality decreased by 5% in 2020 due to reductions in mobility. At the global level, this translates into a reduction of approximately 215,000 deaths.

Fifth, some infectious causes of death may have declined during the pandemic due to the behavioral changes associated with control of the pandemic, including mask use and reduced contact with others. Causes that have clearly declined are influenza,^{8,9} respiratory syncytial virus,¹⁰ measles,¹¹ and possibly other respiratory viruses and viral diarrheas. For example, influenza cases in the United States declined 99.3% from the winter season of 2019-2020 to the winter season of 2020-2021. Combining the reductions reported in different countries in influenza, respiratory syncytial virus, and measles, the global reduction in mortality from these causes may be larger than 400,000 deaths.

Sixth, deaths from some chronic conditions such as ischemic heart disease or chronic respiratory disease declined in some months of 2020, most notably in May and June in Europe. These declines were most likely due to the fact that frail individuals who died from COVID-19 earlier in the year would otherwise likely have died from these chronic conditions. The strongest evidence for this effect is that excess mortality was negative in some countries in Europe in June when the reported COVID-19 death rate was very low. In aggregate, this effect likely reduced mortality by only 2% based on our analysis.

Overall, the evidence suggests reductions of 615,000 deaths, or potentially more, stemming from behavioral changes at the global level. The main potential increases in excess mortality due to deferred care and increases in drug overdose and depression are hard to quantify at this point or are of a much smaller magnitude. Given that there is insufficient evidence to estimate these contributions to excess mortality, for now we assume that excess COVID-19 deaths equal excess mortality. For the reasons presented in this section, we believe that this is likely an underestimate. As the evidence is strengthened in the coming months and years, it is likely that we will revise our estimates of the excess COVID-19 death rate upward in future iterations of this work, once we can properly take into account the drivers described in this section.

3. Estimating the ratio of excess mortality to reported COVID-19 deaths

Based on our analysis, we have generated a ratio of excess mortality to reported COVID-19 mortality for each location. These analyses, based on weekly or monthly mortality data, have been supplemented with published studies for 12 national and subnational locations where the detailed data have not been made publicly available for our analysis. Figure 4 shows the distribution of these ratios in the available data.

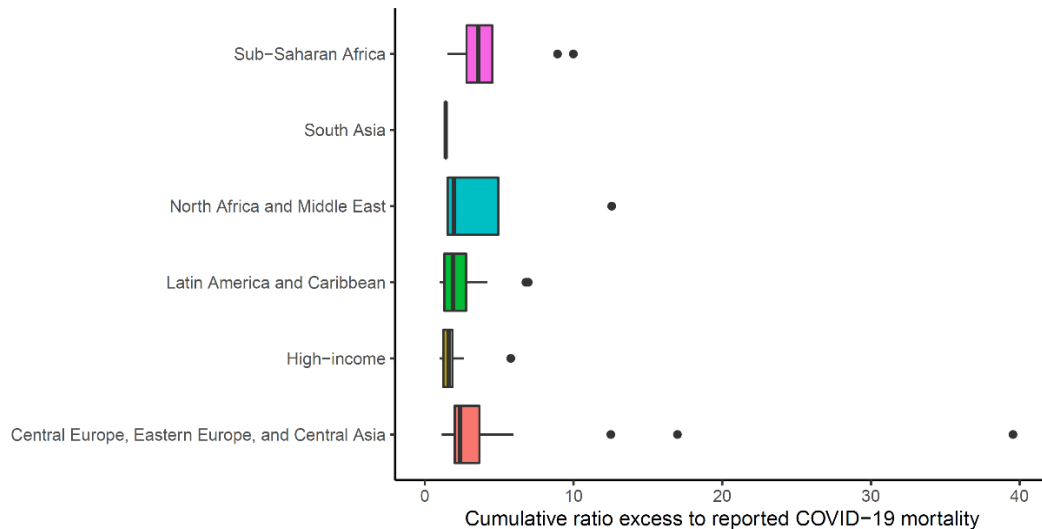


Figure 4. Distribution of weekly/monthly ratios of excess mortality to COVID-19 mortality by Global Burden of Disease super-region

After extensive testing of potential covariates and model specifications, we developed models that predict the ratio of excess COVID-19 mortality to reported COVID-19 mortality as a function of the infection-detection rate and location-specific fixed effects. We use a Bayesian cascade model to allow the relationships between the covariates and the predicted ratio to vary by region and country. More specifically, we use the bounded logit of the ratio as the dependent variable and infection-weighted average of infection-detection rate (IDR) as the covariate. First, a global spline on IDR (lagged by 19 days) is estimated. Then, the residual is fit with location-specific intercepts at four levels: subnational, national, GBD region, GBD super-region, and global.

4. Generating predictions of excess COVID-19 mortality for all locations

Using the same model described in the previous section that relates the ratio of excess mortality to reported COVID-19 mortality as a function of the IDR and location-specific intercept, we predict the ratio of excess COVID-19 mortality to reported COVID-19 mortality for all locations between March 2020 and May 2021. These predictions are a function of the cumulative IDR and location fixed effects and capture, through the Bayesian cascade, location-specific variation in the ratios.

Figure 5 shows a map of the predicted ratio of excess COVID-19 deaths to reported COVID-19 deaths for March 2020 to May 2021. Ratios range from very high levels in many Eastern European and Central Asian countries to ratios that are much closer to 1 in several high-income countries. Notably, for most countries in sub-Saharan Africa, which have to date reported relatively low numbers of COVID-19 deaths, we are estimating that the ratios range from about 1.6 to 4.1, suggesting that the total number of COVID-19 deaths in the region is several times higher than previously thought. Similarly, India, the country with the most recent severe wave of cases and deaths, is estimated to have an overall ratio of

2.97, which implies that the excess COVID-19 death toll to date is much higher than what has been reported.

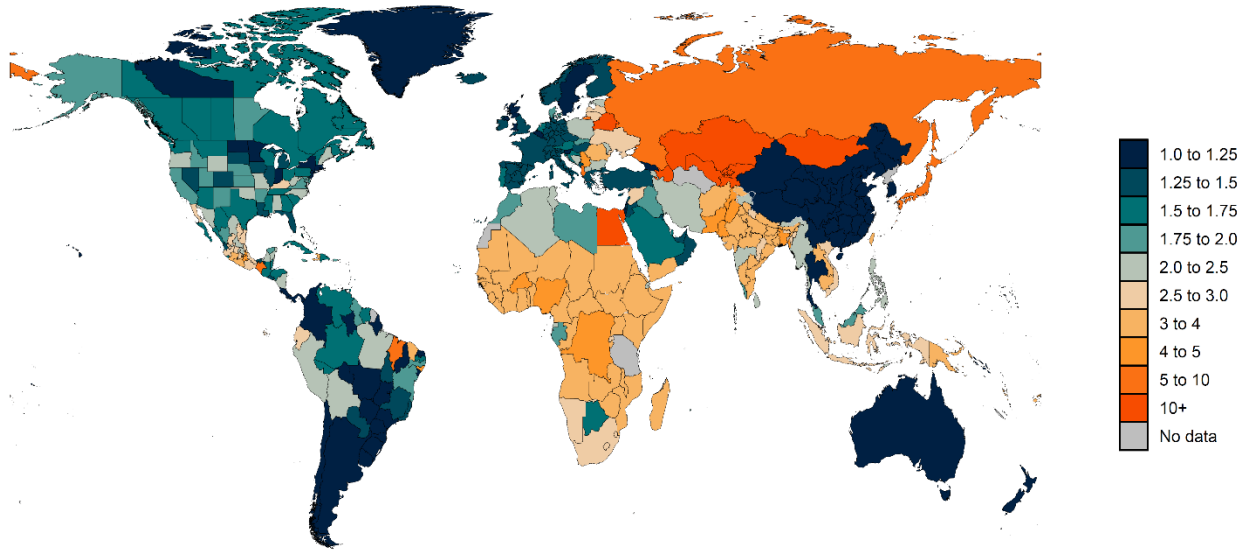


Figure 5. Ratios of excess COVID-19 deaths to reported COVID-19 deaths

Figure 6 shows the adjusted daily deaths for the world once reported deaths are multiplied by these estimated ratios for each location and week. Our analysis estimates that by May 13, 2021, the total number of COVID-19 deaths was 7.1 million, a figure that is more than two times higher than the reported number of deaths of 3.33 million.

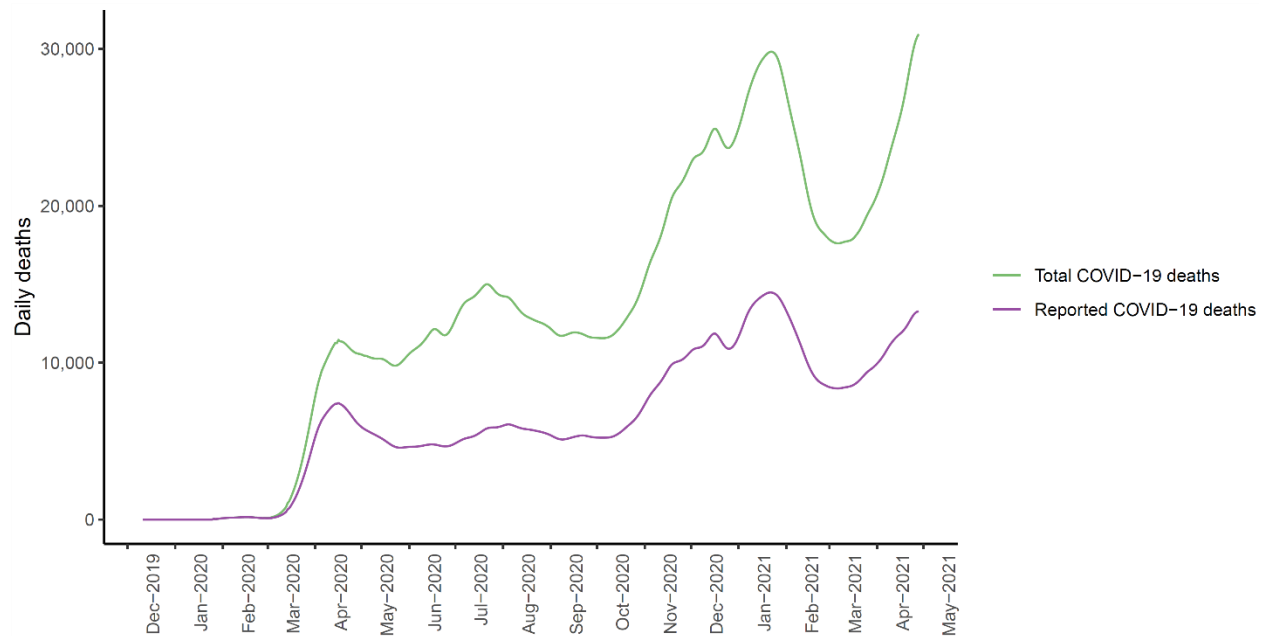


Figure 6. Global reported and excess COVID-19 deaths, 2020-2021

Table 1 below shows the 20 countries with the highest number of excess COVID-19 deaths by May 10, 2021.

Table 1. The 20 countries with the highest numbers of excess COVID-19 deaths, March 2020 to May 2021

Country	Excess COVID-19 deaths	Reported COVID-19 deaths
United States of America	912,345	578,555
India	736,811	248,016
Mexico	621,962	219,372
Brazil	616,914	423,307
Russian Federation	607,589	111,909
United Kingdom	210,076	150,815
Iran (Islamic Republic of)	180,487	75,547
Italy	178,144	122,851
Egypt	175,590	13,970
South Africa	161,504	54,746
Poland	153,626	69,954
Peru	151,939	64,511
Ukraine	143,415	48,393
France	134,400	106,874
Spain	124,449	85,822
Germany	122,977	84,807

Indonesia	118,796	47,150
Romania	89,619	29,020
Kazakhstan	84,453	5,810
Colombia	80,968	78,216

Table 2. The 20 countries with the highest death rates due to COVID-19 (deaths per 100,000), March 2020 to May 2021

Country	Excess COVID-19 death rate	Reported COVID-19 death rate
Azerbaijan	672.7	46.3
Bosnia and Herzegovina	601.1	268.3
Bulgaria	559.9	245.5
Albania	528.9	88.5
Mexico	497.8	175.6
North Macedonia	482.3	237.1
Belarus	472.2	27.8
Romania	465.9	150.9
Kazakhstan	459.2	31.6
Peru	446.9	189.8
Slovakia	436.6	221.2

Russian Federation	414.1	76.3
Lithuania	403.4	144.7
Poland	399.7	182.0
Hungary	397.4	296.1
Czechia	391.1	279.1
Republic of Moldova	384.9	161.5
Montenegro	346.1	247.7
Ukraine	325.6	109.9
Latvia	321.9	115.2

Globally, we observe a high level of heterogeneity in terms of cumulative excess COVID-19 death rates among countries and even subnational units within a country. Figure 7 below shows the cumulative COVID-19 death rate up to May 13, 2021. While at the global level, the cumulative excess COVID-19 death rate is 91.7 per 100,000, Vietnam has the lowest excess COVID-19 death rate at 0.1 per 100,000. Thirteen countries, on the other hand, have excess COVID-19 death rates higher than 400 per 100,000, as shown in Table 2 above.

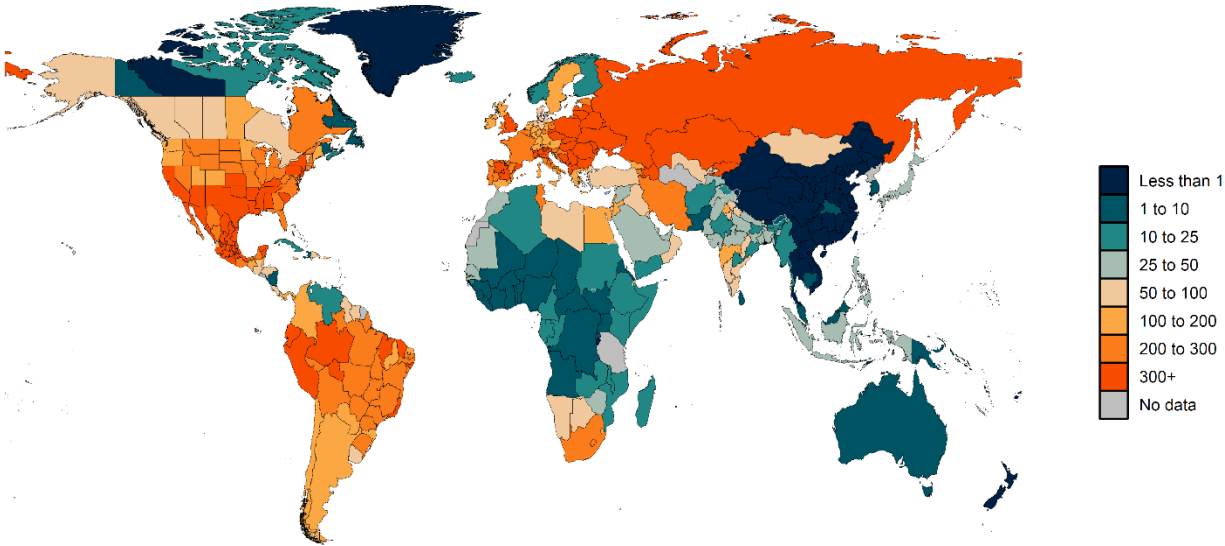


Figure 7. Cumulative excess COVID-19 death rate by May 13, 2021

References:

1. GBD 2019 Demographics Collaborators. Global age-sex-specific fertility, mortality, healthy life expectancy (HALE), and population estimates in 204 countries and territories, 1950-2019: a comprehensive demographic analysis for the Global Burden of Disease Study 2019. *Lancet*. 2020 Oct 17;396(10258):1160-1203. doi: 10.1016/S0140-6736(20)30977-6. PMID: 33069325; PMCID: PMC7566045.
2. Statistics Bureau of the Netherlands. 2021. <https://www.cbs.nl/nl-nl/nieuws/2021/14/3-9-duizend-mensen-overleden-aan-covid-19-in-december-2020>.
3. Dinmohamed AG, Visser O, Verhoeven RHA, Louwman MWJ, van Nederveen FH, Willems SM, Merx MAW, Lemmens VEPP, Nagtegaal ID, Siesling S. Fewer cancer diagnoses during the COVID-19 epidemic in the Netherlands. *Lancet Oncol*. 2020 Jun;21(6):750-751. doi: 10.1016/S1470-2045(20)30265-5. Epub 2020 Apr 30. Erratum in: *Lancet Oncol*. 2020 May 4. PMID: 32359403; PMCID: PMC7252180.
4. McDonald HI, Tessier E, White JM, Woodruff M, Knowles C, Bates C, Parry J, Walker JL, Scott JA, Smeeth L, Yarwood J, Ramsay M, Edelstein M. Early impact of the coronavirus disease (COVID-19) pandemic and physical distancing measures on routine childhood vaccinations in England, January to April 2020. *Euro Surveill*. 2020 May;25(19):2000848. doi: 10.2807/1560-7917.ES.2020.25.19.2000848. PMID: 32431288; PMCID: PMC7238742.
5. Gunnell D, Appleby L, Arensman E, Hawton K, John A, Kapur N, Khan M, O'Connor RC, Pirkis J; COVID-19 Suicide Prevention Research Collaboration. Suicide risk and prevention during the COVID-19 pandemic. *Lancet Psychiatry*. 2020 Jun;7(6):468-471. doi: 10.1016/S2215-0366(20)30171-1. Epub 2020 Apr 21. PMID: 32330430; PMCID: PMC7173821.
6. John A, Pirkis J, Gunnell D, Appleby L, Morrissey J. Trends in suicide during the covid-19 pandemic. *BMJ*. 2020 Nov 12;371:m4352. doi: 10.1136/bmj.m4352. PMID: 33184048.
7. Jesse C. Baumgartner and David C. Radley, "The Spike in Drug Overdose Deaths During the COVID-19 Pandemic and Policy Options to Move Forward," *To the Point* (blog), March 25, 2021. <https://doi.org/10.26099/gyf5-3z49>.
8. Chow, A., Hein, A. A., & Kyaw, W. M. Unintended consequence: influenza plunges with public health response to COVID-19 in Singapore. *The Journal of Infection*. 2020 Aug;81(2): e68-e69. doi: 10.1016/j.jinf.2020.04.035.
9. Sun, J., Shi, Z., & Xu, H. Non-pharmaceutical interventions used for COVID-19 had a major impact on reducing influenza in China in 2020. *Journal of Travel Medicine*. 2020 Dec 23;27(8):taaa064. doi: 10.1093/jtm/taaa064.
10. Haapanen M, Renko M, Artama M, Kuitunen I. The impact of the lockdown and the re-opening of schools and day cares on the epidemiology of SARS-CoV-2 and other respiratory infections in children - A nationwide register study in Finland. *EclinicalMedicine*. 2021 Apr; 34:100807. doi: 10.1016/j.eclinm.2021.100807. Epub 2021 Mar 29. PMID: 33817612; PMCID: PMC8007090.
11. Nicolay N, Mirinaviciute G, Mollet T, Celentano LP, Bacci S. Epidemiology of measles during the COVID-19 pandemic, a description of the surveillance data, 29 EU/EEA countries and the United Kingdom, January to May 2020. *Euro Surveill*. 2020 Aug;25(31):2001390. doi: 10.2807/1560-7917.ES.2020.25.31.2001390. PMID: 32762795; PMCID: PMC7459270.