

Appendix 1: Data Sources and Methods

Joshua R. Goldstein and Ronald d. Lee

April 5, 2020

1. Estimation of Covid-19 mortality age-shares for Figure 1(a)

We calculated the age-shares of Covid-19 mortality using both the age-distribution of Covid-19 attributed deaths and the age-structure of the population. Figure A1 shows the recorded deaths available to us in column 1. The scales differ both because of the scale and stage of the epidemic and the completeness of reporting (compare for example France and Italy and South Korea). The age-pattern of deaths also differs, notably in Hubei, where there are many fewer deaths at the oldest ages. In order to take account the age-structure of each population, we calculate rates using the deaths in column 1 and the population in column 2. These are shown in column 3, labeled unnormalized Covid-19 death rates. Note that magnitudes of these unnormalized rates are also quite different, but the general pattern is similar. For example, comparing Italy and Hubei, the distribution of deaths is more skewed to older ages in Italy than in Hubei, but once the population age-distribution is accounted for, the mortality rate schedule looks quite similar.

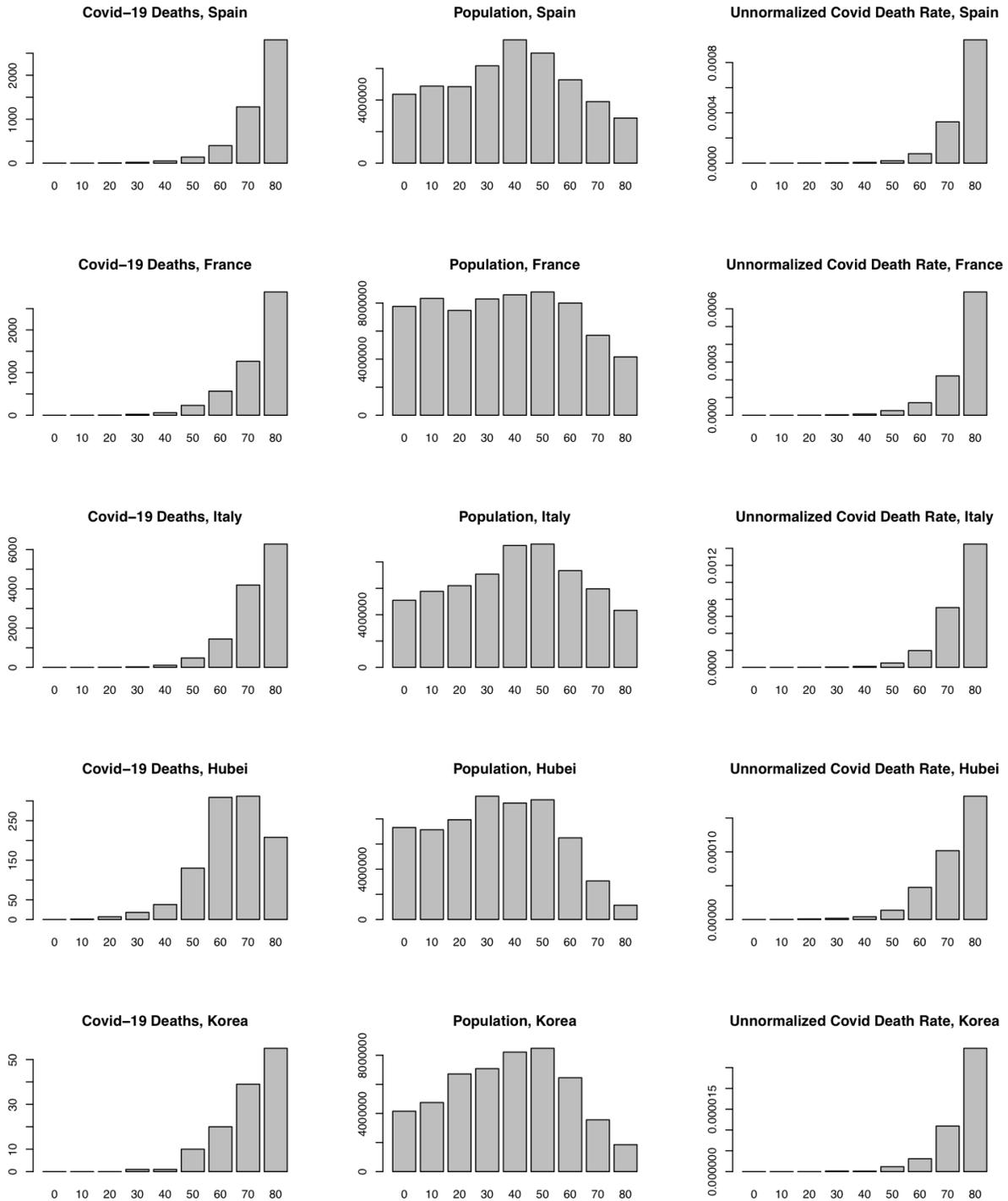


Figure A1 Death and population distributions and unnormalized Covid Death Rates for Hubei, France, Italy, South Korea, and Spain.

2. Calculation of normalized rates

The normalized rates we show in figure 1(b) are calculated for each country separately. We divide the unnormalized rates by the sum over all ages of the normalized rates. Each of the closed age groups (0-9, ..., 70-79) is weighted by the 10 year width of that age group. The open age-group is weighted by $10 / a(80+, i)$. Where, for each country i , we compute an estimate of the average age of death in the open interval 80+ using the UN population estimates for 5-year intervals (80-84, ..., 100-104), weighted by a standard mortality schedule from the United States both-sex period life table in 2017. This goal of this weighting is to account for variation among the countries in the age-structure for those over age 80.

The average normalized rate is calculated as the arithmetic mean across countries of the normalized age-specific mortality rates.

The exponential curve shown in figure 1(b) is a Gompertz hazard curve with exponential rate $b = 1/9$, with the level set so that it will intersect the average normalized rate at age group 70-79.

3. Data Sources for Figures 1 and 2

The following sources were used:

China

The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team (2020) “The Epidemiological Characteristics of an Outbreak of 2019 Novel Coronavirus Diseases (COVID-19) — China” *China CDC Weekly*, 2020, 2(8): 113-122.

France, Spain, and Italy

Ined database, “Demographics of Covid-19 Deaths” Institut National d’Etudes Demographiques <https://dc-covid.site.ined.fr/en/data/>

S. Korea:

<https://www.cdc.go.kr/board/board.es?mid=a30402000000&bid=0030>.

Population counts are drawn from the United Nations World Population Prospects

U.S. mortality schedules are drawn from the Human Mortality Database.

**4. Comparison of Covid-19 mortality to past US epidemics by different metrics for
Figure 3**

The table below shows the parameters used to estimate Figure 3 in the main text.

Table A1. Parameters used to calculate main text Figure 3

Cause	Deaths from epidemic cause (thousands)	Average age of death from epidemic cause	Remaining life expectancy per death from epidemic cause*	Reference year (non-epidemic)	Deaths from all causes in reference year (millions)	Average age of death from all causes in reference year*
Spanish Flu	675	30	38.8	1920 (1933)****	1.4	15.1
HIV	675	40	38.0	1990**	2.1	11.0
Opioid Overdose	770	42	31.6	2010**	2.5	9.8
Covid-19	1,000	80	9.6	2017	3.0	10.2***

* Calculated as life expectancy at average age of death for both epidemic cause and all causes.

** Reference year's include epidemic mortality, but at relatively low levels.

*** Differs from value reported in main text which averaged full distribution across all ages. We use approximations here in order to use the same method for different causes of death.

**** Death count based on 1920, life table based on 1933, the first year available for national death registration.

5. Data Sources for Figure 3

HIV deaths

CDC (August 2016) “CDC Fact Sheet: Today’s HIV/AIDS Epidemic,” p.2 “Since the beginning of the epidemic, nearly 675,000 people with AIDS in the United States have died, ...”

<https://www.cdc.gov/nchhstp/newsroom/docs/factsheets/todaysepidemic-508.pdf>

The mean age of HIV-caused death was estimated by tabulating all HIV deaths from 1979 to 1998 in 5-year age groups and computing the mean age assuming average deaths in mid-interval.

Opioid overdoses

The NIDA website, updated March 20, gives totals for 1999 to 2018. The total for "Number of National Drug Overdose Deaths Involving Select Prescription and Illicit Drugs" is 770,000. The total for Overdose deaths involving any opioid is 446,000.

<https://www.drugabuse.gov/related-topics/trends-statistics/overdose-death-rate>

The mean age of death for opioid overdoses is computed from CDC Wonder data tabulated by KFF for 2010.

<https://www.kff.org/other/state-indicator/opioid-overdose-deaths-by-age-group>

Spanish Influenza

“The number of deaths was estimated to be at least 50 million worldwide with about 675,000 occurring in the United States.”

CDC <https://www.cdc.gov/flu/pandemic-resources/1918-pandemic-h1n1.html>

The mean age of death was estimated from distributions graphed in Gagnon et al. (2013)

Gagnon A, Miller MS, Hallman SA, et al. Age-specific mortality during the 1918 influenza pandemic: unravelling the mystery of high young adult mortality. *PLoS One*. 2013;8(8):e69586. Published 2013 Aug 5. doi:10.1371/journal.pone.0069586