Supplementary appendix to:
Historically high excess mortality during the COVID-19 pandemic in Spain, Sweden and Switzerland

Kaspar Staub[a,†,*], Radoslaw Panczak[b,†], Katarina L. Matthes[a], Joël Floris[a,c], Claudia Berlin[b], Christoph Junker[d], Rolf Weitkunat[d], Svenn-Erik Mamelund[e], Matthias Egger[b,f,g], Marcel Zwahlen[b,‡], and Julien Riou[b,‡]

[a]Institute of Evolutionary Medicine, University of Zurich, Switzerland
[b]Institute of Social and Preventive Medicine, University of Bern, Switzerland
[c]Department of History, University of Zurich, Switzerland
[d]Federal Statistical Office, Neuchâtel, Switzerland
[e]Centre for Research on Pandemics & Society, Oslo Metropolitan University, Norway
[f]Population Health Sciences, Bristol Medical School, University of Bristol, UK
[g]Centre for Infectious Disease Epidemiology and Research, University of Cape Town, Cape Town, South Africa
[†]contributed equally
[‡]contributed equally
[*]Corresponding author (`kaspar.staub@iem.uzh.ch`)

December 8, 2021

Contents

1 Material and methods

1.1 Historical background of excess mortality studies

Demographers have used excess mortality since the 1830s to describe monthly and seasonal mortality fluctuations [1, 2]. In the 1850s, William Farr used expected versus observed crude death rates to identify places and populations that might benefit from sanitary interventions in England [3]. An early application of the concept of excess mortality during a pandemic occurred when Switzerland's federal authorities compared monthly deaths in 1890 with adjacent years in the 1890 influenza pandemic [4].

The concept of excess mortality has frequently been used to assess the overall impact of historical pandemics on specific populations. For example, international comparisons have been made for specific pandemics [5, 6, 7, 8, 9], and different historical pandemics have been compared with each other for particular countries [10, 11]. Recently, a few studies based on non-continuous data have been published comparing excess mortality in 2020 with 1918 [12, 13]. A study comparable to ours based on weekly data focused exclusively on Sweden and showed similar results [14].

1.2 Data sources and availability

Monthly all-cause death numbers for Switzerland (1877-2021) and Sweden (1851-2021) were provided by the Swiss Federal Statistical Office (SFSO)* and Statistics Sweden (SCB)†, they are freely available. Data covering first six months of 2021 in Switzerland were provided directly by SFSO.

For Spain, the monthly death figures 1941-2021 were accessible through the Spanish Statistical Office, the years 1908-1940 had to be transcribed from historical reports provided on their webpage ‡. Data for first six months of 2021 were compiled from experimental weekly data. It should be noted that all statistical authorities still consider the 2020 figures as provisional.

The annual population numbers as well as annual deaths by age group (in 1-year bands) were obtained from the Human Mortality Database (HMD, www.mortality.org). Most recent annual death counts (2019-2021 for Spain; 2020-2021 for Sweden and Switzerland) and population structure (2020-2021 for Spain) unavailable in HMD were obtained from provisional publicly accessible data repositories of the national statistics services of the three countries. All raw and prepared data as well as R code used in this analysis are provided in the study's GitHub repository §.

1.3 Data preparation

Monthly time series for each country until 2020 (Switzerland and Sweden) or 2019 (Spain) were combined with the provisional estimates from each country in most recent years. Yearly deaths by 1-year age groups were groupped into 10-year age bands and combined with relevant population denominator for each country X year X age group stratum.

1.4 Statistical model

We built a model where the linear predictor of deaths count in year i, month j and age group k $D_{i,j,k}$ depended on (1) an intercept by age group α_k, (2) a yearly linear trend β_1, (3) a periodic component $\beta_{2,\dots,5}$ based on four cosine and sine functions to account for temporal trends and seasonal variability in mortality, and (4) the total population in the age group $P_{i,k}$ included as an offset:

$$D_{i,j,k} = \alpha_k + \beta_1 i + \beta_2 \sin\left(\frac{2\pi j}{12}\right) + \beta_3 \sin\left(\frac{4\pi j}{12}\right) + \beta_4 \cos\left(\frac{2\pi j}{12}\right) + \beta_5 \cos\left(\frac{4\pi j}{12}\right) + \log(P_{i,k}) \tag{1}$$

The quantities $D_{i,j,k}$ were treated as latent variables, since deaths counts were not available with this level of precision. Rather, available data consisted of overall death counts by month $\mathbb{M}_{i,j}$ and age group-specific death counts by year $\mathbb{A}_{i,k}$. The model was jointly fitted to both types of data. The sum of $D_{i,j,k}$ by month was fitted to overall death counts by month with a negative binomial likelihood:

*Swiss Federal Statistical Office, Todesfälle nach Monat und Sterblichkeit seit 1803 (2021), (available at https://www.bfs.admin.ch/bfs/de/home/statistiken/kataloge-datenbanken/daten.assetdetail.14387168.html).

†Statistics Sweden (SCB), Births and deaths per month by sex. Year 1851-2020 (2021), (available at https://www.statistikdatabasen.scb.se/pxweb/en/ssd/START__BE__BE0101__BE0101G/ManadFoddDod/).

‡Instituto Nacional de Estadistica (INE), Anuarios Estadísticos (2021), (available at https://www.ine.es/inebaseweb/libros.do?tntp=25687).

§https://GitHub.com/RPanczak/ISPM_excess-mortality/

$$M_{i,j} = \sum_k D_{i,j,k} \tag{2}$$

$$\mathbb{M}_{i,j} \sim \text{neg-bin}\left(M_{i,j}, \phi\right) \tag{3}$$

where ϕ is the overdispersion parameter. Simultaneously, the sum of $D_{i,j,k}$ by age group was fitted to age group-specific death counts by year with a multinomial likelihood:

$$A_{i,k} = \sum_j D_{i,j,k} \tag{4}$$

$$N_i = \sum_j A_{i,k} \tag{5}$$

$$\mathbb{A}_{i,k} \sim \text{multinom}\left(N_i, \frac{A_{i,k}}{N_i}\right) \tag{6}$$

The joint likelihood can thus be expressed as:

$$\Pr(\mathbb{M}, \mathbb{A}|\alpha, \beta_1, \ldots, \beta_5, \phi) = \prod_{i,j} \text{neg-bin}\left(\mathbb{M}_{i,j}|M_{i,j}, \phi\right) \times \prod_{i,k} \text{multinom}\left(\mathbb{A}_{i,k}\middle|N_i, \frac{A_{i,k}}{N_i}\right) \tag{7}$$

This approach of "stratification and joint likelihood" was inspired by a recent work focusing on the COVID-19 infection-fatality ratio [15].

The model was implemented in Stan (the full code is available in the `stan` folder in study's GitHub repository) [16]. We selected weakly informative prior distributions for all model parameters [17, 18], that is normal distributions with mean 0 and standard deviation 10 for intercept and yearly and seasonal slope parameters and a Cauchy distribution with location 0 and scale 5 for the inverse of the overdispersion parameter. We estimated the posterior distributions with Stan's default Hamiltonian Monte Carlo algorithm [19] by sampling 1000 iterations after 1000 iterations for warm-up in four independent chains. We assessed convergence using the Gelman-Rubin convergence diagnostic and sampling efficiency using effective sample size. These diagnostics did not reveal any issues with mixing and convergence in all settings.

1.5 Estimating excess deaths

The procedure to obtain excess deaths for year i (by month or by age group) was as follows:

1. We fitted the model to the five previous years $\{i-5, \ldots, i-1\}$ and obtained posterior samples for all parameters.
2. From these posterior samples, we computed expected values of $D_{i,j,k}$ for year i using equation 1.
3. We then summed these quantities by month or age group (equations 2 or 4), and drew values from the corresponding probability distributions (equations 3 or 6), obtaining a set of expected values of death counts for year i by month or by age group.
4. We then subtracted the observed number of deaths on year i, and summarized the samples by their mean and 2.5% and 97.5% quantiles, obtaining point estimates and 95% credible intervals of excess deaths for year i by month or by age group.

This Bayesian approach ensured a full propagation of uncertainty from the data into the expected values and thus the estimates of excess deaths. Since five years of data is needed to calculate expected counts the estimates of excess deaths are available five years after the start of data collection in each country. We first calculated expected deaths excluding data from pandemic years of 1890, 1918 and 1957, and then also calculated expected deaths using all available data (shown by the example of 1918 in Supplementary Figure S3).

1.6 Monthly vs. weekly data

A comparison with the same seasonality adjustments between weekly[¶] and monthly aggregated calculation of excess mortality (unadjusted for age) for Switzerland, Spain, and Sweden showed small differences in the expected number of deaths (between -0.4% and +0.9%) for the full calendar years where the data was available between 2005 and 2020 (Supplementary Figure S4).

[¶]Short-term Mortality Fluctuations (STMF) data series, (available at `https://www.mortality.org/`).

2 Supplementary table

	Baseline		Sensitivity 1*	Sensitivity 2†		Sensitivity 3‡		Sensitivity 4§	
Country, Year	**Excess**	**95% CrI**	**Excess**	**Excess**	**95% CrI**	**Excess**	**95% CrI**	**Excess**	**95% CrI**
Switzerland									
1889	1,380	(-1,550; 4,255)	1,100	1,300	(-1,650; 4,130)	1,680	(-865; 4,190)	-410	(-3,405; 2,510)
1890	3,455	(705; 6,120)	3,380	3,360	(605; 6,015)	2,395	(-310; 5,020)	1,790	(-1,110; 4,565)
1891	1,920	(-1,265; 5,010)	1,975	1,585	(-1,605; 4,520)	1,950	(-825; 4,670)	990	(-2,140; 4,045)
1917	3,810	(1,155; 6,395)	3,630	3,855	(1,175; 6,505)	4,035	(1,240; 6,670)	1,905	(-925; 4,630)
1918	24,730	(21,925; 27,455)	24,460	24,625	(21,690; 27,465)	25,725	(22,845; 28,480)	23,705	(20,540; 26,705)
1919	3,730	(510; 6,835)	3,820	3,700	(515; 6,790)	4,600	(1,895; 7,275)	3,455	(200; 6,540)
1956	1,535	(-775; 3,750)	1,555	1,470	(-865; 3,725)	1,695	(-505; 3,905)	1,555	(-630; 3,710)
1957	-920	(-3,680; 1,805)	-1,200	-1,035	(-3,835; 1,675)	-180	(-2,650; 2,250)	-25	(-2,440; 2,295)
1958	-3,705	(-6,950; -605)	-3,360	-3,190	(-6,315; -185)	-2,980	(-5,535; -475)	-2,760	(-5,095; -380)
2019	165	(-3,210; 3,400)	280	320	(-2,950; 3,460)	395	(-2,200; 2,895)	600	(-1,865; 2,975)
2020	8,430	(5,255; 11,450)	8,810	8,685	(5,575; 11,675)	8,105	(5,505; 10,630)	7,360	(4,395; 10,240)
2021	-1,000	(-3,085; 955)	-845	-715	(-2,735; 1,220)	-795	(-2,715; 1,010)	-925	(-3,280; 1,375)
Sweden									
1889	2,495	(-1,400; 6,155)	2,785	2,540	(-1,280; 6,290)	290	(-3,385; 3,790)	-1,125	(-5,425; 2,960)
1890	8,810	(4,870; 12,575)	8,720	8,525	(4,475; 12,495)	7,380	(3,750; 10,945)	7,060	(2,930; 11,055)
1891	5,295	(800; 9,480)	6,365	5,570	(1,045; 9,800)	6,515	(2,655; 10,245)	5,875	(1,405; 10,255)
1917	-2,640	(-7,970; 2,330)	-2,980	-2,865	(-8,140; 2,105)	-2,945	(-7,240; 1,245)	-645	(-4,165; 2,725)
1918	25,935	(20,595; 31,085)	25,505	25,695	(20,310; 30,885)	25,735	(21,420; 29,845)	27,700	(23,945; 31,250)
1919	7,245	(1,180; 12,965)	8,260	7,140	(1,020; 12,835)	5,875	(1,265; 10,350)	7,395	(3,625; 11,080)
1956	1,080	(-2,395; 4,445)	1,585	1,450	(-2,085; 4,910)	1,050	(-1,825; 3,820)	880	(-2,410; 4,065)
1957	2,760	(100; 5,340)	3,040	2,990	(395; 5,515)	3,325	(410; 6,220)	3,135	(-470; 6,575)
1958	-360	(-3,505; 2,720)	1,040	620	(-2,465; 3,605)	-315	(-3,565; 2,800)	-800	(-4,875; 3,060)
2019	-4,365	(-8,015; -795)	-4,390	-4,255	(-7,810; -730)	-3,350	(-6,330; -345)	-4,380	(-7,125; -1,695)
2020	7,655	(4,200; 10,985)	8,465	8,405	(4,960; 11,690)	6,760	(3,635; 9,910)	2,600	(-690; 5,750)
2021	--	--	1,700	750	(-1,845; 3,160)	--	--	--	--
Spain									
1917	11,645	(-13,875; 36,180)	12,155	12,020	(-13,400; 36,025)	22,060	(-900; 44,430)	20,950	(1,005; 40,565)
1918	241,660	(215,790; 266,105)	239,680	239,495	(214,035; 264,175)	243,240	(218,610; 266,700)	241,660	(215,790; 266,105)
1919	28,835	(1,100; 55,560)	23,340	22,835	(-4,940; 49,265)	24,670	(-165; 48,125)	45,235	(21,520; 67,910)
1956	32,305	(9,145; 53,685)	39,455	33,085	(10,205; 54,805)	33,355	(14,275; 51,825)	30,730	(15,545; 45,480)
1957	13,225	(-2,660; 28,785)	10,715	11,675	(-4,495; 27,310)	28,360	(7,820; 48,025)	21,360	(4,365; 37,570)
1958	-28,160	(-47,480; -9,435)	-32,630	-28,615	(-48,745; -9,615)	-11,425	(-33,630; 9,845)	-27,265	(-44,020; -10,910)
2019	-13,370	(-34,800; 7,185)	-20,505	-19,885	(-41,320; 1,025)	-8,505	(-26,810; 9,205)	-10,460	(-31,415; 9,695)
2020	72,330	(51,940; 92,410)	65,960	65,495	(44,510; 85,625)	65,045	(47,040; 81,970)	69,225	(47,980; 89,670)
2021	10,865	(-3,425; 24,715)	10,110	10,980	(-3,525; 24,785)	9,350	(-3,005; 21,775)	14,245	(-265; 28,120)

* Serfling, Poisson, unadjusted for age
† Serfling, NB, unadjusted for age
‡ Serfling, NB, age adjusted, alternative time window, no trim
§ Serfling, NB, age adjusted, alternative time window, trim

Table S1: Results of four sensitivity analyses for selected years.

3 Supplementary figures

Figure S1: Monthly numbers of deaths in Spain, Sweden, and Switzerland, displayed as differences between observed and expected number of deaths per 100,000 population (red=more, blue=less). The purple horizontal line marks the level of highest value from 20202021 period.

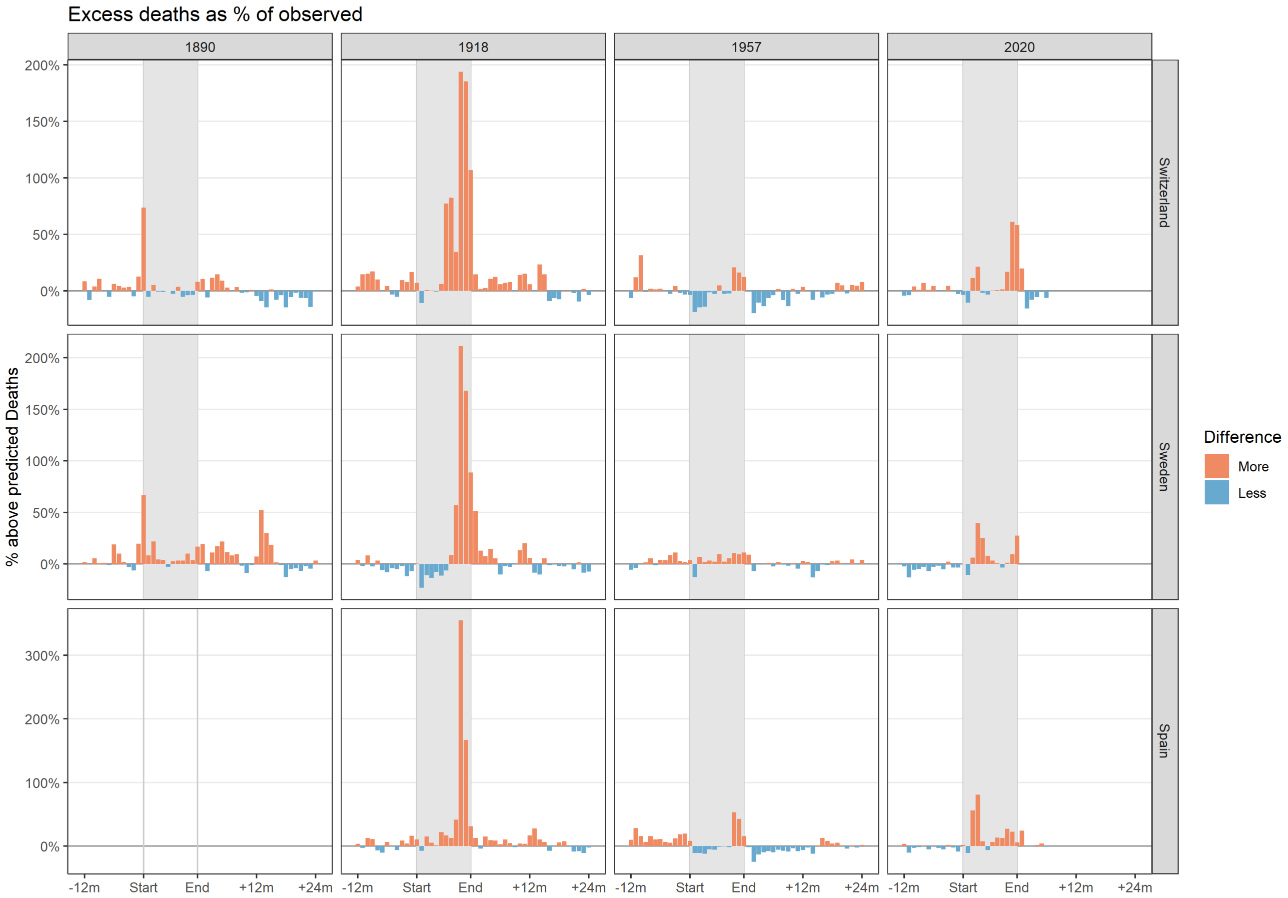

Figure S2: Detailed inspection of the strongest pandemic years 1890, 1918, 1957, and 2020 in all three countries. The differences between observed and expected values (red=more, blue=less) are displayed as percentages.

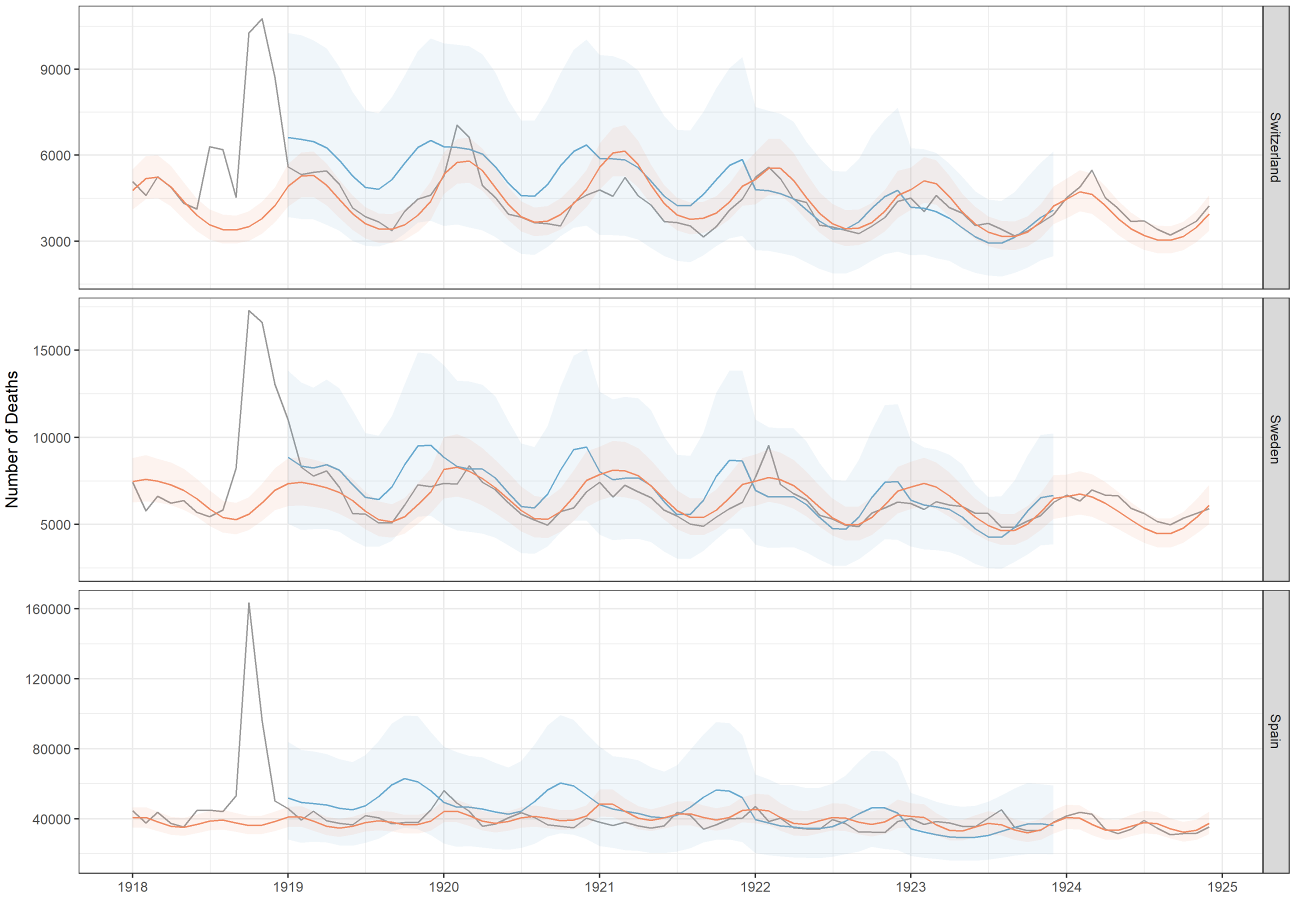

Figure S3: Exemplary visualization based on 1918, how big the difference in the calculated expected number of deaths is for the years 1919 and following, if the year 1918 is included in the calculations (blue line and confidence interval) or not (red line and confidence interval) (Black line: observed deaths). From 1924 onward the results of two methods are the same.

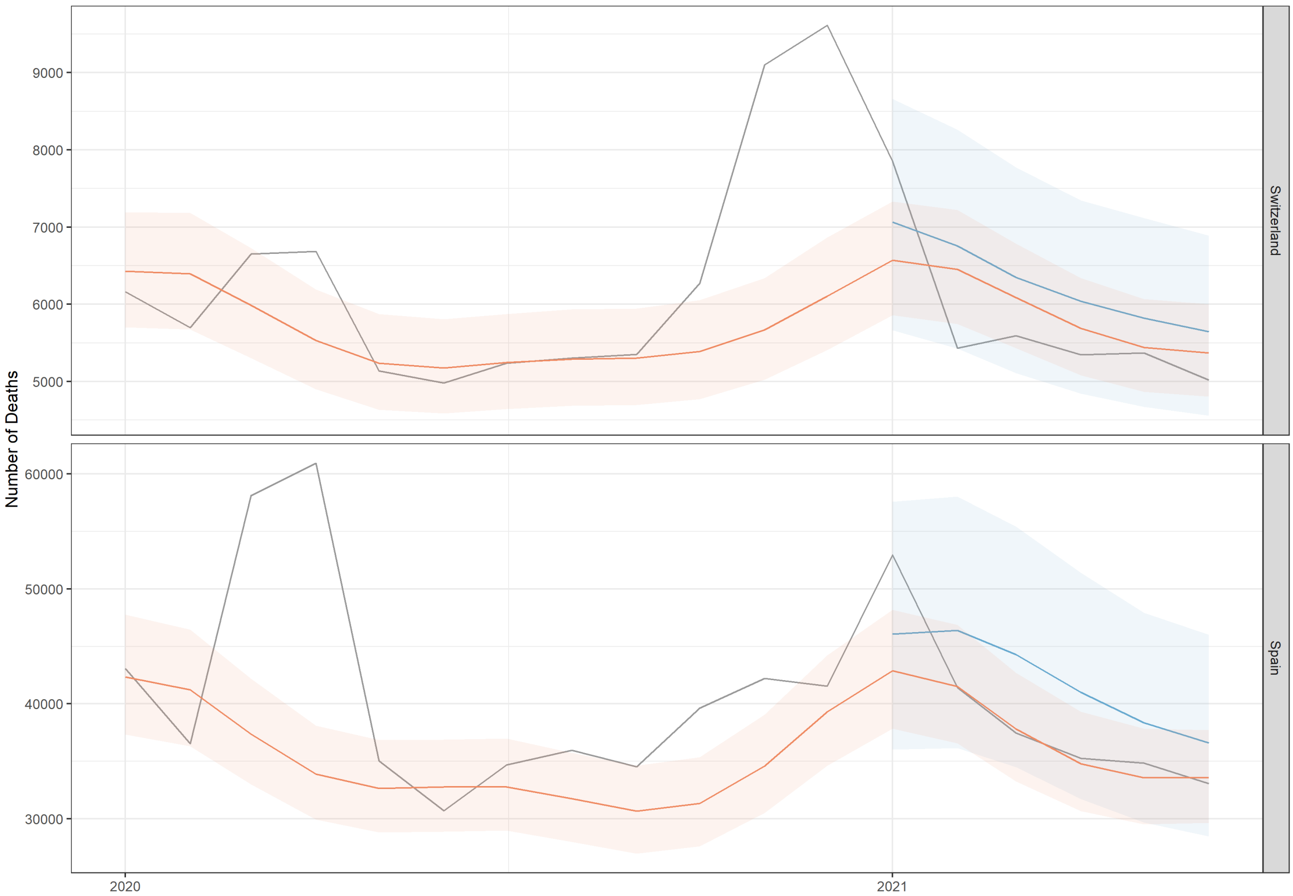

Figure S4: Exemplary visualization based on 2020, how big the difference in the calculated expected number of deaths is for the first six months of the year 2021, if the year 2020 is included in the calculations (blue line and confidence interval) or not (red line and confidence interval) (Black line: observed deaths).

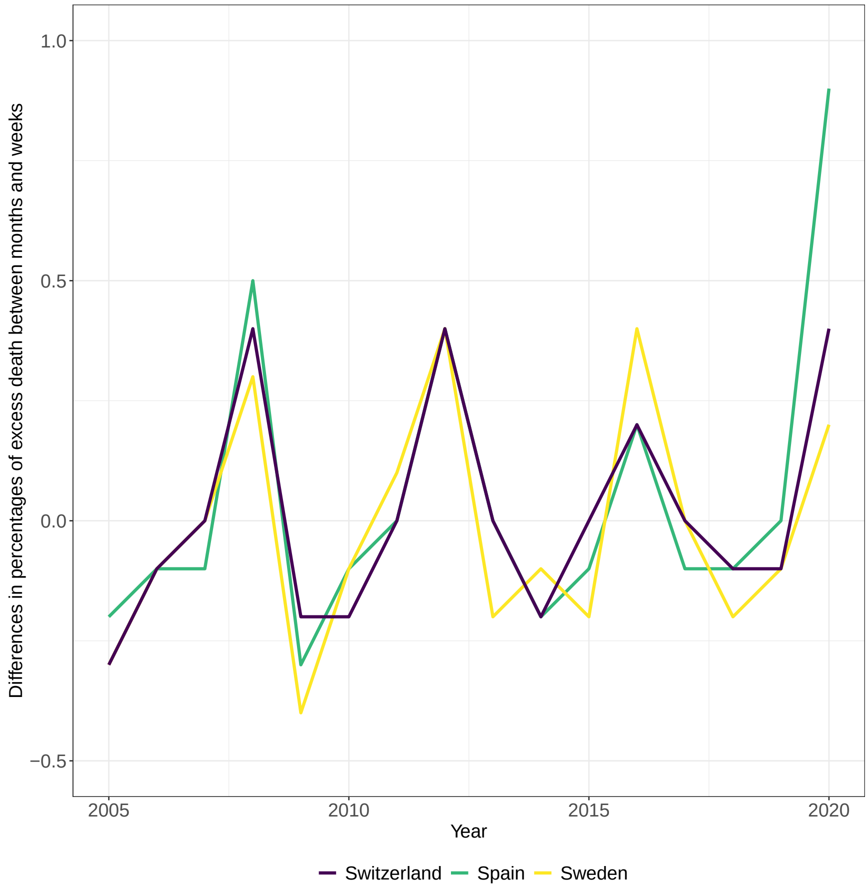

Figure S5: Comparison of weekly (STMF data from Human Mortality Database) vs. monthly (data as described and used in the main analyses) aggregated calculation of excess mortality for Switzerland, Spain, and Sweden and entire calendar years between 2005 and 2020. The differences between weekly and monthly calculated percentages of excess mortality range from -0.4% to 0.9%.

References

[1] Mallet E. Recherches historiques et statistiques sur la population de Genève : son mouvement annuel et sa longévité, depuis le XVIe siècle jusqu'à nos jours (1549-1833). Paris: Paul Renouard; 1837.

[2] Quetelet A. De l'influence des saisons sur la mortalité aux différens âges dans la Belgique. Bruxelles: M. Hayez; 1838.

[3] Mooney G. Professionalization in Public Health and the Measurement of Sanitary Progress in Nineteenth-Century England and Wales. Social History of Medicine. 1997;10(1):53–78.

[4] Schmid F. Die Influenza in Der Schweiz in den Jahren 1889-1894. Francke, editor. Bern; 1895. Available from: `https://digital.zbmed.de/gesundheitspflege/content/structure/5718371`.

[5] Dahal S, Jenner M, Dinh L, Mizumoto K, Viboud C, Chowell G. Excess mortality patterns during 1918–1921 influenza pandemic in the state of Arizona, USA. Annals of Epidemiology. 2018 may;28(5):273–280.

[6] Viboud C, Simonsen L, Fuentes R, Flores J, Miller MA, Chowell G. Global Mortality Impact of the 1957–1959 Influenza Pandemic. Journal of Infectious Diseases. 2016 mar;213(5):738–745.

[7] Ramiro D, Garcia S, Casado Y, Cilek L, Chowell G. Age-specific excess mortality patterns and transmissibility during the 1889–1890 influenza pandemic in Madrid, Spain. Annals of Epidemiology. 2018 may;28(5):267–272.

[8] Viboud C, Grais RF, Lafont BAP, Miller MA, Simonsen L. Multinational Impact of the 1968 Hong Kong Influenza Pandemic: Evidence for a Smoldering Pandemic. The Journal of Infectious Diseases. 2005 jul;192(2):233–248.

[9] Pastor-Barriuso R, Pérez-Gómez B, Hernán MA, Pérez-Olmeda M, Yotti R, Oteo-Iglesias J, et al. Infection fatality risk for SARS-CoV-2 in community dwelling population of Spain: nationwide seroepidemiological study. BMJ. 2020 nov;p. m4509.

[10] Buchholz U, Buda S, Reuß A, Haas W, Uphoff H. Todesfälle durch Influenzapandemien in Deutschland 1918 bis 2009. Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz. 2016 apr;59(4):523–536.

[11] Weinberger DM, Krause TG, Molbak K, Cliff A, Briem H, Viboud C, et al. Influenza Epidemics in Iceland Over 9 Decades: Changes in Timing and Synchrony With the United States and Europe. American Journal of Epidemiology. 2012 oct;176(7):649–655.

[12] Faust JS, Lin Z, del Rio C. Comparison of Estimated Excess Deaths in New York City During the COVID-19 and 1918 Influenza Pandemics. JAMA Network Open. 2020 aug;3(8):e2017527.

[13] Appleby J. UK deaths in 2020: how do they compare with previous years? BMJ. 2021 apr;p. n896.

[14] Ledberg A. Mortality of the COVID-19 Outbreak in Sweden in Relation to Previous Severe Disease Outbreaks. Frontiers in Public Health. 2021 feb;9.

[15] Hauser A, Counotte MJ, Margossian CC, Konstantinoudis G, Low N, Althaus CL, et al. Estimation of SARS-CoV-2 mortality during the early stages of an epidemic: A modeling study in Hubei, China, and six regions in Europe. PLOS Medicine. 2020 jul;17(7):e1003189.

[16] Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M, et al. Stan: A probabilistic programming language. Journal of Statistical Software. 2017 jan;76(1):1–32.

[17] Gelman A, Hill J, Vehtari A. Regression and other stories. Cambridge University Press; 2020.

[18] Gelman A, Jakulin A, Pittau MG, Su YS. A weakly informative default prior distribution for logistic and other regression models. The annals of applied statistics. 2008;2(4):1360–1383.

[19] Hoffman MD, Gelman A, Others. The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. J Mach Learn Res. 2014;15(1):1593–1623.