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

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], 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
[†]contributed equally
[‡]contributed equally
[*]Corresponding author (kaspar.staub@iem.uzh.ch)

January 10, 2022

Contents

1 Material and methods

The data and the statistical code are available at `https://github.com/RPanczak/ISPM_excess-mortality/`.

1.1 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 $\beta_{2,\ldots,5}$ 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:

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

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

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 [1].

The model was implemented in Stan (the full code is available in the `stan` folder in study's GitHub repository) [2]. We selected weakly informative prior distributions for all model parameters [3, 4], 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 [5] 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.

1.2 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, 1957 and 2020, and then also calculated expected deaths using all available data (shown by the examples of 1918 and 2020 in Supplementary Figures S3 and S4).

1.3 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.

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

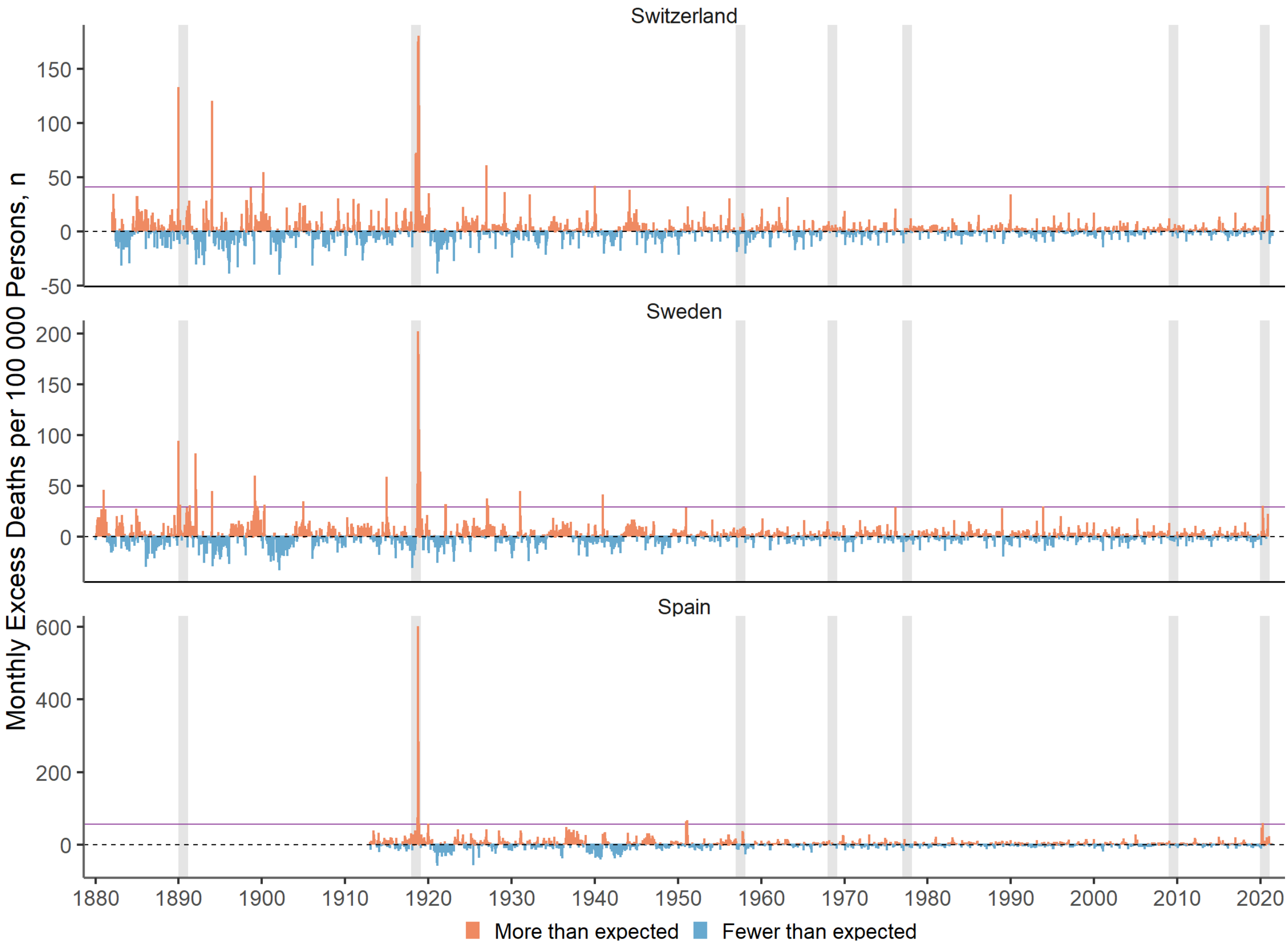

Figure S1: Monthly numbers of deaths in Spain, Sweden, and Switzerland, displayed as differences between observed and expected number of deaths per 100,000 population (red=more, blue=less). The purple horizontal line marks the level of highest value from 20202021 period.

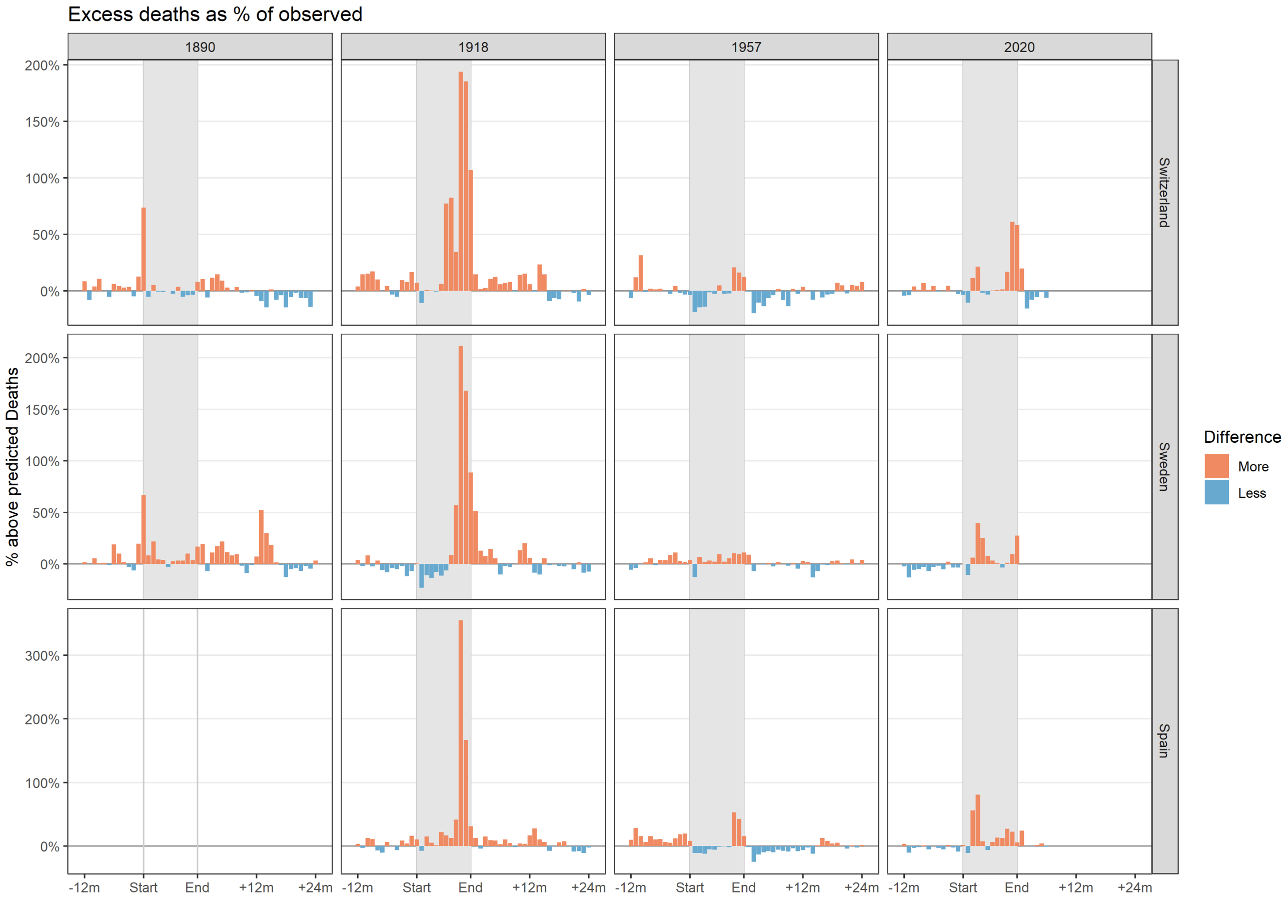

Figure S2: Detailed inspection of the strongest pandemic years 1890, 1918, 1957, and 2020 in all three countries. The differences between observed and expected values (red=more, blue=less) are displayed as percentages.

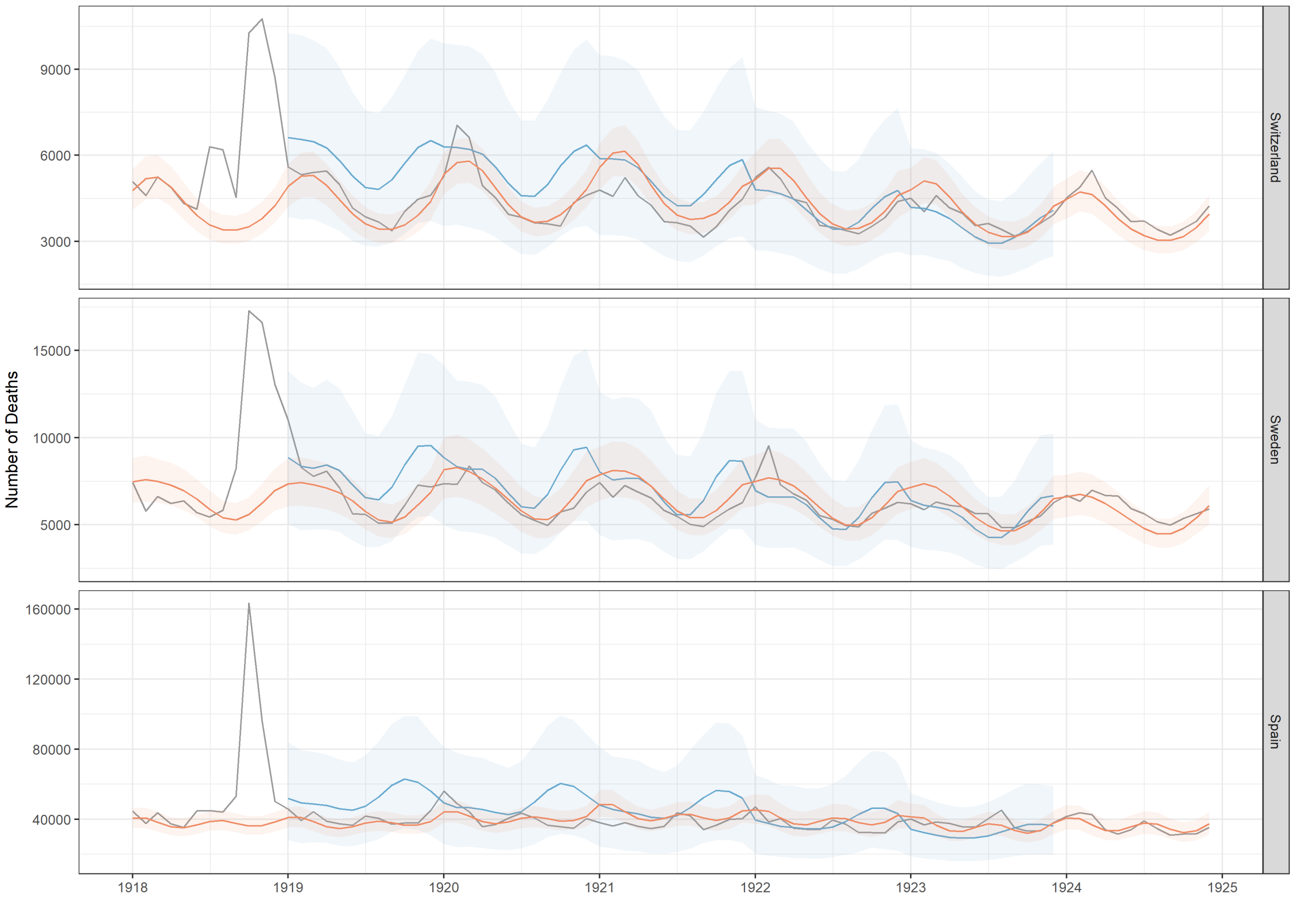

Figure S3: Exemplary visualization based on 1918, how big the difference in the calculated expected number of deaths is for the years 1919 and following, if the year 1918 is included in the calculations (blue line and confidence interval) or not (red line and confidence interval) (Black line: observed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

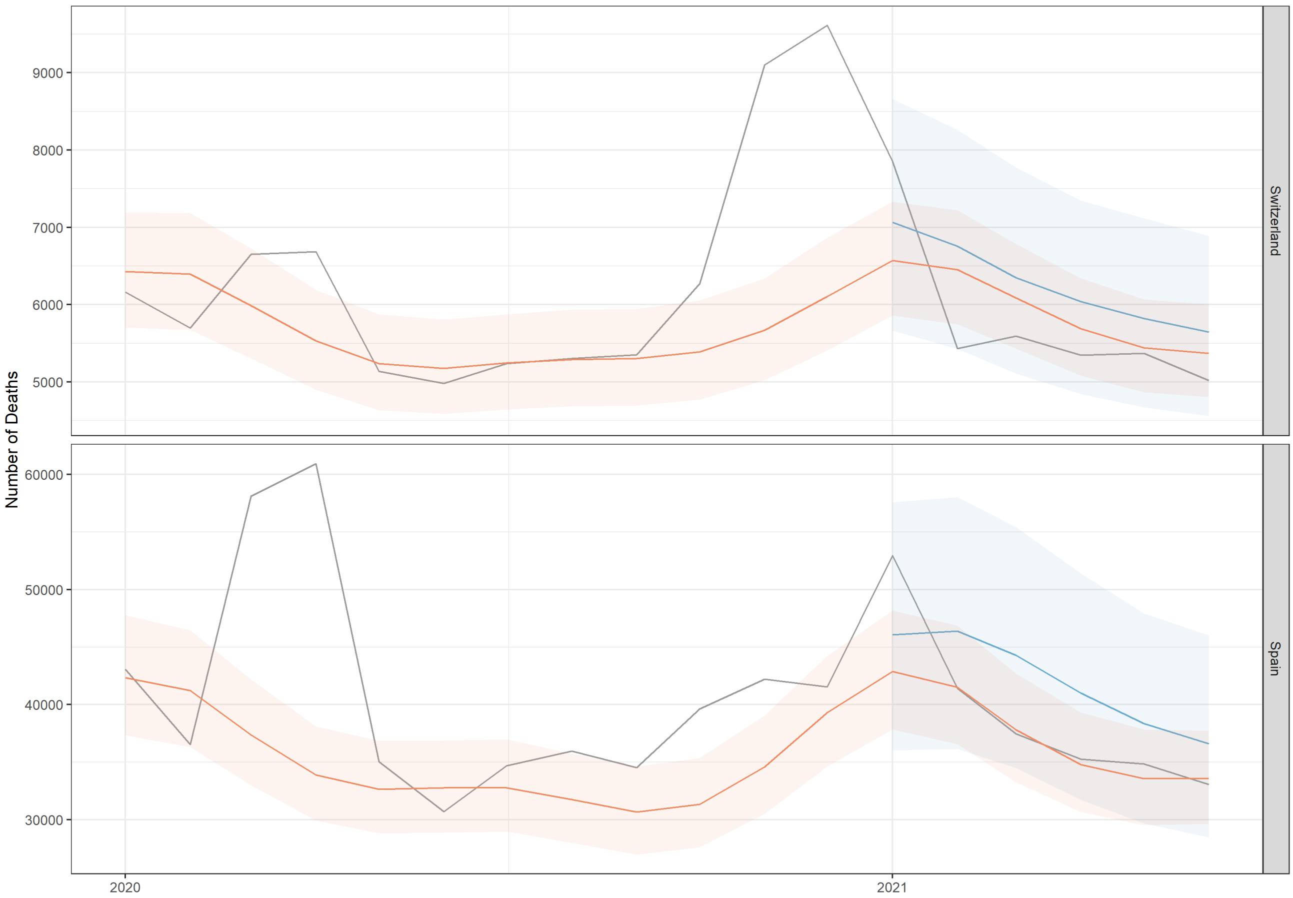

Figure S4: Exemplary visualization based on 2020, how big the difference in the calculated expected number of deaths is for the first six months of the year 2021, if the year 2020 is included in the calculations (blue line and confidence interval) or not (red line and confidence interval) (Black line: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

References

[1] 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.

[2] 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.

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

[4] 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.

[5] 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.