Excess Mortality in Germany 2020-2021

Christof Kuhbandner and Matthias Reitzner

Abstract

Using the period life tables provided by the Federal Statistical Office of Germany and the generation life table provided by the German Association of Actuaries the expected number of deceased in Germany is computed for the years 2020 and 2021, and then compared to the observed number (i.e. excess mortality). In addition, for the year 2021, the expected number of deceased is computed and compared to the observed number for each month separately, yielding the development of the excess mortality along 2021.

1 Introduction

In the last two years, the burden of the COVID-19 pandemic on mortality has been intensively discussed. Basically, it would be expected that the pandemic should lead to a large number of additional deaths. However, estimating the increase of mortality based on the number of reported so-called "COVID-deaths" has proven to be difficult due to the diagnostic problem that it is unclear whether a reported "COVID-death" died *because of* a SARS-CoV-2-infection or only *with* a SARS-CoV-2-infection. An obvious way to solve such diagnostic problems is to compare the number of observed all-cause deaths with the number of expected all-cause deaths.

The following analysis use data provided by the Federal Statistical Office of Germany and the German Association of Actuaries to estimate the expected number of deaths for 2020 and 2021. It turns out that the observed number of deaths in 2020 is very close to the expected number, whereas in 2021 a serious excess mortality is observed.

Examining in which age group this effect occurs, somehow surprisingly shows that the highest excess mortality is not observed in the highest age groups, which would have been expected due to the risk profile of "COVID-19"-disease, but in the middle-aged groups.

In a next step, we model the development of this excess mortality in 2021 from January to December. The results show a significant mortality excess in January which is nearly compensated until March. In April and May, a significant increase in mortality excess is observed, followed by a decrease up to August. However, other than at the beginning of the year, mortality excess remains above zero so that the increase in mortality excess in April and May is not compensated for. In September there is again a significant increase in excess mortality, which is more than doubled in November and December. This development does not seem to fit to the infection and death curves of COVID-19.

For a more detailed discussion and a comparison of the excess mortality in 2020 and 2021 to COVID-19 deaths, long-COVID-effects and side effects of the vaccination we refer to the discussion in [8, App. 5] (in german).

2 Yearly expected mortality: the method

The starting point for our investigations are the period life tables and population demographics available from the Federal Statistical Office of Germany. We denote by

- $l_{x,t}$ the number of x year old at January 1st in year t;
- $d_{x,t}$ the number of deceased in year t that died as x year old;
- $\hat{q}_{x,t}$ an estimate for the probability that someone dies as an x year old in year t.

Note that $d_{x,t}$ also contains deceased that have been $(x-1)$ years old at Januray 1st in year t and died as x year old. To compensate this problem, the 2017/2019 life table of the Federal Statistical Office of Germany [5] uses the method of Farr to estimate $\hat{q}_{x,t}$ (like most german life tables).

$$\hat{q}_{x,2019} = \frac{\sum_{t=2017}^{2019} d_{x,t}}{\frac{1}{2}\sum_{t=2017}^{2019}(l_{x,t} + l_{x,t+1}) + \frac{1}{2}\sum_{t=2017}^{2019} d_{x,t}} \tag{1}$$

A period life table thus takes into account only the mortality probabilities in these three years.

A much more complicated task is to compute generation life tables. Generation life tables observe the mortality development over a long period, roughly 100 years, smoothen the existing data, and in particular estimate the long term behaviour of the mortality probabilities. These probabilities have been decreasing within the last 100 years, and the common ansatz is to set

$$q_{x,t} = q_{x,t_0} e^{-F(x;t,t_0)}.$$

Here q_{x,t_0} is the smoothed life table in year t_0, and the German Association of Actuaries (DAV) models the trend underlying future mortality, the longevity trend function $F(x;t,t_0)$, via regression. In the year 2004 it turned out that the decrease of the mortality probabilities in the previous years has been steeper than expected, therefore the DAV life table DAV 2004 R [6] distinguishes between a higher short-term trend and a lower long-term trend. These trends are of high importance, developed and used for life annuities,whereas for life insurances the trend is mostly ignored. In addition, it seems that the longevity trend was flattening in the last years. Therefore, we have decided to use half the long-term trend function given by the DAV 2004 R,

$$F(x;t,t_0) = \frac{1}{2}(t-2019)F_{l,x}$$

where the numbers $F_{l,x}$ are contained in the DAV 2004 R table. We also decided to use the probabilities $\hat{q}_{x,2019}$ of the life table 2017/2019 by the Federal Statistical Office of Germany as the base life table, thus $t_0 = 2019$. It is well known that mortality probabilities for male and female differ substantially, therefore everything is computed for these two cases separately.

Putting things together, we define the mortality probability of an x year old male in year t by

$$q_{x,t} = \hat{q}_{x,2019} e^{-\frac{1}{2}(t-2019)F_{l,x}},$$

and for a y year old female in year t by

$$q_{y,t} = \hat{q}_{y,2019} e^{-\frac{1}{2}(t-2019)F_{l,y}}.$$

Now, for each individual the probability to die at age x is given by $q_{x,t}$, and hence a population of $l_{x,t}$ individuals produces a binomial distributed random number of deceased, with expected value

$$l_{x,t} q_{x,t}$$

as a first attempt. As is well known (and already discussed above in connection with Farr's method), this formula ignores those individuals which have been of age $(x-1)$ at the beginning of year t, and died as x year olds. To compensate for this missing piece, we follow the procedure proposed by De Nicola et al. [1]. For $x = 0, \ldots, 101$ the random number $D_{x,t}$ of x year old deceased in year t satisfies

$$\mathbb{E} D_{x,t} = \frac{1}{2} \left(l_{x-1,t} \frac{q_{x-1,t} + q_{x,t}}{2} + l_{x,t} \frac{q_{x,t} + q_{x+1,t}}{2} \right) \tag{2}$$

where $l_{x,t}$ is taken from the population table of the Federal Statistical Office of Germany [4]. For $x = 0$ we set $l_{-1,t} = l_{0,t+1}$ if available, $l_{-1,t} = l_{0,t}$ else, and $q_{-1,t} = q_{0,t}$.

Somehow strangely, the 2017/2019 life table by the Federal Statistical Office of Germany contains the data for $x = 0, \ldots, 100$, whereas the underlying population table containing $l_{x,t}$ is published only up to the age of 89. We approximate the missing $l_{90,t}, \ldots, l_{101,t}$ using the life table. For age $x \geq 100$ we set $q_x = q_{100}$. The resulting data should be compared to the observed data $d_{x,t}$ for $t = 2020, 2021$.

Some remarks are in order to contextualize the method and the results.

- There are several methods modelling mortality probabilities. One could use the base life table of the DAV 2004 R instead of the 2017/2019 life table of the Federal Statistical Office of Germany, or a mixture of both. One could use different longevity factors, or ignore them totally. The question, whether a serious excess mortality occurs for 2020 and 2021, heavily depends on this underlying data sets. Yet it seems that in most models the main point coincides with our results [1, 2, 7]: for 2020 the number of deceased is close to the expected value, whereas for 2021 there is a serious excess mortality (in our model of approximately 26.000 deceased).

- Modelling the longevity factors is a challenging task. For example, the Actuarial Association of Austria uses factors involving $\arctan\left(\frac{t}{100} - 20.01\right)$ which has serious advantages. The need for longevity factors depends heavily on the country, it seems for example that in Japan the mortality trend has already vanished and the mortality probabilities are more or less constant.

- As is common, we assume that the number of deceased follows a binomial distribution. This is the most natural assumption. It would imply that the variance is approximately the number of deceased, e.g. in Germany one million, and the standard deviation approximately 1.000 deceased. Yet the empirical standard deviation for the years 2010 to 2019 is approximately 14.000. This shows that the principal mathematical model underlying all life tables seems to lack some further randomization of $q_{x,t}$ which increases the variance.

- The mortality probability not only depends on age and gender, but also heavily on social status, profession, health condition, region, etc. As is common, the german life tables give average mortality probabilities. For a deeper investigation of COVID-19 mortality increase this should be taken into account.

- We use of the life table 2017/2019 by the Federal Statistical Office of Germany although it is not smoothed sufficiently, e.g. it seems that there is some kink around the age of 50. Nevertheless, we preferred this life table to the DAV 2004 R since it is much more recent.

- We have totally ignored migration effects.

- All these remarks raise the question, whether something like the 'precise expected number of deceased' really exists and show that the word "excess mortality" should be used with great care.

3 Yearly expected mortality: the results

Following the computations described in the previous section, we obtain the number of expected deceased in 2020 and 2021. The Federal Statistical Office of Germany provides the (raw) number of deceased in 2021 only in certain age ranges [3]. Therefore, the following table gives the number of deceased in the age ranges

$$\bar{x} \in \{[0,14],[15,29],[30,39],[40,49],[50,59],[60,69],[70,79],[80,89],[90,\infty)\}.$$

We set

$$D_{\bar{x},t} = \sum_{x\in\bar{x}} D_{x,t} \quad \text{and} \quad d_{\bar{x},t} = \sum_{x\in\bar{x}} d_{x,t}.$$

To compare the expected and the observed values, we use the relative difference

$$\frac{d_{\bar{x},t} - \mathbb{E}D_{\bar{x},t}}{\mathbb{E}D_{\bar{x},t}}.$$

age range $\bar{x}$	expected $\mathbb{E}D_{x,2020}$	observed $d_{x,2020}$	rel.diff.		expected $\mathbb{E}D_{x,2021}$	observed $d_{x,2021}$	reldiff.
0-14	3.521	3.306	-6,11 %		3.479	3.477	-0,07 %
15-29	3.933	3.844	-2,26 %		3.817	3.945	3,36 %
30-39	6.607	6.668	0,92 %		6.585	6.923	5,13 %
40-49	15.303	15.507	1,33 %		14.877	16.219	9,02 %
50-59	58.481	57.331	-1,97 %		57.705	59.292	2,75 %
60-69	117.111	118.460	1,15 %		118.456	126.340	6,66 %
70-79	197.847	201.957	2,08 %		190.335	203.908	7,13 %
80-89	377.425	378.406	0,26 %		392.535	396.831	1,09 %
90-∞	201.161	200.093	-0,53 %		205.003	203.767	-0,60 %
total 0-∞	981.389	985.572	0,43 %		992.792	1.020.702	2,81 %

As remarked above, the empirical standard deviation of $D_{[0,\infty),t}$ within the last years has been approximately 14.000 (the precise number clearly depends on the details of the underlying mathematical model), and expressing this standard deviation in terms of the relative error gives $1,42\%$. Thus, in year 2020 the observed number of deceased is extremely close the expected number, whereas in 2021 the difference is of order twice the empirical standard deviation, which is comparable to the maximal deviation observed during the years 2010 – 2019.

The following graph illustrates, that the deviation of the observed mortality from the expected mortality is not uniform over the different age groups, and, in particular, the structure changes from 2020 to 2021. A closer look reveals that the mortality excess observed in 2021 is almost entirely due to an above-average increase in deaths in the age groups between 15 and 79, in some of which high deviations of the observed number of deaths from the expected value are observed. The highest values are reached in the age group 40-49, where an increase in the number of deaths is observed that is nine percent higher than the expected values.

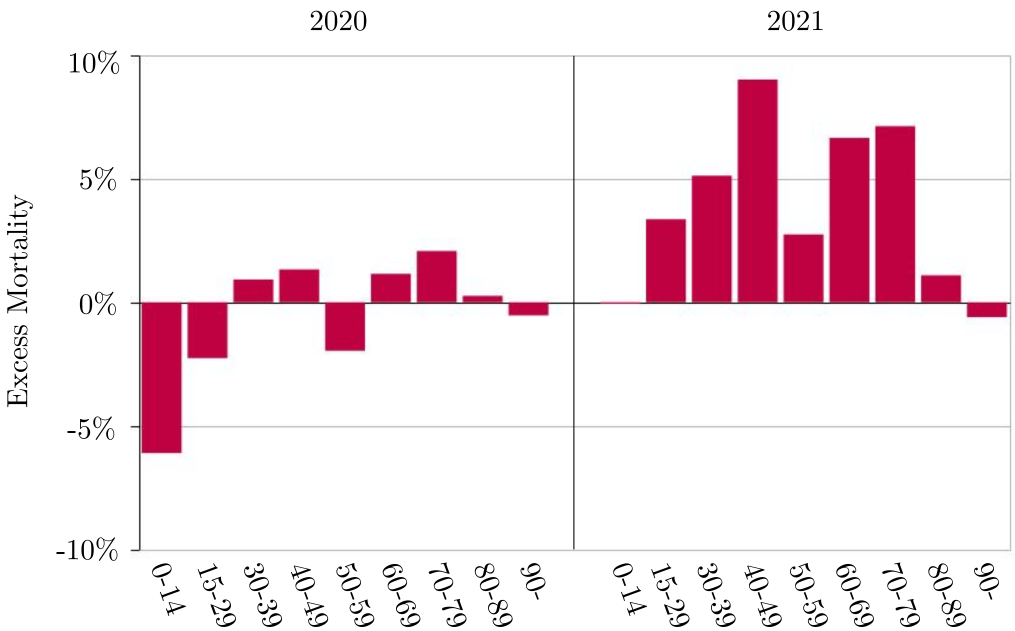

Fig. 1: Excess mortality in 2020 (left panel) and 2021 (right panel) in different age groups

4 Monthly expected mortality: the method

In the following two sections, we present a more detailed analysis of the number of deceased during the year 2021. It is well known that the mortality probabilities are not constant but differ from month to month with peaks at the beginning and the end of the year and also sometimes in summer when the weather is too hot (and depending on many other circumstances).

Unfortunately, the data basis for such investigations provided by the Federal Statistical Office of Germany is even weaker than the data on a yearly basis. Therefore, again several approximation steps have to be applied. We denote by

- $l_{x,t,m}$ the number of x year old at the 1st of month m in year t;
- $d_{x,t,m}$ the number of deceased in year t in month m that died as x year old;
- $\hat{q}_{x,t,m}$ an estimate for the probability that someone dies as an x year old in month m in year t.

We already know $l_{x,t} = l_{x,t,1}$. The Federal Statistical Office of Germany offers tables for $d_{\bar{x},t,m}$ in the age ranges $\bar{x} \in \{[0,14],[15,29],[30,39],[40,49],[50,59],[60,69],[70,79],[80,89],[90,\infty)\}$ which we use for the years $t = 2016, \ldots, 2021$, see [3]. We set

$$l_{\bar{x},t,m} = l_{\bar{x},t} - \sum_{i=1}^{m-1} d_{\bar{x},t,i}$$

ignoring again migration issues. As already discussed in Section 2 in principle the method of Farr or similar formulas have to be applied to estimate $\hat{q}_{x,t,m}$. Because of the thin data basis we will use a brute force method. We put directly

$$\hat{q}_{\bar{x},t,m} = \frac{d_{\bar{x},t,m}}{l_{\bar{x},t,m}} = \frac{d_{\bar{x},t,m}}{l_{\bar{x},t} - \sum_{i=1}^{m-1} d_{\bar{x},t,i}}$$

where we take into account a small error which is acceptable because the mortality probabilities are calculated for a whole age range, and thus the boundary effects are smaller than in (1) and (2). For the mortality trend we use the mean value of the mortality trends in each age group $F_{l,\bar{x}} = \frac{1}{|\bar{x}|} \sum_{x \in \bar{x}} F_{l,x}$ and set

$$q_{\bar{x},2021,m} = \frac{1}{4} \sum_{t=2016}^{2019} \hat{q}_{\bar{x},t,m} e^{-(2021-t)F_{l,\bar{x}}}.$$

A first approximation for the expected number of deceased in January should be

$$\tilde{D}_{\bar{x},2021,1} = l_{\bar{x},2021} q_{\bar{x},2021,m} e^{-\frac{m-1}{12} F_{l,\bar{x}}}$$

taking into account the mortality trend. For the other months $m = 2, \ldots, 12$, we in addition have to take into account that the number $l_{\bar{x},2021}$ should be decreased by the expected number of people who already died. This yields an iterative procedure. Set $\tilde{D}_{\bar{x},2021,1}$ as above, and for $m \geq 2$

$$\tilde{D}_{\bar{x},2021,m} = \left(l_{\bar{x},2021} - \sum_{i=1}^{m-1} \tilde{D}_{\bar{x},2021,i} \right) q_{\bar{x},2021,m} e^{-\frac{m-1}{12} F_{l,\bar{x}}}.$$

Yet due to our approximation errors, the sum $\sum_{m=1}^{12} \tilde{D}_{\bar{x},2021,m}$ does not coincide with our one year result $\mathbb{E}D_{\bar{x},2021}$, hence we have to normalize our approximation to be consistent with our results from Section 2. Thus we get the following approximation for the expected number of deceased,

$$\mathbb{E}D_{\bar{x},2021,m} = \frac{\mathbb{E}D_{\bar{x},2021}}{\sum_{i=1}^{12} \tilde{D}_{\bar{x},2021,i}} \tilde{D}_{\bar{x},2021,m}.$$

The resulting data should be compared to the observed data $d_{\bar{x},2021,m}$ for $m = 1, \ldots, 12$. The remarks made at the end of Section 2 apply similarly to the computations made in this section.

5 Monthly expected mortality: the results

Following the computations described in the previous section we obtain the number of expected deceased $\mathbb{E}D_{\bar{x},2021,m}$ for all months $m = 1, \ldots, 12$ in the age ranges

$$\bar{x} \in \{[0,14],[15,29],[30,39],[40,49],[50,59],[60,69],[70,79],[80,89],[90,\infty)\}.$$

Since it turned out in Section 3 that the highest mortality deviation occurs in the age groups $15 - 79$ we state the results for this age range. To compare the expected and the observed values, we again use the relative difference

$$\frac{d_{[15,79],2021,m} - \mathbb{E}D_{[15,79],2021,m}}{\mathbb{E}D_{[15,79],2021,m}}.$$

	m=1	m=2	m=3	m=4	m=5	m=6
$\mathbb{E}D_{[15,79],2021,m}$	35.879	34.110	36.983	32.383	32.083	30.388
$d_{[15,79],2021,m}$	40.027	33.092	34.298	35.171	34.854	32.204
reldiff.	11,56 %	-2,98 %	-7,26 %	8,61 %	8,64 %	5,98 %

	m=7	m=8	m=9	m=10	m=11	m=12
$\mathbb{E}D_{[15,79],2021,m}$	31.704	31.422	29.761	31.797	31.618	33.647
$d_{[15,79],2021,m}$	32.289	31.695	31.871	34.225	36.586	40.315
reldiff.	1,85 %	0,87 %	7,09 %	7,64 %	15,71 %	19,82 %

The results show a significant mortality excess in January which is nearly compensated until March. That is, by the end of March, mortality excess was close to zero. In April and May, a significant increase in mortality excess is observed, followed by a decrease up to August. However, other than at the beginning of the year, mortality excess remains above zero so that the increase in mortality excess in April and May is not compensated for. In September there is again a significant excess mortality, which is more than doubled in November and December. This finding is illustrated in Figure 2.

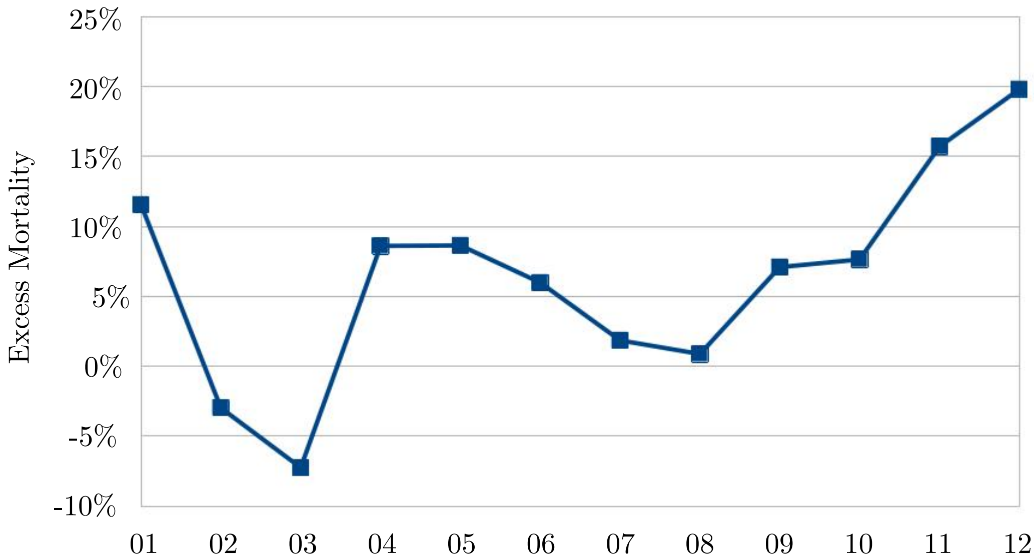

Fig. 2: Development of the monthly excess mortality from January to December 2021

References

[1] De Nicola, G., Kauermann, G., Höhle, G.: On assessing excess mortality in Germany during the COVID-19 pandemic. *AStA Wirtsch Sozialstat Arch*, link.springer.com/article/10.1007/s11943-021-00297-w (2022)

[2] De Nicola G., Kauermann G.: Altersadjustierte Übersterblichkeit 2021. In: CODAG Bericht Nr. 26, 20–21 www.covid19.statistik.uni-muenchen.de/pdfs/codag-bericht-26.pdf (2022)

[3] Federal Statistical Office of Germany (destatis): Death statistics. www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Sterbefaelle-Lebenserwartung/Tabellen/sonderauswertung-sterbefaelle.html (21.2.2022)

[4] Federal Statistical Office of Germany (destatis): Population statistics. https://www-genesis.destatis.de/genesis//online?operation=table&code=12411-0005&bypass=true&levelindex=1&levelid=1647003146226#abreadcrumb (21.2.2022)

[5] Federal Statistical Office of Germany (destatis): Life table 2017/2019. www-genesis.destatis.de/genesis//online?operation=table&code=12621-0001&bypass=true&levelindex=0&levelid=1647001436791 (21.2.2022)

[6] German Association of Actuaries (DAV): life table DAV 2004R. www.aktuar.de

[7] Kowall B., Standl F., Oesterling F., Brune B., Brinkmann M., Dudda M., Pflaumer P., Jöckel K., Stang A.: Excess mortality due to Covid-19? A comparison of total mortality in 2020 with total mortality in 2016 to 2019 in Germany, Sweden and Spain. *PLoS ONE* 16(8): e0255540. doi.org/10.1371/journal.pone.0255540 (2021)

[8] www.7argumente.de, Appendix 5: Excess mortality in 2021 (Die erhöhte Sterblichkeit im Jahr 2021), p.46-54 (2022)

Christof Kuhbandner
Universität Regensburg
Institut für Experimentelle Psychologie
93040 Regensburg
Germany

Matthias Reitzner
Universität Osnabrück
Institut für Mathematik
49069 Osnabrück
Germany