Mortality of the COVID-19 outbreak in Sweden in relation to previous severe disease outbreaks

Anders Ledberg*

Department of Public Health Sciences, Stockholm University
SE-106 91 Stockholm, Sweden

June 1, 2020

Abstract

Influenza viruses have caused disease outbreaks in human societies for a long time. Influenza often has rapid onset and relatively short duration, both in the individual and in the population. The case fatality rate varies for different strains of the virus, as do the effects on total mortality. Outbreaks related to coronavirus infections have recently become a global concern but much less is known about the dynamics of these outbreaks and their effects on mortality.

In this work, disease outbreaks in Sweden, in the time period of 1860-2020, are characterized and compared to the currently ongoing COVID-19 outbreak. The focus is on outbreaks with a sharp increase in all-cause mortality. Outbreak onset is defined as the time point when deaths counts starts to increase consistently for a period of 10 days. The duration of the outbreak is defined as the time period in which mortality rates are elevated. Excess mortality is estimated by standard methods.

In total there were 15 outbreaks detected in the time period, the first 14 were likely caused by influenza virus infections, the last by SARS-CoV-2. The mortality dynamics of the SARS-CoV-2 outbreak is shown to be similar to outbreaks due to influenza virus, and in terms of the number of excess deaths, it is the worst outbreak in Sweden since the 'Spanish flu' of 1918-1919.

1 Introduction

Influenza viruses of type A are known to have caused disease outbreaks at least since 19th century, and probably for much longer (Potter 2001). The severity of the disease caused by an influenza virus infection depends both on the properties of the virus (which strain) and on the acquired immunity and general health status of the infected individual (e.g., Cobey & Hensley 2017). The great majority recover completely from an infection, however, each year a number of persons die from consequences of influenza infections. The death toll at the

*anders.ledberg@su.se or anders.ledberg@gmail.com

level of the population depends on vaccination programs and nonpharmaceutical interventions aiming to reduce the spread of the virus (Bell et al. 2006).

Outbreaks caused by coronaviruses are thought to be a more recent phenomenon; the first reported outbreak was caused by SARS-CoV-1 in 2003 (Peiris et al. 2003), and the most recent is the still ongoing pandemic caused by SARS-CoV-2 (Huang et al. 2020). Vaccines against coronaviruses are still under development (Amanat & Krammer 2020) and measures available to control the outbreaks are currently limited to nonpharmaceutical interventions.

It has long been recognized that influenza outbreaks often are associated with an increase in all-cause mortality that exceeds the increase directly attributed to influenza and pneumonia (e.g., Collins 1932). Indeed, influenza seasons, and outbreaks, can be reliably detected from all-cause mortality data (e.g., Collins 1932, Simonsen et al. 1997). In fact, many countries and regions of the world monitor influenza by detecting when, and by how much, the number of deaths per week exceed a preset, model based, threshold. Outbreaks such as the 1918-1920 influenza pandemic are characterized by a high attack rate and often lead to a rapidly increasing number of deaths during a short time period. Consequently, with access to daily mortality counts, it should be possible to detect outbreaks by looking at the local rate of change of the number of recorded deaths. Here such an approach is developed and applied to daily death counts from Sweden in the time period of 1860-2020. The excess mortality caused by the ongoing outbreak of COVID-19 is related to the 14 most severe outbreaks during the previous 160 years.

2 Methods

2.1 Data

Two sources of daily counts of deaths from all causes were used. For the years 1860-2014, data were obtained from 'Swedish Book of Death' issued by the The Federation of Swedish Genealogical Societies (Sveriges Släktforskarförbund 2017). This is a database compiled from a range of official sources and contains information on times and places of births and deaths for persons that have died in Sweden since 1860. The coverage is almost complete. Data from 2015 until present were obtained from the website of Statistics Sweden (`www.scb.se`). Since the COVID-19 outbreak is still ongoing (at the time of writing), data from the most recent days are preliminary and will most likely be adjusted upwards as more data comes in. Data on total population size was obtained from Statistics Sweden. The complete time series used in this work will be made available upon request.

2.2 Outbreak detection

The onsets of putative outbreaks were detected by a two-step procedure: First the rate of change (derivative) of death numbers was computed from a smoothed version of the time series of daily death counts. Second, time sequences, where the derivative exceeded a threshold for a number of consecutive days, were identified, and the first of these days were taken as the onset. The optimal values of the three parameters involved in this procedure: i.e, width of the

smoothing kernel, threshold of the derivative, and the number of consecutive days, were determined by applying the procedure to data from 1860 to 2017, i.e., not involving the COVID-19 outbreak. The offset of the outbreaks were taken as the first time point, after the peak of the outbreak, where the derivative returned to zero. Two of the outbreaks detected using this method, both occurring before 1886, were after visual inspection determined not to qualify as outbreaks (the mortality did not exceed the background level). In some cases, the offsets of the outbreaks were adjusted manually since the automatic detection tended to overestimate the duration. For the COVID-19 outbreak, the real offset could not be determined since the outbreak is still ongoing; here the last data point available was taken as the offset.

2.3 Excess mortality estimation

In order to estimate the excess mortality during an outbreak, a baseline level of expected mortality is needed. Here a variant of the method described by (Simonsen et al. 1997) was used. This consists in fitting a model to data where the time points of the outbreaks have been removed, and subsequently using this model to predict (forecast) what values would have been observed if the outbreak had not happened. In particular, the number of deaths per day, N_t, was assumed to follow a Poisson distribution with a time-dependent expected value obeying the following model

$$\log\{E(N_t)\} = \mu + \beta t + \alpha \text{month}(t), \tag{1}$$

where month(t) is a categorical variable denoting the month corresponding to time t. Note that time, t, is expressed in units of days. Separate models for each outbreak were fitted to data from five years prior to the onset of the outbreak.

Excess mortality was estimated as the difference between observed and expected mortality. Since the size of the Swedish population has changed substantially over the time period, excess mortality was also expressed in terms of per 100,000 population. The number for the total population size was taken as the population the last of December the year before the onset of the outbreak.

3 Results

Fifteen outbreaks were detected in the 160 years of mortality data analyzed (Fig 1). These outbreaks correspond to the 15 highest peaks in the data. The method used to estimate excess mortality is illustrated in Fig 2, which also shows the COVID-19 outbreak in more detail. Data characterizing the 15 outbreaks are tabulated in Table 1. In terms of excess deaths, the COVID-19 outbreak (last outbreak) is the worst since the outbreak in 1931, and when standardized by the total population size, it is the worst outbreak since 1941.

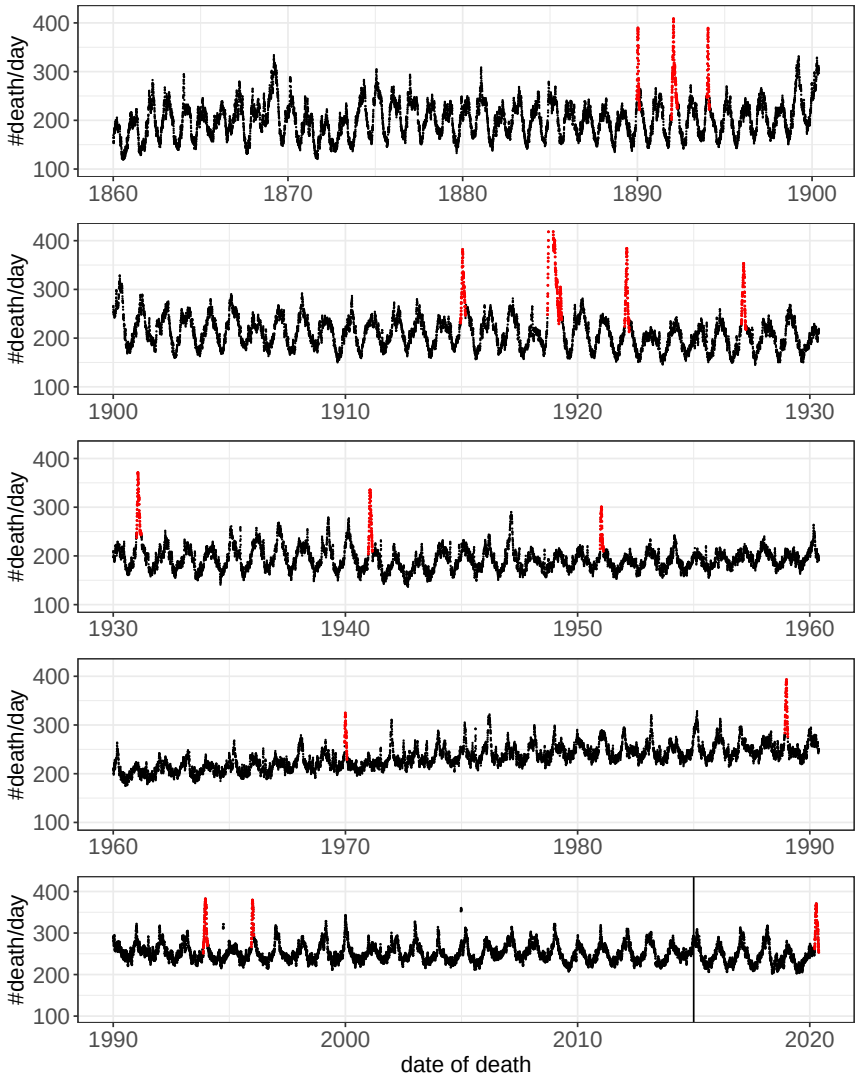

Figure 1: Number of deaths per day as a function of time. For visualization the original count data have been smoothed by rectangular window of length seven. That is, each value represent the mean value of the neighboring seven values. In red are shown the outbreaks detected. Note that the peak of the outbreak of 1918-1919 has been truncated. The vertical line indicates the time point at which there was a change in data source (see Methods).

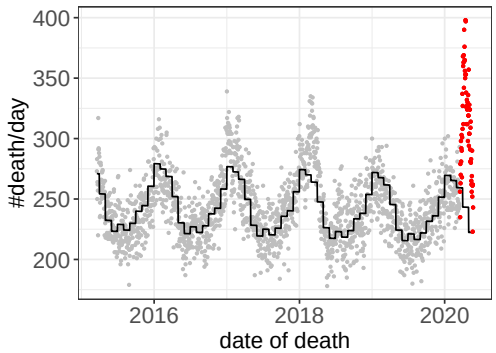

Figure 2: Death counts prior to, and during, the COVID-19 outbreak. Observed values prior to the outbreak in gray. Values during the outbreak in red. The predicted values based on the Poisson model (Eq 1) are shown in black.

onset	offset	duration (days)	excess mortality	excess mortality (per 10^5 pop.)
1889-12-21	1890-02-13	54	5011	105.5
1891-12-14	1892-04-14	122	7089	148.2
1893-12-27	1894-02-13	48	2665	55.4
1914-12-17	1915-03-04	77	5145	91.3
1918-09-15	1919-04-25	222	38795	668.8
1922-01-11	1922-03-23	71	3557	59.7
1927-01-21	1927-04-03	72	4587	75.5
1931-01-04	1931-03-15	70	4571	74.4
1940-12-28	1941-03-01	63	3147	49.6
1950-12-22	1951-02-10	50	2310	33.1
1969-12-20	1970-01-25	36	1335	16.8
1988-12-05	1989-01-19	45	2789	33.1
1993-11-18	1994-01-24	67	2913	33.5
1995-12-14	1996-01-20	37	2579	29.2
2020-03-18	2020-05-24	67	4857	47.0

Table 1: Outbreaks detected in the data. Excess mortality is expressed as the difference between observed and expected deaths (see Methods), and is show both in absolute numbers as well as standardized to the size of the total population.

4 Discussion

Using daily death counts from 1860 until present, 15 disease outbreaks, characterized by rapidly increasing all-cause mortality, were detected. From previous work, and available official Swedish records on causes of death, the 14 first outbreaks can be attributed to influenza type A viruses. The last, and ongoing, outbreak is caused by SARS-CoV-2. In terms of all-cause mortality the time-course is similar for all the 15 outbreaks; a rapid onset and a slower return to baseline. Most outbreaks were relatively short in duration; all except two were less than three months long. The 1918-1919 outbreak (part of the 'Spanish flu' pandemic) was exceptional both in excess mortality and duration, and lasted for more than half a year. The duration of the current wave of the COVID-19 outbreak remains to be determined.

That all detected outbreaks have rapid increases in mortality is to some extent a consequence of the method used, it was after all designed to detect such rapid increases. However, it is clear from Fig 1 that the most prominent peaks in the 160 years of data were detected using this method. Note that the singular mortality increases caused by the sinking of MS Estonia the 28th of September 1994 and by the Indian Ocean tsunami of 26th of December 2004 were not classified as outbreaks by the method, even if the casualties, more than 500 at each occasion, led to a very sharp increase in the number of deaths around these dates. This shows that using only local (in time) information it is possible to reliably detect onsets of disease outbreaks from all-cause mortality data. Many of the outbreaks detected have been previously characterized, however, several have not, including some of the quite severe ones in the first three decades of the 20th century.

There is no clear threshold at which an increase in deaths become an "outbreak"; changing the three parameters of the method would lead to more (or fewer) peaks being so classified. Data and code are available upon request, and the curious reader can easily try out other parameter combinations.

Of the eleven outbreaks that were detected in the 20th century the five first were the most severe, and they all occurred before 1932. This decrease in severity probably has several causes, one of which is the development of influenza vaccines. The SARS-CoV-2-related outbreak in 2020 seems to be an exception from this trend of decreasing severity. In terms of absolute number of excess deaths this outbreak is the most severe since the Spanish flu in 1918-1919. At the time of writing the daily number of reported deaths in Sweden are still above the predicted number, and it is thus likely that the relative severity of the COVID-19 outbreak will continue to increase.

Acknowledgments

I thank Sveriges Släktforskarförbund for letting me use data from Dödboken and Jenny Öqvist for helpful comments on previous versions of the text.

References

Amanat, F. & Krammer, F. (2020), 'SARS-CoV-2 Vaccines: Status Report', *Immunity* **52**(4), 583–589.

Bell, D., Nicoll, A., Fukuda, K., Horby, P., Monto, A. & Grp, W. H. O. W. (2006), 'Nonpharmaceutical interventions for pandemic influenza, national and community measures', *Emerging Infectious Diseases* **12**(1), 88–94.

Cobey, S. & Hensley, S. E. (2017), 'Immune history and influenza virus susceptibility', *Current Opinion In Virology* **22**, 105–111.

Collins, S. (1932), 'Excess mortality frotn causes other than influenza and pneumonia during influenza epidemics', *Public Health Reports* **47**(46), 2159–2179.

Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., Cheng, Z., Yu, T., Xia, J., Wei, Y., Wu, W., Xie, X., Yin, W., Li, H., Liu, M., Xiao, Y., Gao, H., Guo, L., Xie, J., Wang, G., Jiang, R., Gao, Z., Jin, Q., Wang, J. & Cao, B. (2020), 'Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China', *Lancet* **395**(10223), 497–506.

Peiris, J., Lai, S., Poon, L., Guan, Y., Yam, L., Lim, W., Nicholls, J., Yee, W., Yan, W., Cheung, M., Cheng, V., Chan, K., Tsang, D., Yung, R., Ng, T., Yuen, K. & Grp, S. S. (2003), 'Coronavirus as a possible cause of severe acute respiratory syndrome', *Lancet* **361**(9366), 1319–1325.

Potter, C. (2001), 'A history of influenza', *Journal Of Applied Microbiology* **91**(4), 572–579.

Simonsen, L., Clarke, M., Williamson, G., Stroup, D., Arden, N. & Schonberger, L. (1997), 'The impact of influenza epidemics on mortality: Introducing a severity index', *American Journal Of Public Health* **87**(12), 1944–1950.

Sveriges Släktforskarförbund (2017), *Sveriges dödbok 1860–2017*, Sveriges Släktforskarförbund, Stockholm.