A biometric analysis of infant mortality and temperature, northern Sweden 1895-1950, including the Stensele station and the inland

Göran Broström and Tommy Bengtsson

2022-02-03 23:05:48

Abstract

The effect of extreme temperatures on infant mortality in the Umeå region 1895-1950 is studied in a biometric analysis setting. More precisely, the effect of climate and weather, measured by temperature, on infant mortality is investigated. It turns out that climate is more important than weather, low average temperatures are more important than temporary dips in temperature. The investigated geographical region is situated too far north for bad effects of high temperatures to show.

Contents

1 Introduction

The impact of ambient temperature variations on infant mortality is studied for a northern Sweden coastal area, the Umeå and Skellefteå regions, during the first half of the twentieth century. Two recent papers (Junkka et al., 2021; Karlsson et al., 2021) studied neonatal mortality and temperature variations in a larger geographical area containing the present one during the years 1880–1950. Climate and mortality in general is a research area that has generated great interest over the last years, see Bengtsson and Broström (2010).

The effect of seasonal variation and the occurrence of extreme monthly temperatures is studied and interacted with sex, social class, and legitimacy. Studies are performed separately for neonatal and postneonatal mortality, and for winter and summer seasons, and the classification into endogenous and exogenous factors will be discussed.

One important reason for studying neonatal and postneonatal mortality separately is the empirical findings by Bourgeois-Pichat (Bourgeouis-Pichat, 1951a,b) about endogenous and exogenous mortality and the log-cube transform.

2 Data

We have two sources of data which we combine into one data set suitable for our purpose. The first is demographic data obtained from the *Centre for Demographic and Ageing Research* (CEDAR, https://cedar.umu.se), the second is daily temperature measurements obtained from the *Swedish Meteorological and Hydrological Institute* (SMHI, https://www.smhi.se).

2.1 Infant mortality

Individual data with all births between 1 January 1895 and 31 December 1950 in two coastal areas of north Sweden, Skellefteå (51560 births) and Umeå (31213 births). They were followed until death or age 365 days, whichever came first. The following *static* characteristics were observed on each child:

birthdate Date of birth.

sex Girl or boy.

exit Number of days under observation.

event Logical, *TRUE* if a death is observed.

socBranch Working branch of father (if any).

socStatus Social status of family, based on HISCLASS.

illeg Mother unmarried?

parity Order among siblings.

Some crude statistics about infant, neonatal, and postneonatal mortality are shown in Figures.

Figure 1 shows the average monthly crude infant mortality, and a clear seasonal pattern is visible.

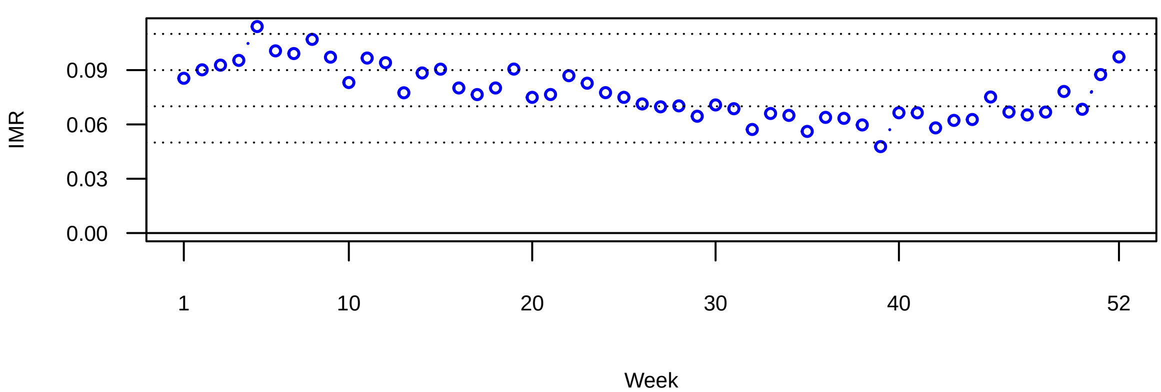

Figure 1: Crude infant mortality by week of year, Umeå/Skellefteå 1895–1950.

The average monthly neonatal mortality is shown in Figure 2.

The seasonal pattern is similar to the one we found above for infant mortality.

The average monthly postneonatal mortality is shown in Figure 3.

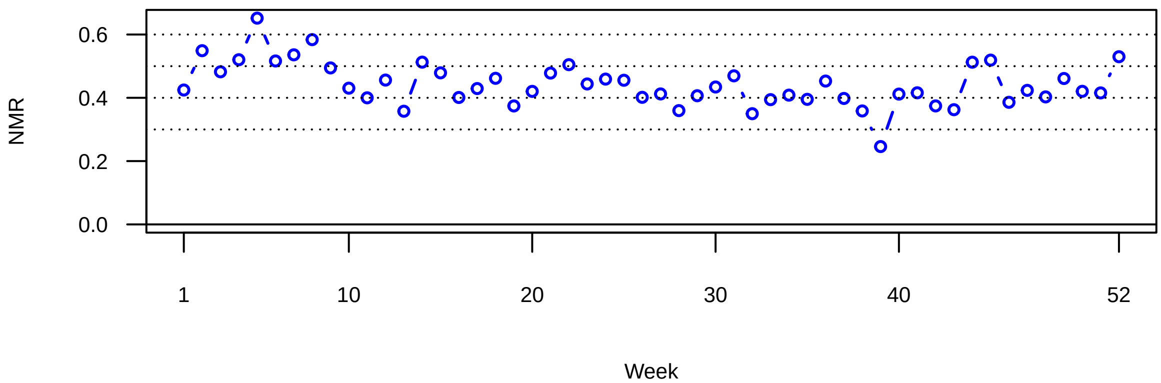

Figure 2: Crude neonatal mortality by week of year, Umeå/Skellefteå 1895–1950.

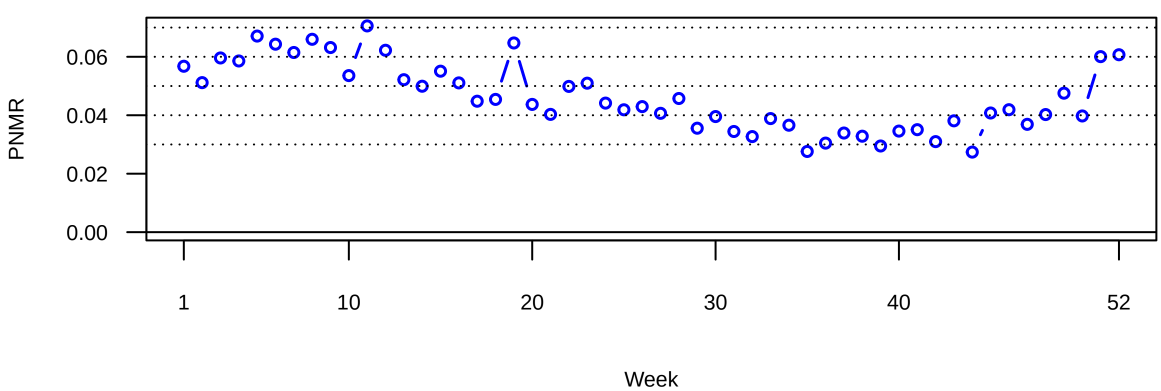

Figure 3: Crude postneonatal mortality by week of year, Umeå/Skellefteå 1895–1950.

The seasonal pattern is once again similar to the one we found for infant mortality. Next, the decline over the years in Figures 4 and 5.

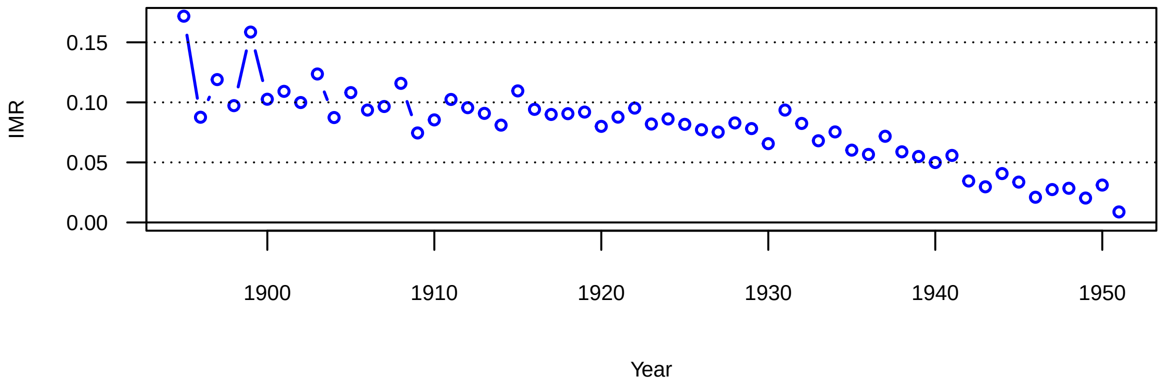

Figure 4: Crude IMR by year, Umeå-Skellefteå 1895–1950.

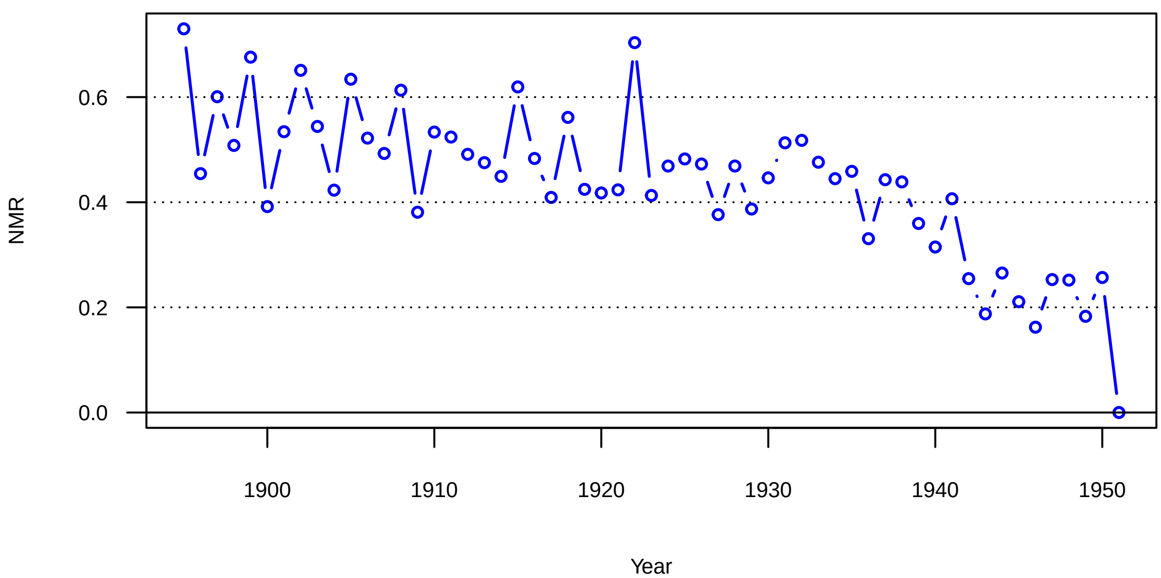

Figure 5: Crude NMR by year, Umeå-Skellefteå 1895–1950.

2.2 Temperature

Temperature data are collected from two weather stations, *Umeå* and *Bjuröklubb* (used with population data from Skellefteå coastal area). Both stations deliver daily temperature data covering our time period, usually three measures per day, morning, noon, and evening. In Table 1, the Umeå data from the week 1–7 January, 1923 is shown.

There are three measurements per day, or 21 per week. In the forthcoming analyses, the weekly data are summarized in a few measurements, see Table 2.

Weekly averages (`mintemp`, `maxtemp`, `meantemp`) are calculated by week and year, and deviations from the averages (`emintemp`, `emaxtemp`, `emeantemp`) of the weekly averages are used as

Table 1: Raw temperature data from first week of 1923, Umeå.

Date	**Time**	**Temperature**	**Quality**
1923-01-01	07:00:00	0.4	G
1923-01-01	13:00:00	0.6	G
1923-01-01	20:00:00	0.0	G
1923-01-02	07:00:00	-1.4	G
1923-01-02	13:00:00	-1.4	G
1923-01-02	20:00:00	-1.2	G
1923-01-03	07:00:00	0.4	G
1923-01-03	13:00:00	0.8	G
1923-01-03	20:00:00	1.2	G
1923-01-04	07:00:00	1.4	G
1923-01-04	13:00:00	1.2	G
1923-01-04	20:00:00	1.0	G
1923-01-05	07:00:00	-1.4	G
1923-01-05	13:00:00	-3.2	G
1923-01-05	20:00:00	-3.4	G
1923-01-06	07:00:00	1.0	G
1923-01-06	13:00:00	0.4	G
1923-01-06	20:00:00	0.4	G
1923-01-07	07:00:00	0.6	G
1923-01-07	13:00:00	0.4	G
1923-01-07	20:00:00	0.4	G

Table 2: Weekly summarized temperature data: Umeå 1923, first week.

week	**year**	**mintemp**	**maxtemp**	**meantemp**	**emintemp**	**emaxtemp**	**emeantemp**
1	1923	-3.4	1.4	-0.1	-17.73	-0.36	-7.54

time-varying *communal covariates*. As an example, see Figure 6, where the variation around the average minimum temperature (`emintemp`) week 1 is shown.

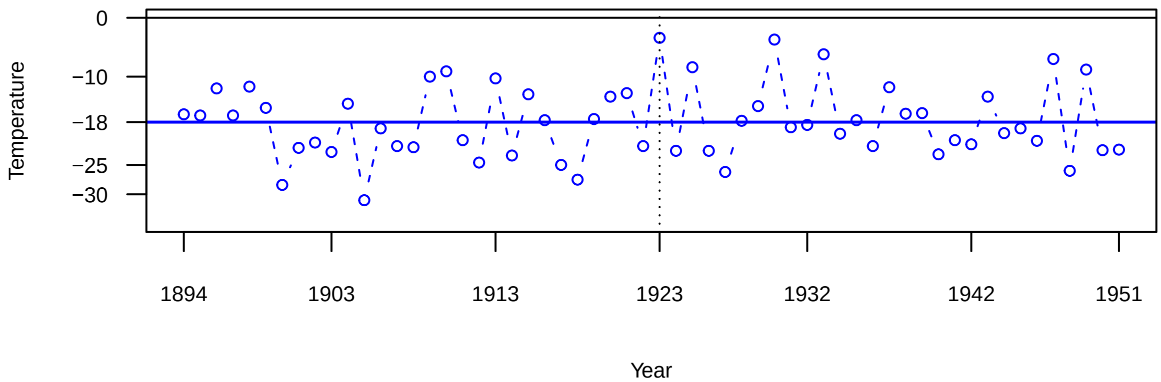

Figure 6: Minimum temperature the first week of each year.

Curiously, our randomly selected year 1923 turns out contain the warmest first week of all years, see Figure 6.

Figure 7 shows the average monthly distribution over all years. The subregional patterns and levels are very similar.

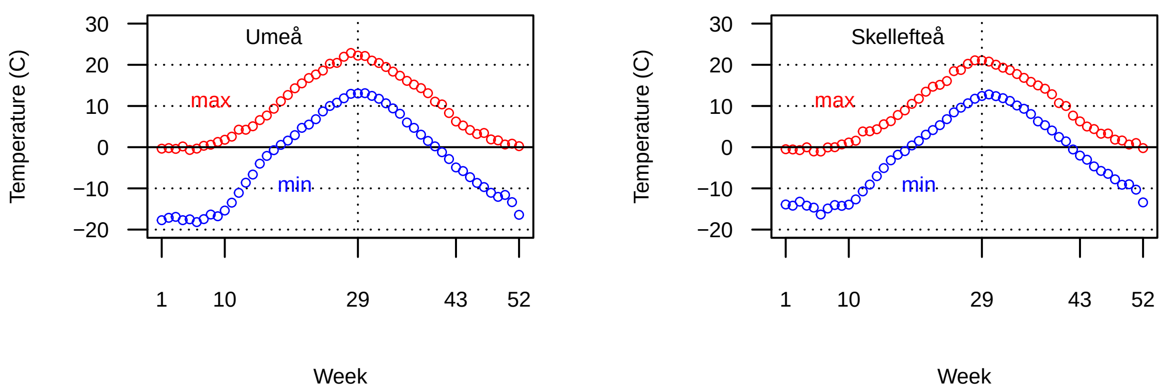

Figure 7: Weekly max and min temperature averages, 1895–1950.

Time trends of yearly average temperatures, see Figure 8.

2.3 Temperature as communal covariates

The two data sets, mortality and weather, are combined into one by treating temperature data as a communal covariate and incorporate it as such in the mortality data set. The function *make.communal* in the **R** (R Core Team, 2021) package *eha* (Broström, 2021a,b) is used for that purpose. Resulting data drame is partly shown in Table 3.

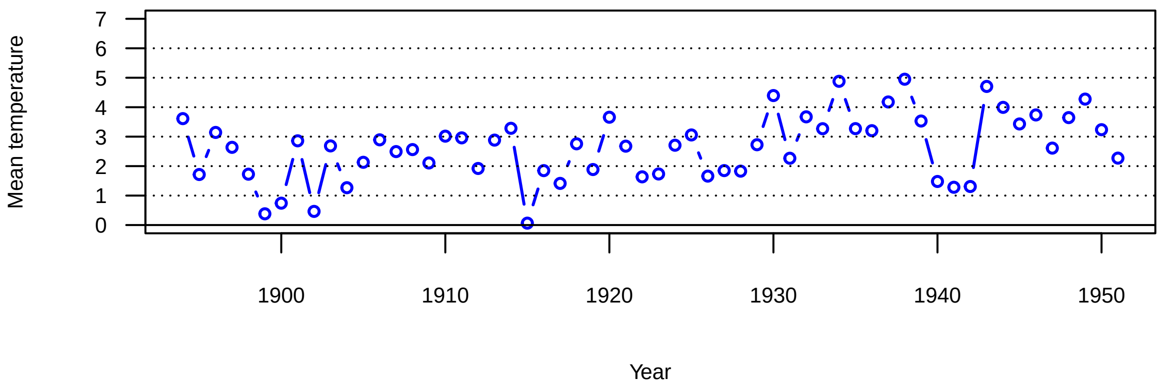

Figure 8: Yearly average temperatures, Umeå and Skellefteå.

Table 3: Data with communal covariates.

enter	exit	event	lowTemp	highTemp	aver	emeantemp	week	year
0.0000000	0.0180327	0	FALSE	FALSE	-1	10.248276	35	1900
0.0180327	0.0372634	0	FALSE	FALSE	-4	8.863793	36	1900
0.0372634	0.0564942	0	FALSE	FALSE	0	7.582759	37	1900
0.0564942	0.0757250	0	FALSE	FALSE	1	6.910345	38	1900
0.0757250	0.0949557	0	FALSE	FALSE	-2	5.315517	39	1900
0.0949557	0.1141865	0	FALSE	FALSE	0	3.610345	40	1900

3 Statistical modelling

The analyses are performed on the *log-cube* scale, following the hints of Bourgeois-Pichat (Bourgeouis-Pichat, 1951a,b), with proportional hazards modelling. Note that the property of proportional hazards are preserved under a strictly monotone increasing time transform. However, the estimates of baseline distribution characteristics will change, of course.

It turns out that extremely low temperature (`lowTemp`) is bad during all seasons except summer, and extremely high temperature (`highTemp`) is bad during summer, but good otherwise. So we group season into two categories, *summer* and *not summer*. In each case separate analyses for neonatal and postneonatal mortality are performed.

4 Results

The results regarding neonatal mortality is much in accordance with the results found by Junkka et al. (2021). However, they used temperature in a "hockey-stick" regression with a breakpoint at 14.5 degrees Celsius and a negative slope (decreasing risk) to the left and a positive slope (increasing risk) to the right. Instead, we are using the average weekly temperature for the 52 weeks of a year, for each week averaging over all the years in the study, as our "reference points" ("climate"), adding deviances up and down ("weather") as "short-term temperature stress". This is similar to the way prices and mortality were related in for instance Bengtsson and Broström (2011), that is, a time series split into long time trend and short term variation.

In accordance with the results of Bourgeouis-Pichat (1951a) and Bourgeouis-Pichat (1951b), we separate the investigation into two parts, *neonatal* and *postneonatal* mortality.

4.1 Neonatal mortality

The analyses are split into two parts by season, *winter + spring* is one, and *summer + fall* the other.

4.1.1 Winter and Spring

This period refers to the months *December to May.* A Cox regression involves as interesting variables *highTemp*, an indicator of temperature at least four degrees above the expected for at least two weeks in a row, *emeantemp* the expected temperature the actual weeek, and *aver* the *excess temperature* the actual week.

Table 4: Neonatal mortality, winter and spring. Adjusted for social branch, sex, illegitimacy, parity, time period, and subregion.

Covariate	**Mean**	**Coef**	**H.R.**	**S.E.**	**L-R p**
highTemp					0.013
FALSE	0.915	0	1	(reference)	
TRUE	0.085	−0.244	0.783	0.101	
aver	−0.013	−0.004	0.996	0.006	0.562
emeantemp	−2.839	−0.002	0.998	0.006	0.795
sex					0.000
boy	0.514	0	1	(reference)	
girl	0.486	−0.260	0.771	0.047	
season					0.041
winter	0.489	0	1	(reference)	
spring	0.511	−0.145	0.865	0.071	
subreg					0.681
ume	0.295	0	1	(reference)	
ske	0.481	0.024	1.025	0.055	
inland	0.224	−0.028	0.972	0.069	
illeg	0.055	0.352	1.422	0.091	0.000
parity					0.000
1	0.257	0	1	(reference)	
2-4	0.454	−0.231	0.794	0.059	
5+	0.288	0.100	1.105	0.063	
socBranch					0.065
office	0.104	0	1	(reference)	
farming	0.470	0.195	1.215	0.087	
worker	0.426	0.133	1.143	0.087	
Events	1896	TTR	4013		
Max. logLik.	−20564				

4.1.2 Summer and Fall

This period refers to the months *April to October*. A Cox regression involves as interesting variables *highTemp*, an indicator of temperature at least four degrees above the expected for at least two weeks in a row, *emeantemp* the expected temperature the actual weeek, and *aver* the *excess temperature* the actual week.

Table 5: Neonatal mortality, summer and fall. Adjusted for social branch, sex, illegitimacy, parity, time period, and subregion.

Covariate	**Mean**	**Coef**	**H.R.**	**S.E.**	**L-R p**
aver	0.004	−0.021	0.980	0.010	0.032
emeantemp	8.803	−0.012	0.988	0.007	0.090
sex					0.000
boy	0.512	0	1	(reference)	
girl	0.488	−0.175	0.840	0.050	
season					0.092
summer	0.502	0	1	(reference)	
fall	0.498	−0.152	0.859	0.090	
subreg					0.218
ume	0.295	0	1	(reference)	
ske	0.492	0.101	1.106	0.060	
inland	0.213	0.039	1.040	0.075	
illeg	0.054	0.538	1.712	0.094	0.000
parity					0.000
1	0.272	0	1	(reference)	
2-4	0.452	−0.170	0.844	0.064	
5+	0.276	0.101	1.106	0.068	
socBranch					0.003
office	0.108	0	1	(reference)	
farming	0.460	0.318	1.374	0.099	
worker	0.432	0.233	1.262	0.098	
Events	1599	TTR	3911		
Max. logLik.	−17292				

4.2 Postneonatal mortality

This is the simplest part, because Bourgeouis-Pichat (1951a) predicts that the baseline distribution (given the log-cube transformation **g**) is *exponential*, that is, the hazard function is *constant*. To give substance to this claim, a short introduction to the ideas of Bourgeois-Pichat is given.

4.2.1 Bourgeois-Pichat and the log-cube transform

The *log-cube transform* g, suggested by Bourgeouis-Pichat (1951a), is defined as

$$g(t) = \log^3(1 + t), \quad 0 \le t \le 365, \tag{1}$$

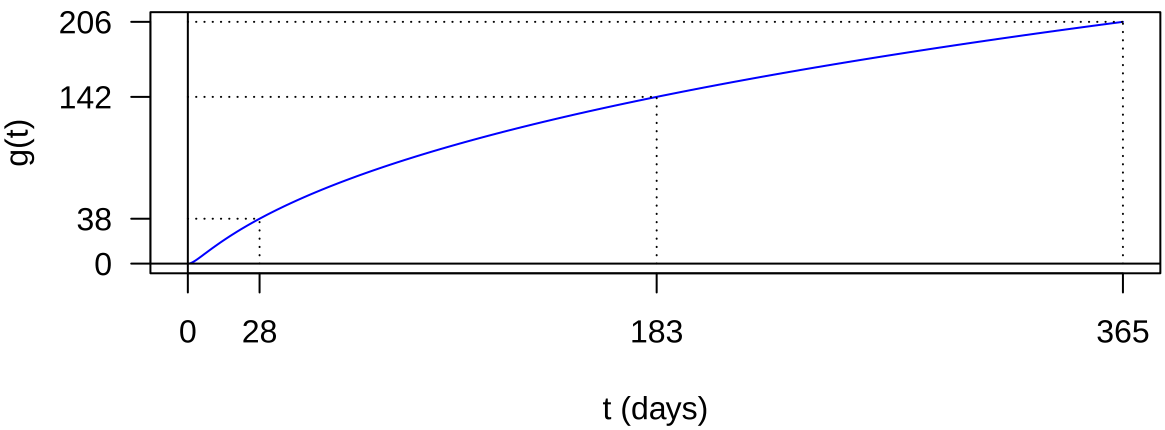

Figure 9: The g transform.

where t is age in days. A graph of the transform is shown in Figure 9.

It was used as a tool for dividing infant mortality into *endogenous* and *exogenous* causes, and a an important part in that venue was the observations (i) postneonatal mortality is purely exogenous, and (ii) the distribution of exogenous mortality is *truncated uniform* (right truncated at $t = g(365)$) on the g-transformed time scale.

Today, assumption (ii) is slightly outdated, for three reasons: (a) The uniform distribution is not a very practical tool in survival analysis, (b) replacing the uniform with an *exponential* not only fits better (usually), it also is a survival distribution that is extremely easy to work with in survival analysis, and (c) "all models are wrong, but some are useful (Box, 1976)".

For now, let us look at the basic distribution of postneonatal life on the g time scale. We do it nonparametrically by calculating the *Nelson-Aalen* estimator of the cumulative hazards function. See Figure 10, where the two curves are indistinguishable.

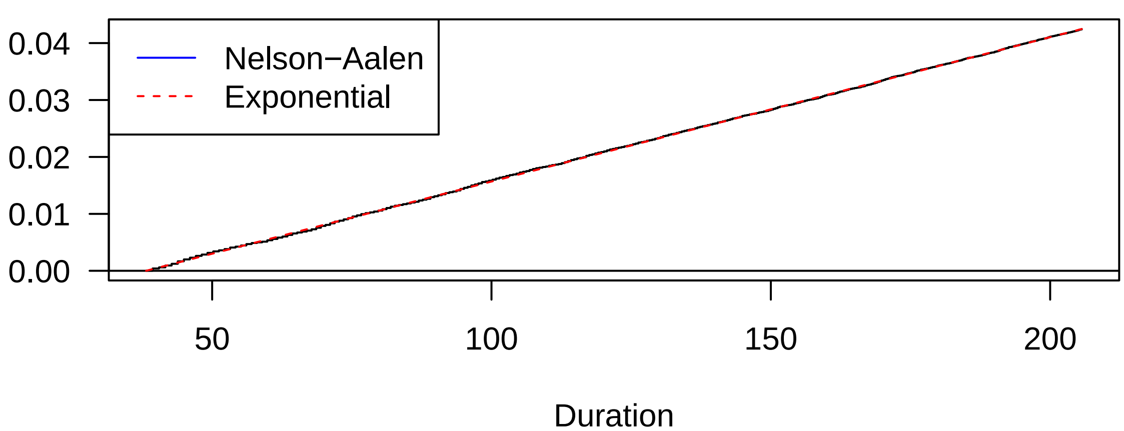

Figure 10: Nelson-Aalen estimator and exponential cumulative hazards of postneonatal survival on the g scale.

Table 6: Postneonatal mortality, winter and spring. Adjusted for social status and branch, sex, illegitimacy, parity, time period, and subregion.

Covariate	**Mean**	**Coef**	**H.R.**	**S.E.**	**L-R p**
lowTemp					0.026
FALSE	0.875	0	1	(reference)	
TRUE	0.125	0.146	1.157	0.065	
highTemp					0.020
FALSE	0.913	0	1	(reference)	
TRUE	0.087	−0.200	0.819	0.087	
aver	0.004	−0.010	0.991	0.006	0.114
emeantemp	−2.932	−0.016	0.984	0.004	0.000
sex					0.000
boy	0.512	0	1	(reference)	
girl	0.488	−0.228	0.796	0.040	
socBranch					0.000
office	0.107	0	1	(reference)	
farming	0.465	0.651	1.918	0.092	
worker	0.428	0.544	1.722	0.084	
socStatus					0.000
high	0.587	0	1	(reference)	
low	0.388	0.199	1.220	0.053	
unknown	0.025	−0.301	0.740	0.155	
illeg	0.053	0.583	1.792	0.082	0.000
parity					0.000
1	0.265	0	1	(reference)	
2-4	0.455	0.192	1.211	0.055	
5+	0.280	0.609	1.838	0.057	
subreg					0.000
ume	0.294	0	1	(reference)	
ske	0.488	−0.322	0.725	0.046	
inland	0.218	−0.274	0.761	0.057	
Events	2549	TTR	8351030		
Max. logLik.	−23001				

4.2.2 Winter and Spring

The result in Table 6 shows that *climate* (`emintemp`) is more important than *weather* (`excessTemp`). Moreover, no signs of interaction between weather or climate and the rest of covariates (not shown).

4.2.3 Summer and Fall

The result in Table 7 shows that *climate* (`emintemp`) is more important than *weather* (`excessTemp`). Moreover, no signs of interaction between weather or climate and the rest of covariates (not shown).

Table 7: Postneonatal mortality, winter and spring. Adjusted for social status and branch, sex, illegitimacy, parity, time period, and subregion.

Covariate	**Mean**	**Coef**	**H.R.**	**S.E.**	**L-R p**
lowTemp					0.008
FALSE	0.975	0	1	(reference)	
TRUE	0.025	0.394	1.482	0.143	
highTemp					0.112
FALSE	0.982	0	1	(reference)	
TRUE	0.018	0.283	1.327	0.172	
aver	0.011	−0.010	0.990	0.010	0.329
emeantemp	8.654	0.005	1.005	0.004	0.178
sex					0.000
boy	0.511	0	1	(reference)	
girl	0.489	−0.176	0.838	0.049	
socBranch					0.000
office	0.107	0	1	(reference)	
farming	0.463	0.722	2.059	0.112	
worker	0.430	0.587	1.799	0.102	
socStatus					0.000
high	0.586	0	1	(reference)	
low	0.389	0.339	1.404	0.065	
unknown	0.025	0.269	1.309	0.155	
illeg	0.054	0.566	1.762	0.098	0.000
parity					0.000
1	0.266	0	1	(reference)	
2-4	0.454	0.277	1.319	0.068	
5+	0.280	0.672	1.957	0.071	
subreg					0.000
ume	0.294	0	1	(reference)	
ske	0.486	−0.419	0.658	0.055	
inland	0.220	−0.507	0.602	0.071	
Events	1711	TTR	8445678		
Max. logLik.	−16127				

5 The biometric analysis of infant mortality

A modern version of the Bourgeois-Pichat biometric modelling (Bourgeouis-Pichat, 1951a,b) is given here, *without* the “log-cube transform”.

6 Conclusion

Remains to be written.

References

Bengtsson, T. and Broström, G. (2010). Mortality crisis in rural southern Sweden 1766–1860. In Kurosu, T., Bengtsson, T., and Campbell, C., editors, *Demographic Response to Economic and Environmental Crisis*, pages 1–16. Reitaku University Press, Kashiwa.

Bengtsson, T. and Broström, G. (2011). Famines and mortality crises in 18th to 19th century southern Sweden. *Genus*, 67:119–139.

Bourgeouis-Pichat, J. (1951a). La mesure de la mortalité infantile. I. Principes et métodes. *Population (French Edition)*, 6:233–248.

Bourgeouis-Pichat, J. (1951b). La mesure de la mortalité infantile. II. Les causes de décès. *Population (French Edition)*, 6:459–480.

Box, G. E. P. (1976). Science and statistics. *Journal of the American Statistical Association*, pages 791–799.

Broström, G. (2021a). *eha: Event History Analysis*. R package version 2.9.0. https://CRAN.R-project.org/package=eha.

Broström, G. (2021b). *Event History Analysis with R, Second Edition*. Chapman & Hall/CRC, Boca Raton.

Junkka, J., Karlsson, L., Lundevaller, E., and Schumann, B. (2021). Climate vulnerability of Swedish newborns: Gender differences and time trends of temperature-related neonatal mortality, 1880–1950. *Environmental Research*, 192. article id 110400.

Karlsson, L., Junkka, J., Schumann, B., and Häggström Lundevaller, E. (2021). Socioeconomic disparities in climate vulnerability: neonatal mortality in northern Sweden, 1880–1950. *Population and Environment*, 10.1007.

R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org.