Accounting for temporal variation in morbidity measurement and projections

Alyson van Raalte*[1] and Tim Riffe[1]

[1]Max Planck Institute for Demographic Research

September 25, 2015

Abstract

This is important stuff!

G is a bad health condition that varies as a function of time to death, y and not as a function of chronological age, a. However, there will still be an apparent age function, $g'(a)$, given that $g(y)$ is regular and mortality is sort of stable, but not really. $g'(a)$, in this case, is an aggregate based on both mortality and the real underlying time-to-death process:

$$g'(a) = \frac{\int_0^\omega g(y)N(a,y)\,\mathrm{d}y}{N(a)} \tag{1}$$

$$= \frac{\int_0^\omega g(y)N(a)\mu(a+y)\frac{\ell(a+y)}{\ell(a)}\,\mathrm{d}y}{N(a)} \tag{2}$$

$$= \int_0^\omega g(y)f(y|a)\,\mathrm{d}y \tag{3}$$

a little exercise we still need to do: find the $g'(a)$ that belongs to $g(y)$ in our canned example. It will be different for males and females because they have different mortality schedules. In this case, we can make the healthy life expectancy function be based on mortality and $g'(a)$ and see what would be the prediction if $g'(a)$ is held constant and we induce mortality improvement. The answer is that mortality improvement will appear to increase the proportion of remaining life expectancy that is unhealthy: also the absolute years spent unhealthy, but the change in sex gap is maybe ambiguous (gotta check, maybe not), depending on the changes induced.

brief interlude .... This is a simple caes of $g(y)$, but in reality morbidity often varies as a function of both chronological and thanatological age, and we ought to have a function $g(a, y)$.

*vanraalte@demogr.mpg.de

1 temp section, out of sync, just for latex

Given the numbers from Figures X, there are various methods that one can use to calculate period and cohort lifetables. For the sake of reproducibility for our toy example, we describe steps as follows, firs tfor periods, then for cohorts.

1.1 Period quantities

We use event exposure lifetables, though it would be possible to jump straight to death quotients from the given Lexis diagram. Exposures for age x in year t, $E(x,t)$, are calculated, per the HMD Methods Protocol (cite) as:

$$E(x,t) = \frac{C(x,t) + C(x,t+1)}{2} + D_L(x,t) - D_U(x,t)6 \qquad , \tag{4}$$

where $C(x,t)$ is the census population in age interval $[x, x+1)]$ on January 1 of year t, $D_L(x,t)$ are deaths in the lower Lexis triangle of age x in year t, i.e., belonging to the cohort born in the year interval $[t-x, t-x+1)$. $D_U(x,t)$ are deaths in the upper Lexis triangle of age x in year t, i.e., belonging to the cohort born in the year interval $[t-x-1, t-x)$. All standard period lifetable steps are followed from the HMD Methods Protocol, with the exception of the $a(x)$ assumption. The HMD assumes period $a(x)$ values of $\frac{1}{2}$. Instead, we apply the following formula:

$$a(x,t) = \frac{D_L(x,t)\frac{1}{3} + D_U(x,t)\frac{2}{3}}{D_L(x,t) + D_U(x,t)} \qquad . \tag{5}$$

We then proceed to calculate all columns through $e(x)$.

The average value of the unhealthy condition G at age $x$in year t, $g(x,t)$ is calculated as follows. We first convert counts unhealthy on birthdays to proportions, and then take the arithmetic average of the proportion unhealthy at age x and age $x+1$. Expectancies are then calculated as follows:

$$e(0,t) = \sum_{0}^{2} L(x,t) \tag{6}$$

$$e_U(0,t) = \sum_{0}^{2} L(x,t)g(x,t) \tag{7}$$

$$e_H(0,t) = e(0,t) - e_U(0,t) \qquad , \tag{8}$$

where $e(0,t)$ is the life expectancy at birth in year t, $e_U(0,t)$ is unhealthy life expectancy, and $e_H(0,t)$ is healthy life expectancy.

1.2 Cohort quantities

For cohorts we proceed directly from within-cohort age interval survival probabilities, $p(x, c)$, as follows:

$$p(x, c) = \frac{B(x+1, c)}{B(x, c)} \quad , \tag{9}$$

where B are birthdays (horizontal counts), x indexes the lower age interval bound, and c is the cohort born in the year interval $[c, c+1)]$. Starting with a radix of 1, we calculate $l(x, c)$ as:

$$l(x, c) = \prod_{0}^{x} p(a, c) \tag{10}$$

$L(x, c)$ is calculated as the arithmetic average of lower $l(x, c)$ and $l(x+1, c)$. The average value of G for the AC parallelogram is calculated similarly as for period squares, except that we take the arithmetic average of the proportions unhealthy at birthday x and $x + a$ within the same cohort. Expectancies are then calculated in the same way.

1.3 Values assumed for $g(y)$ in toy example

If deaths in Lexis triangles are distributed uniformly, then the average year lived of those dying in the triangle is 1/3. Since deaths in the example Lexis diagram are given in triangles, we assume values of $g(y)$ for the average years lived for the deaths in each triangle. For example, 20 of the babies born in the year 2003 will die within the year, and these have an average lifespan of 1/3, and so they contribute 20×0.90 unhealthy people to the birth cohort at age 0.

TTD	$g(y)$
1/3	0.90
2/3	0.60
1 1/3	0.20
1 2/3	0.10
2 1/3	0.05
2 2/3	0.02

Table 1: $g(y)$ used in toy example.

2 Morbidity as a function of thanatological age

Imagine a bad health condition, G, that varies as a function of time to death, y, and not as a function of chronological age, a. Since the distribution of

times to death is empirically regular, there will still be an apparent age function, $g^\star(a)$. In this case $g^\star(a)$ is an aggregate based on both mortality and the real underlying time-to-death process:

$$g^\star(a) = \frac{\int_0^\omega g(y)N(a,y)\,\mathrm{d}y}{N(a)} \tag{11}$$

$$= \frac{\int_0^\omega g(y)N(a)\mu(a+y)\frac{\ell(a+y)}{\ell(a)}\,\mathrm{d}y}{N(a)} \tag{12}$$

$$= \int_0^\omega g(y)f(y|a)\,\mathrm{d}y \qquad , \tag{13}$$

where $N(a)$ is the population aged a, $\ell(a)$ is the survival function, and $\mu(a)$ is the force of mortality. $f(y|a)$ can be interpreted as the probability of dying in y years given survival to age a, and thus the final expression is purged of population structure. That is, the population of age a that has condition G does not depend on population structure at all, but only on the force of mortality and the time-to-death pattern of G, $g(y)$.

A function such as $g(y)$ would have implications for the interpretation of period age patterns of morbidity, and by extension, healthy life expectancy (HLE). If a function such as $g(y)$ holds, it is tautologically true that the measurement of HLE in completed cohorts (or populations with fixed mortality) will be identical whether estimated on the basis of $g^\star(a)$ or $g(y)$. Period HLE is also unproblematic in a fixed mortality setting, as long as future mortality is also held fixed. Difficulties only arise in the interpretation of period HLE under changing mortality. For a time-to-death process, forthcoming improvements in mortality will have the effect of decreasing morbidity today. This artifact can stymie the projection of today's age patterns of morbidity into the future: Under changing mortality and a fixed $g(y)$, the age pattern of morbidity will change even as the morbidity process does not.

Further, since morbidity is partly a function of mortality, it is difficult to compare the age patterns of morbidity for populations with different mortality levels or patterns. Under these circumstance, it is trickier than it seems to partition differences period HLE into true morbidity and mortality components, because the morbidity component will rely on some unknown future mortality quantity that is a driver of the apparent age pattern of morbidity.

We first illustrate this concept with a toy example. Assume the population process outlined in the Lexis diagram of Figure 1. This diagram represents quantities that will induce the aforementioned inconsistency. Blue numbers on the horizontal age bars represent the number of birthdays, while black numbers represent census counts. Red numbers represent deaths in the Lexis triangle, and finally, green numbers next to birthdays represent

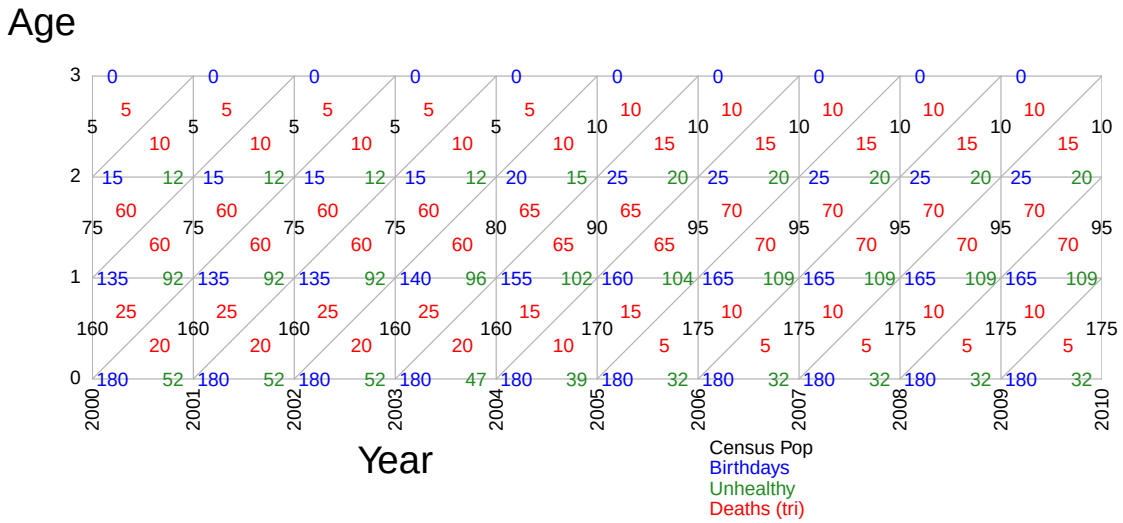

Figure 1: The Lexis diagram

the number of unhealthy people that would be alive assuming the pattern of $g(y)$ found in Figure 4 (given in Appendix Table 1) and the given mortality pattern, assuming deaths are distributed uniformly in Lexis triangles, and the average values given in Table 1 apply. The example moves from a higher mortality stationary setting to a lower mortality stationary setting.

To calculate the numbers unhealthy in any given birthday line, first note that the average years lived in a Lexis triangle by those dying in the triangle is 1/3. For example, 20 of the babies born in the year 2003 will die within the year. These deaths have an average lifespan of 1/3, and so they contribute 20×0.90 unhealthy people to the birth cohort at age 0 (value from appendix table 1. Those dying from the same cohort before age 1 in the year 2004 contribute 15×0.6 unhealthy persons to age 0 in 2003, and so forth iteratively up the cohort.

In this controlled setting, where the underlying $g(y)$ is held fixed, we can calculate both the period and cohort HLE values.[1] Figure 3 shows trends in period and cohort total and healthy life expectancy. In the initial and final stages, period and cohort healthy and total life expectancies agree, because these are the stationary start and end of the series. As is usually the case, there is no perfect way to compare cohort and period line graphs, since the x-axis refers to both cohort and period.

The gap between year 2001 and 2005 marks an inconsistency. In this case, the inconsistency is entirely driven by mortality change. Further the

[1]See Appendix explanations for the details used to calculate lifetables and mean values of $G(a)$.

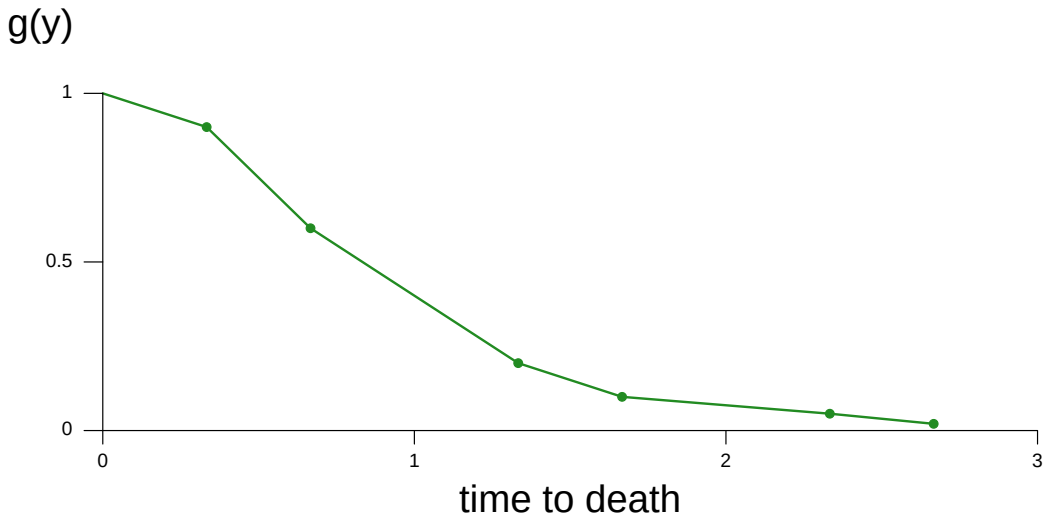

Figure 2: The time-to-death profile of disability

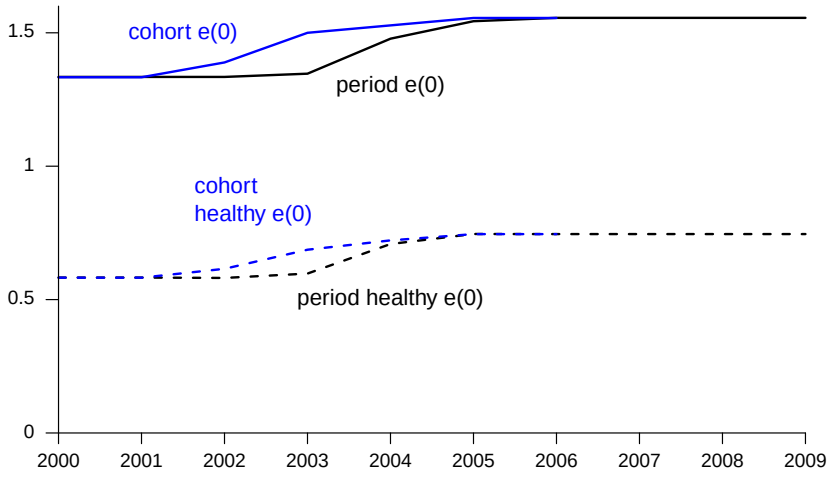

Figure 3: Cohort and period life expectancy and healthy life expectancy.

gap in HLE is due not only to the change in the survival function used to calculate HLE, but also to changes in the pattern of $g^{\star}(a)$. That is, if we were to decompose the difference in 2001 and 2007 HLE using conventional methods, there would appear to be a morbidity contribution to the difference, even though morbidity in this case is held fixed. We demonstrate such an inconsistency using somewhat more realistic data in a following section.

In reality, not all end-of-life health conditions are exclusive functions of time-to-death, but morbidity often varies as a function of both chronological and thanatological age, and it is best to express morbidity as a function of both age and time-to-death, $g(a, y)$. There is great variety in the temporal variation of late-life health conditions (cite Riffe 2015 MPIDR working paper.) Further, the function $g(a, y)$ changes over time and it is not fixed as in our example. The distortion demonstrated is however likely to arise in everyday practice when comparing health trends over age between populations and over time, since many health conditions appear to show strong time-to-death components. Trends in mortality may offset or amplify changes in morbidity. Therefore, in order to separate effects, more careful measurements are required than is typically the case. We provide another expository example in a more realistic setup.