Max-Planck-Institut für demografische Forschung

Max Planck Institute for Demographic Research

The decomposition of the difference between two healthy life expectancies. Which formula is right?

Vladimir M. Shkolnikov I shkolnikov@demogr.mpg.de
Evgeny M. Andreev

The decomposition of the difference between two healthy life expectancies. Which formula is right?

Vladimir M. Shkolnikov and Evgeny M. Andreev

Background

About two months ago, a colleague asked me to look at a preliminary version of a working paper written by him together with two other researchers. The topic of the paper was how patterns of morbidity prevalence influence the length of healthy life. Among other analyses, the authors presented an age decomposition of a change in healthy life expectancy (HLE). The text of the working paper included references to two studies published in the early 2000s, which were the first to introduce formulas for decomposing the difference between two HLE values by age, and by health and mortality within each age. The (chronologically) first study (Andreev, Shkolnikov, Begun 2002) deduced new formulas for HLE decompositions by applying a universal decomposition method called the "general algorithm of stepwise replacement" to health expectancies. The focus of the second study (Nusselder and Looman 2004) was entirely on the decomposition of HLE. In line with the first study and independently of it, the authors of this study presented a new formula for the same decomposition by age, health, and mortality. In addition, they showed how the respective components can be further split by medical causes.

The colleague who had sent me the working paper mistakenly thought that the two studies of 2002 and 2004 provided variants of one and the same decomposition method, and that these variants returned approximately the same results. However, after taking a closer look, we realized that there was an *important difference between the two decomposition methods* that leads to substantial differences in their quantitative results.

In this note, we intend to clarify this apparent discrepancy by demonstrating 1) the correctness of the method by Andreev, Shkolnikov, and Begun 2002; and 2) the flaws in the method related to mortality components by Nusselder and Looman 2004.

Sullivan's HLE

The index we will decompose is Sullivan's HLE. This is the most widely used measure of healthy life expectancy (also called "life expectancy free of disease/disability"). Using the standard life table notation, Sullivan's HLE is:

$$h_a = \overset{w}{\underset{x=a}{\overset{\circ}{a}}} \left({}_nL_x \times {}_np_x \right). \tag{1}$$

This familiar formula, which is based on the "observed prevalence life table" approach, is used to determine the expected length of healthy life by combining the life table years lived within elementary age intervals [x, $x+n$) and the prevalence of the healthy state based on

1

self-reported information on ill health, (or disability or physical performance) (Sullivan 1964; Robine, Romieu, Cambois 1999). For the sake of simplicity, in the rest of this paper we assume that a=0, n=1, and l_0 = 1. For the same reason, we will skip the left index for variables defined on the elementary age intervals [x, x+1).

The decomposition task is to split the difference in the HLE values between two populations or the temporal change in one population according to the contributions of age-specific mortality and health. Formally, the difference

$$\Delta h_0 = \sum_{x=0}^{w} L_x^2 p_x^2 - \sum_{x=0}^{w} L_x^1 p_x^1 \qquad (2)$$

should be presented as

$$\Delta h_0 = \sum_{x=0}^{w} (MRT_x + HEL_x), \qquad (3)$$

where MRT_x and HEL_x are the contributions to the total difference Δh_0 produced by the differences in the mortality rates, and in the prevalence values within an elementary age interval [x, x+1), respectively.

Decomposition by Andreev, Shkolnikov, and Begun (2002)

The stepwise replacement algorithm was introduced in a *Demographic Research* paper by Andreev, Shkolnikov, and Begun (2002). A brief and easy-to-understand summary of the method can also be found in Online Resource 1 of the recent article by Jdanov, Shkolnikov, van Raalte, and Andreev (2017). The method expresses the total change in the dependent index as the sum of the effects of the sequential replacement of age-specific rates progressing from age zero to the highest age. This order of replacement of the age-specific death rates guarantees that in the case of life expectancy, the resulting age components are exactly the same as those calculated according to the earlier (and the most commonly used) decomposition formula by Arriaga-Andreev-Pressat (Arriaga 1984; Andreev 1982; Pressat 1985).

According to the algorithm, the mortality component MRT_x produced by age [x, x+1) is the HLE change resulting from the replacement of the death rate M_x^1 by the rate M_x^2 [1] in the vector of death rates M consisting of the elements M_y^2 at ages 0≤y≤x-1 and elements M_y^1 at ages y≥x. The health component HEL_x is the HLE change resulting from the similar replacement of the healthy state prevalence p_x^1 by the healthy state prevalence p_x^2 in the similarly composed vector of prevalences Π. The derivation of the formal expressions on pp. 504-505 of the paper by Andreev, Shkolnikov and Begun (2002) benefited from the following presentation of h_0:

---

[1] More precisely, the values of the MRT_x and HEL_x components were obtained by averaging the HLE changes resulting from the replacements $M_x^2 \circledR M_x^1$ and $M_x^1 \circledR M_x^2$.

$$h_0 = \sum_{y=0}^{x-1} L_y p_y + P_x p_x l_x + h_{x+1} l_{x+1} ,$$ (4)

where $P_x = \dfrac{L_x}{l_x}$.

It is worth noting that in the latter equation, the first term represents a component of HLE that depends on mortality and health at ages younger than x; the second term depends on mortality and health within the age interval $[x, x+1)$ and on mortality at ages under x; and the third term depends on mortality and health at ages $x+1$ and older and on mortality at ages under $x+1$. In order to derive formulas for the health and the mortality components produced by the health and the mortality differences within the age interval $[x, x+1)$, the stepwise replacement algorithm was applied to the right-hand side expression in (4).

The final formulas for the age-specific mortality and health components of $\mathrm{D}h_0$ are:

$$MRT_x = \frac{1}{4}(l_x^1 + l_x^2)(P_x^2 - P_x^1)(p_x^1 + p_x^2) + \frac{1}{2}(h_{x+1}^1 l_x^2 + h_{x+1}^2 l_x^1)(q_x^1 - q_x^2) ,$$ (5)

$$HEL_x = \frac{1}{4}(l_x^1 + l_x^2)(P_x^1 + P_x^2)(p_x^2 - p_x^1) .$$ (6)

If all individuals in populations 1 and 2 are healthy and $p_x^1 = p_x^2 = 1$ at every age x, then healthy life expectancies h_0^1 and h_0^2 are equal to conventional life expectancies e_0^1 and e_0^2. It is easy to see that in such a case, the health components (6) are equal to zero, and the mortality components (5) are equal to the conventional age-specific components in the conventional decomposition by Andreev (1982), Arriaga (1984), and Pressat (1985).

Decomposition by Nusselder and Looman (2004)

In the work by Nusselder and Looman (2004), the derivation of the formulas for the components MRT_x and HEL_x is more straightforward. The right-hand side of equation (2) was considered as the sum (over ages) of age-specific differences $L_x^2 p_x^2 - L_x^1 p_x^1$. Similarly to the classic Kitagawa's decomposition (Kitagawa 1955), this difference was presented as:

$$L_x^2 p_x^2 - L_x^1 p_x^1 = \left(L_x^2 - L_x^1\right) \times \left(\frac{p_x^1 + p_x^2}{2}\right) + \left(p_x^2 - p_x^1\right) \times \left(\frac{L_x^1 + L_x^2}{2}\right)$$ (7)

Nusselder and Looman then considered the first additive term on the right-hand side of equation (7) as the mortality component, and the second additive term as the health component corresponding to the mortality and the health differences between the two populations :

$$MRT_x = \frac{1}{2}(\pi_x^1 + \pi_x^2)(L_x^2 - L_x^1),$$ (8)

$$HEL_x = \frac{1}{2}(L_x^1 + L_x^2)(\pi_x^2 - \pi_x^1).$$ (9)

We can, however, see that if all of the individuals in populations 1 and 2 are healthy ($\pi_x^1 = \pi_x^2 = 1$ at every age x), the mortality components (8) do not correspond to the conventional age components of the life expectancy decomposition.

Reasons for the differences between the ASB and the NL decompositions

After comparing the mortality components in equations (5) and (8), we can detect one main difference between them. While the former equation depends only on mortality between age x and age $x+1$, the latter equation depends on mortality at age x and at all younger ages (since L_x depends on l_x). This implies that mortality component (8) is not a contribution of the difference $M_x^2 - M_x^1$ to the total difference Δh_0, as it should be by definition. This components depends also on the differences between death rates at ages 0 to x-1.

Health components (6) and (8) are much more similar. Health component (6) is slightly more precise because it takes into account curvilinear survival between ages x and $x+1$. However, when (6) is presented as:

$$HEL_x = \frac{1}{4}(L_x^1 + L_x^2)(\pi_x^2 - \pi_x^1) + \frac{1}{4}(\frac{l_x^2}{l_x^1}L_x^1 + \frac{l_x^1}{l_x^2}L_x^2)(\pi_x^2 - \pi_x^1),$$ (6a)

it becomes clear that in most cases, the difference between (6) and (8) is small.

Let us consider an empirical example of the decomposition of the difference in healthy life expectancy among men aged 30 in France and in Poland in 2011. For our calculations, we use the age-specific mortality rates from the *HMD* (Human Mortality Database 2017) and the age-specific prevalence among individuals without chronic diseases at ages 30-34, 35-39, ..., 85+ from the SILK surveys (Statistics on Income and Living Conditions), which are available on the *EuroHEx* website (Advanced Research on the European Health Expectancies 2017).

Calculations of life expectancy free of chronic disease at age 30 for the two male populations are given in the enclosed Excel file. In 2011, the e_{30} values for these men were 49.4 years in France and 43.7 years Poland; a gap of 5.7 years. In the same year, life expectancy free of chronic disease was 28.4 years in France and 26.2 years in Poland; a difference of 2.2 years. This suggests that at age 30, French men could expect to live 5.7 years longer than their Polish counterparts, but that they would spend more than half of these additional time with chronic diseases. Table 1 compares the mortality and health components of the health expectancy gap of 2.2 years between the ASB and the NL decompositions. Calculation of these components is given in the enclosed Excel file.

Table 1. Two decompositions of the difference in the length of life free of chronic disease among men aged 30 and older in France and in Poland in 2011.

	Health components		Mortality components	
	ASB[1]	NL	ASB	NL
30-34	-0.254	-0.254	0.070	0.005
35-39	-0.302	-0.302	0.101	0.020
40-44	-0.269	-0.269	0.141	0.043
45-49	-0.072	-0.072	0.173	0.075
50-54	0.028	0.028	0.208	0.113
55-59	0.076	0.076	0.235	0.150
60-64	0.261	0.262	0.259	0.205
65-69	0.079	0.079	0.258	0.258
70-74	0.426	0.427	0.213	0.285
75-79	0.060	0.060	0.158	0.288
80-84	0.054	0.054	0.098	0.246
85+	0.077	0.078	0.057	0.280
Total	0.164	0.167	1.970	1.967

[1] ASB denotes the Andreev-Shkolnikov-Begun decomposition.
NL denotes the Nusselder-Looman decomposition.

Looking at this table, we can see that the totals (with respect to age) of health and mortality contributions in the two decompositions are very similar. At the same time, the age-specific mortality components of the two decompositions differ greatly (Figure 1). Indeed, the NL components increase with age and form plateau (with some fluctuations) at the oldest ages (75 and older). The ASB components increase up to ages 60-69, and then steeply decrease with age, reaching their lowest point at age 85 and older.
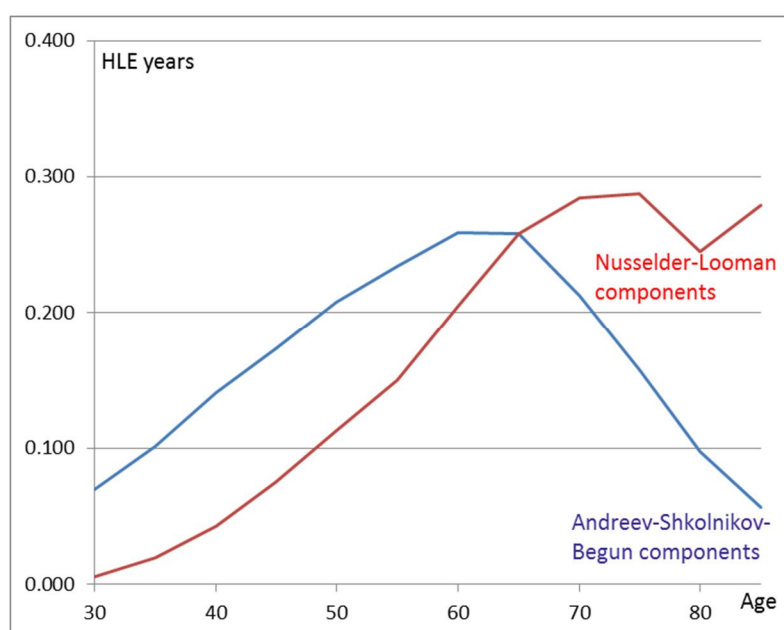

Figure 1. Mortality components of the gap in the expected length of life free of chronic disease between 30-year-men in France and in Poland in 2011.

Conclusion

The age decomposition formulas by Andreev, Shkolnikov, and Begun are correct for both the health and the mortality components. The age decomposition formulas by Nusselder and Looman are correct for the health components, but are incorrect for the mortality components.

We hope that this working paper provides a helpful clarification for those interested in the decomposition of differences or changes in healthy life expectancy.

References

Advanced Research on the European Health Expectancies 2017. Available at http://www.eurohex.eu/index.php?option=ehleisproject . Last visited on 21 August 2017.

Andreev, E.M. 1982. *Metod komponent v analize prodoljitelnosty zjizni.* [The method of components in the analysis of length of life]. *Vestnik Statistiki*, vol. 9, pp. 42-47.

Andreev, E.M., Shkolnikov, V.M., Begun, A.Z. 2002. Algorithm for decomposition of differences between aggregate demographic measures and its application to life expectancies, healthy life expectancies, parity-progression ratios and total fertility rates. *Demographic Research*, vol. 7, article 14, pp. 499-521, www.demographic-research.org/

Arriaga, E. 1984. Measuring and explaining the change in life expectancies. *Demography* vol. 21(1), pp. 83-96.

Human Mortality Database 2017. Available at www.mortality.org . Last visited on 21 August 2017.

Jdanov, D.A., Shkolnikov, V.M., van Raalte, A., Andreev, E.M. 2017. Decomposing current mortality differences into initial differences and differences in trends: the contour decomposition method. *Demography* 54: 1579-1602

Kitagawa E.M. 1955. Components of a difference between two rates. *Journal of the American Statistical Association*, vol. 50(272), pp. 1168-1194.

Nusselder, W.J. and Looman, C.W.N. 2004. Decomposition of differences in health expectancy by cause. Demography, vol. 41(2), pp. 315-334.

Pressat, R. 1985. Contribution des écarts de mortalité par âge à la différance des vies moyennes. *Population*, 4-5, pp. 766-770.

Robine, J.-M., Romieu, I., Cambois, E. 1999. Health expectancy indicators. *Bulletin of the World Health Organisation*, 77(2), pp. 181-185

Sullivan, D.F. 1964. A single index of mortality and morbidity. *HSMHA Health Report*, 86, pp. 347-354.