Demography (2012) 49:1207–1230
DOI 10.1007/s13524-012-0133-9

# A New Method for Determining Why Length of Life is More Unequal in Some Populations Than in Others

**Claudia Nau · Glenn Firebaugh**

Published online: 26 September 2012


**Abstract** Why is there greater variability in individual longevity in some populations than in others? We propose a decomposition method designed to address that question by quantifying the effects of population differences in the spread, allocation, and timing of the principal causes of death. Applying the method to the United States and Sweden, we find that spread effects account for about two-thirds of the greater variance in age at death among American adults, meaning that two-thirds of the U.S.-Sweden difference would persist if the two countries differed only with respect to within-cause variance among adults. The remainder of the difference is due largely to allocation effects, with the greater incidence of homicides and fatal traffic accidents alone accounting for more than one-fourth of the greater variance in age at death among adults in the United States.

**Keywords** Variance decomposition · Mortality · Cause of death · Life span inequality · Compositional effects

## Introduction

Variation in length of life is one of the most fundamental inequalities in human populations. The objective of this article is to introduce a variance-decomposition technique for determining why the variance in individual longevity differs from one population to another. Although current methods can point to the causes of death that contribute most heavily to those differences, they are not designed to tell us how

**Electronic supplementary material** The online version of this article (doi:10.1007/s13524-012-0133-9) contains supplementary material, which is available to authorized users.

C. Nau (✉)
Johns Hopkins Bloomberg School of Public Health, 615 N Wolfe Street, Baltimore, MD 21205, USA
e-mail: cnau@jhsph.edu

G. Firebaugh
Department of Sociology, Pennsylvania State University, 206 Oswald Tower, University Park, PA 16802, USA
e-mail: firebaugh@psu.edu

those causes produce the greater heterogeneity in some populations. In a standard cause-of-death decomposition, for example, we might find that heart disease accounts for much of the greater variance in longevity among the members of country A (versus country B), but we would not know whether heart disease matters because it strikes across a wider age swath in country A (spread effect), claims more victims (allocation effect), or strikes at different average ages (timing effect). By quantifying the effects of population differences in the spread, allocation, and timing of the principal causes of death, our method is designed to advance understanding of why heart disease and other causes of death matter (or not) for the degree of variation in the duration of life in human populations—and thus where efforts to reduce this variation are most likely to produce results.

To illustrate, we use our decomposition method to compare variance in the longevity of adults in the United States versus Sweden.[1] Because Sweden enjoys the lowest variance in longevity of nations in the West (Edwards and Tuljapurkar 2005), it serves as the best case against which to compare the United States.

## Variance in Longevity Within and Between Countries

The past century witnessed a notable rise in life expectancy throughout the world (Becker et al. 2005; White 2002) that was accompanied by convergence in the length of individuals' lives within countries as well as convergence in the average length of life across countries (Peltzman 2009; Wilson 2001).[2] The decline in mortality inequality within and between countries was propelled in its earlier stages by the "epidemiologic transition" (Omran 1971) that reduced deaths attributable to infectious and parasitic diseases that affected especially the young. In the United States, as in most high-income countries, this mortality compression lost momentum only after levels of infant and child mortality had decreased sufficiently to leave future mortality reductions to be made at older ages (Wilmoth and Horiuchi 1999). Mortality, then, became increasingly concentrated among the old, producing greater homogeneity in length of life as the "rectangularization" of the survival curve increased the proportion of people who survived to ages where mortality rates are high (Kannisto 2000; Wilmoth and Horiuchi 1999).

Mortality nonetheless remains less age-compressed in some high-income countries than in others (Mackenbach et al. 2008), and an emerging line of research asks why (Edwards and Tuljapurkar 2005; Ho and Preston 2010; van Raalte et al. 2010). To measure heterogeneity in age at death among *adults* (the issue we address here), Edwards and Tuljapurkar (2005) used a measure they called $S_{10}$: the standard deviation (SD) of life table ages at death among individuals who survive past age 10. Their central finding is that although adult age-at-death distributions for Britain, Canada, Denmark, France, Japan, Sweden, and the United States demonstrated a strong tendency toward convergence since 1960, "stark differences" (Edwards and

[1] Because variance is a measure of inequality when means are equal, in this study the decomposition of *variance* is roughly the same as the decomposition of *inequality* because the means of the age-at-death distributions of the United States and Sweden are not greatly different. Cross-national studies of inequality most often have focused on income, not mortality (e.g., Firebaugh 2003).

[2] The within- and between-country convergence in length of life might have stalled or reversed recently. See Shkolnikov (2004) and Goesling and Firebaugh (2004).

Tuljapurkar 2005:669) persist in the levels and trends in $S_{10}$ for these countries. Surprisingly, these differences in $S_{10}$ are not highly correlated with country differences in educational or income inequality. As a result, the sources of differential variability in age at death within low-mortality countries "remain unclear and await further research" (Edwards and Tuljapurkar 2005:669).

### Methods for Studying Differential Variance in Longevity Within Populations

Our study begins where the Edwards-Tuljapurkar study ended, with the challenge of accounting for national differences in $S_{10}$. Like their study, our study is based on life table deaths after age 10. Unlike Edwards and Tuljapurkar, however, we measure heterogeneity in age at death using variance—not standard deviation—because the difference in variance for two life table populations can be more conveniently decomposed into cause-specific spread, allocation, and timing components. Because we isolate the spread, allocation, and timing components of variance *within* the principal causes of death, our decomposition partitions variance along both a cause-of-death axis and a spread-allocation-timing (S-A-T) axis.

We could find no examples of prior mortality research that partitions along both axes. To be sure, many studies describe the differential mortality of—or difference in life expectancy across—*social groups*; and some of these studies also break down the group disparities by cause of death (Kallan 1997; Kitagawa and Hauser 1973; Mackenbach et al. 2008; Smith et al. 1990; Wong et al. 2002). A newer line of work on human longevity focuses on how cause-specific mortality relates to variability in the length of life of *individuals*. In their study of variance in individual longevity, for example, Edwards and Tuljapurkar (2005) recalculated $S_{10}$ after removing so-called external causes of death (death attributable to accidents, homicides, and suicides). That approach—estimating the effect of a cause by deleting it—is a common strategy in the cause-of-death literature. (See Beltrán-Sánchez et al. (2008) for discussion and references.)

Studies also use standardization (Chevan and Sutherland 2009; Kitagawa 1964) or replacement methods to compare the mortality regimes of two populations, cause by cause. Shkolnikov et al. (2003) described such a method for decomposing inequality in length of life as measured by the Gini coefficient. Specifically, they assessed how the difference in the Ginis for two populations changes as they progressively replace one cause-age–specific mortality rate after another. One complication of this type of decomposition is that "there are many different ways to replace group- and age-specific mortality rates and composition by group of one population by respective rates and composition by group of the other population" (Shkolnikov et al. 2003:332). Because S-A-T decomposition is not based on such sequential replacement, it permits a more definitive isolation of the sources of differential variance in longevity for two populations.

## The S-A-T Decomposition Scheme

Recall that we want to illustrate our method by showing why there is greater variance in age at death among adults in the United States than in Sweden. The S-A-T method

is based on the idea that the greater variance for Americans can stem from three principal mechanisms:

(1) Americans and Swedes die of causes at different rates. This would amplify the variance of the American (relative to the Swedish) mortality age distribution if Americans were more likely than Swedes to die of causes that tend to strike the young as well as the old (that is, causes that are highly variable with respect to age) or of causes that are centered far off the overall mean age at death. A good example is traffic accidents, a cause of death that disproportionately affects the young. Even if the mean and variance for age at death attributable to traffic accidents were exactly the same in the United States and Sweden, a higher rate of traffic deaths in the United States would nonetheless contribute to the greater variability in age at death in the United States because of the distinctive youthful age profile of this cause of death. In this way, then, variability of the mortality age distribution can be affected (increased *or* reduced) by the allocation of the causes of death in a society. We call this an *allocation effect*.

(2) Alternatively, Americans and Swedes might die of the same causes at the same rates, yet the variability in age at death could be greater for any or all causes in the United States. Under these circumstances, even if the average age at death for each cause were the same in the United States and Sweden, there would be greater heterogeneity in age at death in the United States because of this *spread effect*.

(3) Third, imagine that cause-specific death rates were the same for the American and Swedish populations (no allocation effects) and that the cause-specific age distributions had the same variance (no spread effects). Under those circumstances, the mortality age distributions for the United States and Sweden still could differ if the age distributions of specific causes were centered on different mean ages in the United States and in Sweden. In that case, variability in age at death will be greater in the country where the cause-specific means diverge the most. We call this a *timing effect*.

Figure 1 presents stylized representations of spread effects, allocation effects, and timing effects for Nations A and B. To simplify, we assume only four causes of death in each nation. Age is the *x*-axis, and number of life table deaths is the *y*-axis in Fig. 1, so the distributions in Fig. 1 depict the age distribution of deaths for each of the four causes. In each of the comparisons of Nation A with Nation B, overall variability in age at death is greater in Nation A, but for different reasons in the top, middle, and bottom panels. In the case of pure spread effects (top panel), individuals in the two countries die in the same proportions for the same causes; and victims of each cause die, on average, at the same age in both countries (as indicated by the dotted lines). However, in Nation A there is greater variance around those means. In the case of pure allocation effects, again victims of each cause die, on average, at the same age in both countries (as indicated by the dotted lines); but in this case, the within-cause variance is the same. Nevertheless, variability differs for the two nations because citizens in Nation A are more likely (than those in Nation B) to die of causes that disproportionately strike the young or the old. Finally, in the bottom panel, the only difference between Nations A and B is that the age distributions are centered over different means, with the distributions in Nation B overlapping more than the distributions in Nation A, resulting in pure timing effects.

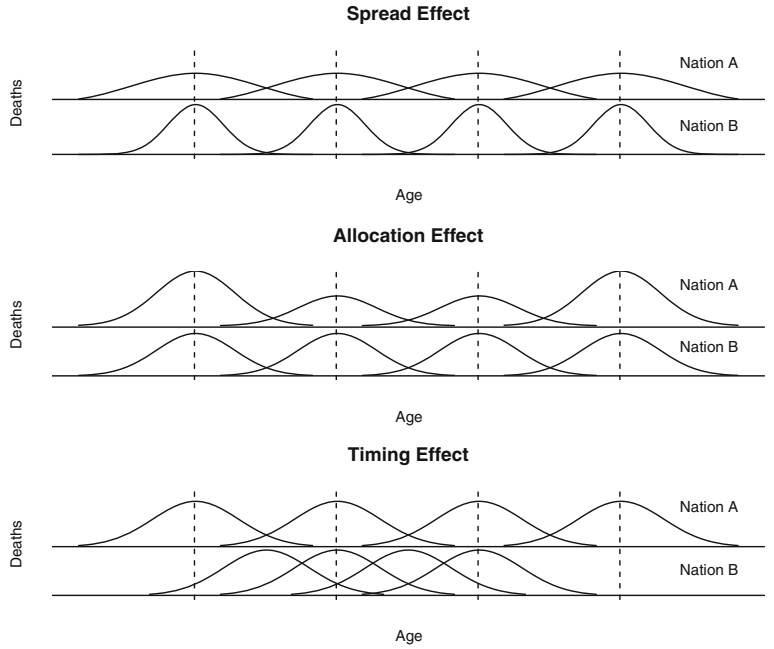


**Fig. 1** Graphic presentation of S-A-T decomposition

## The S-A-T Equations

S-A-T decomposition equations are based on the insight that the difference in the age variance of life table deaths for any two populations is a function of (1) population differences in the cause-specific within- and between-age variance of life table deaths weighted by (2) population differences in the number of life table deaths occurring within each cause category.[3] Let $x$ denote age at death, $\overline{x}_A$ and $\overline{x}_B$ denote the mean age of life table deaths for populations $A$ and $B$, respectively; $N$ denote number of life table deaths, and $\sigma^2$ denote variance of the age distribution of life table deaths. By definition, the *difference* in the variance of the age distribution of life table deaths between populations $A$ and $B$ is

$$\sigma_A^2 - \sigma_B^2 = \sum_{i=1}^{N_A}\left(x_{iA} - \overline{x}_A\right)^2 \Big/ N_A - \sum_{i=1}^{N_B}\left(x_{iB} - \overline{x}_B\right)^2 \Big/ N_B, \tag{1}$$

where there are $i = 1, 2, \ldots, N_A$ life table deaths in $A$ and $i = 1, 2, \ldots, N_B$ life table deaths in $B$. Because life table deaths are used, $N_A = N_B = N$, and Eq. (1) can be rewritten as

$$N\left(\sigma_A^2 - \sigma_B^2\right) = \sum_{i=1}^{N}\left(x_{iA} - \overline{x}_A\right)^2 - \sum_{i=1}^{N}\left(x_{iB} - \overline{x}_B\right)^2. \tag{2}$$

[3] Douglas L. Anderton earlier suggested the idea of decomposing the mean age at death to investigate the role of mean age and incidence changes of specific causes of death in the extension of the length of life (Nau and Beemer 2004). We extend that strategy to isolate S-A-T effects with respect to differential variance in the ages at death of two populations.

Focusing on the right side of Eq. (2), observe that within each population, we can array the data by cause of death and, for each cause of death $c$, sum $(x_i - \bar{x})^2$ across the $i = 1,2, \ldots, n_c$ life table deaths to calculate the contribution of that cause to $N\sigma^2$. Because all deaths are assigned a cause, the contributions of the $c = 1,2, \ldots, C$ causes sum to $N\sigma^2$ for each population, and the right side of Eq. (2) decomposes by cause of death as follows:

$$\begin{aligned}
&\sum_{i=1}^{N}\left(x_{iA}-\bar{x}_A\right)^2-\sum_{i=1}^{N}\left(x_{iB}-\bar{x}_B\right)^2= \\
&\sum_{c=1}^{C}\sum_{i=1}^{n_{cA}}\left(x_{icA}-\bar{x}_A\right)^2-\sum_{c=1}^{C}\sum_{i=1}^{n_{cB}}\left(x_{icB}-\bar{x}_B\right)^2= \\
&\sum_{i=1}^{n_{1A}}\left(x_{i1A}-\bar{x}_A\right)^2-\sum_{i=1}^{n_{1B}}\left(x_{i1B}-\bar{x}_B\right)^2+\sum_{i=1}^{n_{2A}}\left(x_{i2A}-\bar{x}_A\right)^2-\sum_{i=1}^{n_{2B}}\left(x_{i2B}-\bar{x}_B\right)^2+\ldots \\
&+\sum_{i=1}^{n_{CA}}\left(x_{iCA}-\bar{x}_A\right)^2-\sum_{i=1}^{n_{CB}}\left(x_{iCB}-\bar{x}_B\right)^2.
\end{aligned} \tag{3}$$

The pairs of differences in Eq. (3) are the gross contributions of each cause: they show how much each cause of death contributes to the overall difference in the variance in longevity for members of populations $A$ and $B$. To express these differences as proportions, we divide by $N\left(\sigma_A^2-\sigma_B^2\right)$ because it follows from Eqs. (2) and (3) that

$$\sum_{c=1}^{C}\left[\sum_{i=1}^{n_{cA}}\left(x_{icA}-\bar{x}_A\right)^2-\sum_{i=1}^{n_{cB}}\left(x_{icB}-\bar{x}_B\right)^2\right]\Big/N\left(\sigma_A^2-\sigma_B^2\right)=1.$$

Thus, for example, the term

$$\left[\sum_{i=1}^{n_{cA}}\left(x_{icA}-\bar{x}_A\right)^2-\sum_{i=1}^{n_{cB}}\left(x_{icB}-\bar{x}_B\right)^2\right]\Big/N\left(\sigma_A^2-\sigma_B^2\right)$$

is the proportion of the overall difference in the variance in longevity that is due to the first cause.

Next we decompose the contribution of the $c$th cause of death into its spread, allocation, timing, and joint components. The key insight is that $x_{icA}-\bar{x}_A = x_{icA}-\bar{x}_{cA}+\bar{x}_{cA}-\bar{x}_A$, where $\bar{x}_{cA}$ is the mean age of life table deaths for the $c$th cause in population $A$, so

$$\begin{aligned}
&(x_{icA}-\bar{x}_A)^2=[(x_{icA}-\bar{x}_{cA})+(\bar{x}_{cA}-\bar{x}_A)]^2= \\
&(x_{icA}-\bar{x}_{cA})^2+2(x_{icA}-\bar{x}_{cA})(\bar{x}_{cA}-\bar{x}_A)+(\bar{x}_{cA}-\bar{x}_A)^2.
\end{aligned}$$

Because $(\bar{x}_{cA}-\bar{x}_A)$ is constant for a given cause, the middle term, $2(x_{icA}-\bar{x}_{cA})(\bar{x}_{cA}-\bar{x}_A)$, drops out when we sum over $i$:

$$\sum_{i=1}^{n_c}2\left(x_{icA}-\bar{x}_{cA}\right)\left(\bar{x}_{cA}-\bar{x}_A\right)=2\left(\bar{x}_{cA}-\bar{x}_A\right)\sum_{i=1}^{n_c}\left(x_{icA}-\bar{x}_{cA}\right)=0.$$

Thus for the $c$th cause of death,

$$\begin{aligned}
&\sum_{i=1}^{n_{cA}}\left(x_{icA}-\bar{x}_A\right)^2-\sum_{i=1}^{n_{cB}}\left(x_{icB}-\bar{x}_B\right)^2= \\
&\sum_{i=1}^{n_{cA}}\left(x_{icA}-\bar{x}_{cA}\right)^2+n_{cA}\left(\bar{x}_{cA}-\bar{x}_A\right)^2-\left\{\sum_{i=1}^{n_{cB}}\left(x_{icB}-\bar{x}_{cB}\right)^2+n_{cB}\left(\bar{x}_{cB}-\bar{x}_B\right)^2\right\}= \\
&n_{cA}\sigma_{cA}^2+n_{cA}\Delta\bar{x}_{cA}^2-\left(n_{cB}\sigma_{cB}^2+n_{cB}\Delta\bar{x}_{cB}^2\right).
\end{aligned} \tag{4}$$

The term $\sigma_{cA}^2$ stands for $\sum_{i=1}^{n_{cA}}\left(x_{icA}-\bar{x}_{cA}\right)^2\Big/n_{cA}$. The expression $\Delta\bar{x}_{cA}^2$ is shorthand for $(\bar{x}_{cA}-\bar{x}_A)^2$; here, as before, $\bar{x}_A$ is the mean age of life table deaths across all causes in $A$, whereas $\bar{x}_{cA}$ is the mean age of life table deaths for the $c$th cause in $A$.

Choosing population $B$ as the reference group, Eq. (4) can be rewritten as follows (see the appendix for proof):

$$\begin{aligned}
&n_{cA}\sigma_{cA}^2+n_{cA}\Delta\bar{x}_{cA}^2-\left(n_{cB}\sigma_{cB}^2+n_{cB}\Delta\bar{x}_{cB}^2\right)= \\
&\left(\sigma_{cA}^2-\sigma_{cB}^2\right)n_{cB}+\left(n_{cA}-n_{cB}\right)\left(\sigma_{cB}^2+\Delta\bar{x}_{cB}^2\right)+\left(\Delta\bar{x}_{cA}^2-\Delta\bar{x}_{cB}^2\right)n_{cB} \\
&+\left(n_{cA}-n_{cB}\right)\left[\left(\sigma_{cA}^2-\sigma_{cB}^2\right)+\left(\Delta\bar{x}_{cA}^2-\Delta\bar{x}_{cB}^2\right)\right].
\end{aligned} \tag{5}$$

The terms in the right side of Eq. (5) are the spread, allocation, timing, and joint components, respectively, for the $c$th cause. To express these components as proportions of the total difference in the variances of the two populations, we divide each term by $N\left(\sigma_A^2-\sigma_B^2\right)$:

$$\begin{aligned}
&\text{spread effect}=\left(\sigma_{cA}^2-\sigma_{cB}^2\right)n_{cB}\Big/N\left(\sigma_A^2-\sigma_B^2\right) \\
&\text{allocation effect}=\left(n_{cA}-n_{cB}\right)\left(\sigma_{cB}^2+\Delta\bar{x}_{cB}^2\right)n_B\Big/N\left(\sigma_A^2-\sigma_B^2\right) \\
&\text{timing effect}=\left(\Delta\bar{x}_{cA}^2-\Delta\bar{x}_{cB}^2\right)n_{cB}\Big/N\left(\sigma_A^2-\sigma_B^2\right) \\
&\text{joint effect}=\left(n_{cA}-n_{cB}\right)\left[\left(\sigma_A^2-\sigma_B^2\right)+\left(\Delta\bar{x}_{cA}^2-\Delta\bar{x}_{cB}^2\right)\right]\Big/N\left(\sigma_A^2-\sigma_B^2\right).
\end{aligned} \tag{6}$$

The S-A-T method provides an exact decomposition of the difference in variance between two populations because the S-A-T components and the joint effect for each cause sum to the gross contribution of that cause, and those gross cause-specific contributions in turn sum to 100 % of the difference in the variance of the mortality age distributions (see the appendix).

To illustrate the S-A-T equations in Eq. (6), consider the contribution of traffic deaths to the difference in the variance of the mortality age distributions of the United States and Sweden (hereafter called *the U.S.-Sweden difference in variance* or simply *the difference in variance*). In 2006 (the year we use), the sum of squares to be decomposed is $N\left(\sigma_A^2 - \sigma_B^2\right)$ = 100,000 ($15.38763^2 - 12.68128^2$) = 7,596,429. From age-specific mortality rates, we compute 833.84 male life table traffic deaths in the United States compared with 291.07 in Sweden, with a difference of $\sigma_{cA}^2 - \sigma_{cB}^2$ = −22.16 for traffic deaths. Substituting these values into the spread effect equation in Eq. (6), we have (−22.16×291.07) / 7,596,429=−0.0008, multiplied by 100 to convert to a percentage, resulting in a minuscule spread effect of −0.08 % for traffic deaths. The more important effect here is the allocation effect: from Eq. (6), we have $\left(n_{cA} - n_{cB}\right)\left(\sigma_{cB}^2 + \Delta\bar{x}_{cB}^2\right)\Big/ N\left(\sigma_A^2 - \sigma_B^2\right)$ = (833.84 – 291.07) × (1,720.158) / 7,596,429 = 0.1229, or 12.29 %, where $\sigma_{cB}^2 + \Delta\bar{x}_{cB}^2$ = 1,720.158 is the sum of the Swedish within- and between-cause variance for traffic accidents. An allocation effect of 12.29 % indicates that 12.29 % of the difference in variance would have persisted if both countries had differed only in terms of the incidence of mortality from traffic accidents. To be very clear on this point, when we say that we estimate the allocation effect *using Sweden as the standard*, we mean that we recalculate the U.S. variance in age at death by substituting the Swedish values for all the components *except* allocation. Any difference that remains, then, is attributable to allocation, since the two countries now differ *only* with respect to cause-specific death rates. The other effects are calculated in similar fashion.

What if we had used population *A* instead of *B* as the reference in Eq. (6)? Observe in Eq. (6) that spread, allocation, and timing effects do not change sign (because reversing *A* and *B* in Eq. (6) reverses the signs of both numerator and denominator), but they can change in magnitude. The joint effect, by contrast, remains the same but reverses sign as the numerator is constant, while the denominator changes sign. Because joint effects remain the same size but reverse sign, and because we know that spread, allocation, timing, and joint effects must sum to the same number whether we use *A* or *B* as the reference, it follows that to balance the equation, the change in S-A-T effects must sum to twice the joint effect when we switch back and forth between *A* and *B* as the reference. (Online Resource 1 explains how joint effects can be used then to calculate the decomposition when the reference is reversed.)

The important point here is that joint effects are linked to the reference group issue. As Eq. (6) shows, the numerator of the joint effect is the sum of two terms: an allocation-spread interaction term, $\left(n_{cA} - n_{cB}\right)\left(\sigma_{cA}^2 - \sigma_{cB}^2\right)$, and an allocation-timing interaction term, $\left(n_{cA} - n_{cB}\right)\left(\Delta\bar{x}_{cA}^2 - \Delta\bar{x}_{cB}^2\right)$. (Only reference group *Ns*—not variances—are used as weights, so there is no spread-timing interaction term.) When both interaction effects are small, as they are in the present analysis, the choice of reference does not affect the conclusions.[4]

[4] Small joint effects could mask large but offsetting allocation-spread and allocation-timing interactions, so researchers should check for that possibility. Online Resource 1 shows how joint effects can be used to place conservative bounds around the S-A-T components.

## Data

This study uses mortality data for the United States and Sweden for the year 2006. Death counts by age, sex, and cause of death for the United States come from the multiple-cause-of-death data files available through the National Center for Health Statistics (NCHS 2009). The equivalent Swedish mortality data were provided upon request from the Swedish Death Registry. Both mortality databases provide the underlying cause of death, defined as "the disease or injury which initiated the train of morbid events leading directly to death, or the circumstances of the accident or violence which produced the fatal injury" (World Health Organization 2010). For both countries, cause of death is coded using the latest version (10th revision) of the International Classification of Disease (ICD) (see Online Resource 2).

We classify cause of death by adapting the coding scheme used to tabulate the 15 leading causes in the National Vital Statistics Report for 2005 (Heron et al. 2009). Our final categorization comprises 15 cause groups that are consistent for Sweden and the United States but reflect our data constraints and findings from prior analysis. Our classification consists of (1) the 10 leading causes of deaths in the two countries, (2) two key infectious disease categories (infectious diseases transmitted sexually (STDs) or through the shared use of needles (NTDs), and a remainder category of infectious diseases), (3) traffic accidents and homicides broken apart from other external causes of death, and (4) a residual category aggregating those causes that were not elsewhere classified (n.e.c.). The n.e.c. group assembles ill-defined causes of death as well as all causes that had too few deaths to stand on their own and could not be regrouped with any of the 14 specific categories. The multiple-cause-of-death files provide four-digit codes, the most detailed level of diagnosis; the Swedish Death Registry, for reasons of privacy protection, would provide only the three-digit codes for deaths that were tabulated in the one-year age intervals that we desired for our analysis. Slight changes were necessary to the NCHS cause list in the coding of the categories STDs/NTDs and traffic accidents. (See Online Resource 2 for a detailed description of the cause groupings.)

We used the 15 cause-of-death categories to calculate a multiple-decrement life table for each country. For these life tables, we used one-year age groups to calculate age-sex-cause–specific life table rates of decrement for Sweden and the United States for the year 2006.[5] Using life table deaths solves the problem of differences in population size and age composition across national populations, which would otherwise distort our results. The multiple-decrement life table for each country permits us to calculate how deaths would have been distributed across the 15 cause categories for men and women at each age assuming that two hypothetical cohorts of the same size had been subject to each country's age-sex-cause–specific rates of 2006 until their extinction. We assume that deaths occur, on average, in the middle of the one-year age intervals.

Because disease and mortality processes differ for children and adults, research on general health inequality or mortality processes often focuses either on infant and

[5] Age-specific estimates of the Swedish and U.S. populations are from the *Human Mortality Database* (HMD n. d.). We could not use the HMD as our source for mortality data because it does not provide information on cause-specific mortality.

child mortality or on adult mortality. We focus on the latter and start the life table at age 10.[6] Our analysis, then, bears on the health and mortality of aging populations, an issue of much current interest (Crimmins et al. 2010; Ho and Preston 2010; Lynch and Brown 2001; Rogers and Crimmins 2011).

Figure 2 shows the age distributions of life table deaths (occurring after age 10) for the United States and Sweden in 2006. The countries differ in two key respects. First, life expectancy at age 10 is 68.65 years for Americans versus 71.08 years for Swedes. Second, for those who survive past age 10, the variance in longevity among adults is 47 % greater in the United States than in Sweden (52 % greater when infant and child deaths are included). We turn now to the results of S-A-T decomposition to shed light on why $S_{10}$ is 15.39 for the United States versus 12.68 for Sweden.

## Results

### Cause-of-Death Results: Undercontributors and Overcontributors

Table 1 summarizes the decomposition results for the United States versus Sweden (total, and sex-specific). The 15 cause groups are listed, top to bottom, according to the proportion of life table deaths that they account for in the United States. Two causes—heart disease and cancer—account for over half of the life table adult deaths in the United States (28.3 % and 22.2 %, respectively), and they are listed first. At the bottom of our list are traffic accidents (1.2 % of all life table deaths), infectious diseases not transmitted sexually or by needles (0.6 %), infectious diseases transmitted sexually or by sharing needles (0.5 %), and homicides (0.4 %). To avoid an unwieldy table, a large number of minor causes are aggregated under the heading "other causes, not elsewhere classified (n.e.c.)." Collectively, these minor causes account for 19.1 % of all U.S. life table deaths and for 15.2 % of the greater variance in age at death among Americans.

Consider first the cause- and sex-specific contributions to the difference in variance. One important finding is that the various causes of death contribute disproportionately to the difference in variance, with some causes "overcontributing" and other causes "undercontributing" *relative to their incidence* (comparing the first two columns of results in Table 1). The biggest undercontributor is cancer, which accounts for 22.2 % of adult life table deaths in the United States but for almost none (−0.4 %) of the U.S.-Sweden difference in variance. Indeed, the four leading cause groups (heart disease, cancer, cerebrovascular diseases, and minor causes n.e.c.) are all undercontributors to some extent; although they collectively account for more than three-fourths of all life table deaths in the United States, their total contribution to the difference in variance is about 39 %.

Other causes, then, must "overcontribute" to the difference in variance. Overcontributors mostly are causes of death that disproportionately strike the young, resulting in age profiles disfavoring the country where they are more common. The principal overcontributors to the difference are traffic accidents (accounting for 1.2 %

[6] Results (available upon request) are the same when we construct life tables beginning at age 10 based on the observed survivors to age 10 in each country (99,591.42 for Sweden and 99,192.51 for the United States).

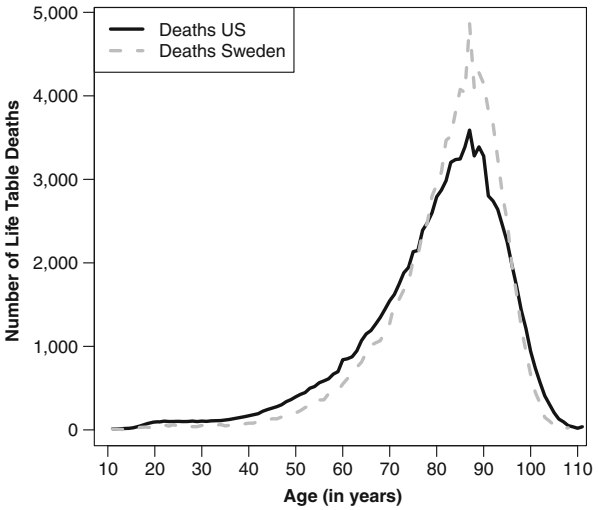


**Fig. 2** Life table deaths occurring after age 10, United States versus Sweden (based on age-specific death rates from 2006)

of life table deaths in the United States but 15.8 % of the U.S.-Sweden difference in variance), homicides (accounting for 0.4 % of U.S. deaths but 10.6 % of the difference in variance), other external causes (accounting for 3.6 % of U.S. deaths but 8.2 % of the difference in variance), and diseases transmitted sexually or by needles (accounting for 0.5 % of U.S. deaths but 4.8 % of the difference in variance).

In line with the findings of Edwards and Tuljapurkar (2005), we find no evidence that the overall U.S.-Sweden difference in variance is sex-specific. Summing over the 15 cause groups, we find that about 48 % of the difference would persist if only women differed in their mortality experience (that is, if American and Swedish women differed in their mortality experience, but American and Swedish men did not: see Table 1). Apparently, men and women each contribute roughly their share to the countries' difference in the age concentration of mortality.

## Beyond Cause of Death: Findings From the S-A-T Decomposition

Standard cause-of-death analyses would end here. To better understand the dynamics shaping the difference in the age concentration of deaths in the United States and Sweden, we used the S-A-T equations in Eq. (6) to decompose the overall and sex-specific contributions of each cause. Four findings stand out, as shown in Table 1. First, because joint effects are small, we can cleanly isolate the effects of spread, allocation, and timing.

Second, spread effects are primary: fully two-thirds (66 %) of the greater variance in the longevity of Americans would persist if the American and Swedish populations differed only in terms of *within-cause* variance in age at death. In fact, about 56 % of the variance gap would remain if the only differences between Americans and Swedes were the differences we observe for the spread effects of the four leading causes of death (heart disease, cancer, cerebrovascular disease, and causes not elsewhere classified (n.e.c.)).

**Table 1** S-A-T decomposition results for the difference in the variance of life table ages at death for the United States versus Sweden, based on age-specific rates from 2006

| | | Gross Contribution by Cause (%) | | | Spread Effects (%) | | | Allocation Effects (%) | | | Timing Effects (%) | | | Joint Effects (%) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | % of Adult Deaths in U.S. | Total | Female | Male | Total | Female | Male | Total | Female | Male | Total | Female | Male | Total | Female | Male |
| Heart | 28.27 | 23.06 | 14.77 | 8.29 | 25.19 | 12.77 | 12.42 | −3.48 | −0.73 | −2.75 | 3.38 | 3.24 | 0.14 | −2.03 | −0.51 | −1.52 |
| Cancer | 22.19 | −0.35 | −0.88 | 0.53 | 6.22 | 2.68 | 3.54 | −3.58 | −1.52 | −2.06 | −2.72 | −2.00 | −0.72 | −0.27 | −0.04 | −0.23 |
| n.e.c.[a] | 19.06 | 15.15 | 10.38 | 4.77 | 17.90 | 11.17 | 6.73 | −2.26 | −0.77 | −1.49 | 0.39 | 0.37 | 0.02 | −0.88 | −0.39 | −0.49 |
| Cerebro. | 6.20 | 0.97 | 1.32 | −0.34 | 6.34 | 3.83 | 2.51 | −4.08 | −2.18 | −1.90 | 1.20 | 1.18 | 0.02 | −2.48 | −1.51 | −0.97 |
| Respiratory | 5.31 | 4.76 | 2.71 | 2.05 | 1.20 | 0.66 | 0.55 | 2.55 | 1.43 | 1.11 | 0.04 | 0.04 | 0.00 | 0.97 | 0.58 | 0.39 |
| Alzheimers | 3.72 | 5.74 | 4.81 | 0.92 | 0.01 | 0.05 | −0.04 | 1.81 | 1.45 | 0.36 | 1.76 | 1.39 | 0.37 | 2.16 | 1.92 | 0.24 |
| Other external | 3.60 | 8.19 | 0.56 | 7.63 | 4.01 | 2.10 | 1.91 | −3.65 | −1.85 | −1.80 | 9.23 | 0.86 | 8.38 | −1.40 | −0.55 | −0.85 |
| Diabetes | 2.89 | 3.14 | 1.73 | 1.41 | 1.36 | 0.86 | 0.50 | 1.17 | 0.64 | 0.53 | 0.12 | −0.08 | 0.20 | 0.49 | 0.31 | 0.18 |
| Infl./Pneu. | 2.69 | 2.32 | 1.73 | 0.59 | 1.80 | 1.21 | 0.59 | 0.33 | 0.32 | 0.01 | 0.02 | 0.04 | −0.02 | 0.17 | 0.17 | 0.00 |
| Nephritis | 1.98 | 3.21 | 1.78 | 1.43 | 0.51 | 0.24 | 0.28 | 1.68 | 0.94 | 0.74 | 0.04 | 0.04 | 0.00 | 0.97 | 0.56 | 0.41 |
| Septicemia | 1.41 | 2.49 | 1.41 | 1.07 | 0.55 | 0.24 | 0.31 | 1.16 | 0.75 | 0.41 | 0.02 | 0.00 | 0.02 | 0.76 | 0.42 | 0.34 |
| Traffic | 1.21 | 15.81 | 4.89 | 10.92 | −0.03 | 0.05 | −0.08 | 17.29 | 5.00 | 12.29 | −0.47 | −0.08 | −0.39 | −0.97 | −0.08 | −0.89 |
| Other Infect. | 0.55 | 0.12 | −0.06 | 0.19 | 1.07 | 0.63 | 0.44 | −0.48 | −0.33 | −0.14 | −0.16 | −0.18 | 0.02 | −0.31 | −0.18 | −0.13 |
| STD/NTD | 0.49 | 4.82 | 1.45 | 3.37 | 0.03 | 0.00 | 0.03 | 3.51 | 1.10 | 2.41 | 0.16 | 0.06 | 0.11 | 1.12 | 0.30 | 0.82 |
| Homicide | 0.43 | 10.57 | 1.35 | 9.22 | 0.02 | 0.04 | −0.02 | 7.60 | 1.41 | 6.19 | 0.32 | −0.06 | 0.38 | 2.63 | −0.04 | 2.67 |
| Total % | 100 | 100 | 47.95 | 52.05 | 66.19 | 36.53 | 29.66 | 19.56 | 5.67 | 13.90 | 13.34 | 4.82 | 8.51 | 0.92 | 0.94 | −0.02 |

*Notes:* Joint effects are the sum of an allocation-spread and allocation-timing interaction term. In this analysis, both interaction terms are small, with the biggest being the male allocation-timing interaction of homicides with 2.82 %. Large interaction terms with opposite signs could mask simultaneous differences in cause-specific allocation-spread and allocation-timing effects. Therefore, before presenting joint effects, researchers should check for this possibility.

[a]The n.e.c. ("not elsewhere classified") group consists of ill-defined causes of death as well as all causes that had too few deaths to stand on their own and that could not be placed in the other cause groups. *Cerebro.* refers to cerebrovascular diseases, *Respiratory* to diseases of the lower respiratory tract, *Infl./Pneu.* to influenza and pneumonia, *Other Infect.* to other infectious diseases, and *STD/NTD* to sexually and needle-transmitted diseases.

Third, spread effects are nearly always positive, indicating stubbornly greater cause-specific variance in age at death for Americans. The spread effect for heart disease alone accounts for one-fourth of the total U.S.-Sweden difference in variance. Interestingly, the spread effects are remarkably similar for women and men (12.8 vs. 12.4), suggesting that the risk factors increasing the vulnerability to heart disease along the age range in the United States are not sex-specific. Spread effects for minor causes not elsewhere classified (n.e.c.) account for another 18 % of the U.S.-Sweden difference in the variances of the ages at death; but in this instance, the differences are much larger for women than for men.

The fourth finding that stands out is that spread effects, although important, do not fully account for the U.S.-Sweden difference in variance. Despite the relative similarity of the cause-of-death structures of the United States and Sweden (Himes 1994), allocation effects sum to nearly 20 % of the difference. Timing effects account for virtually all the remainder; joint effects account for less than 1 %, with the allocation-spread and allocation-timing interaction effects (not shown) being negligible.

Unlike spread effects, which are consistently positive or negligible across all causes of death, allocation effects and timing effects are sometimes negative. The presence of negative effects is noteworthy because it indicates that moving the U.S. mortality regime toward Sweden's would in some instances serve to *increase, not narrow*, the greater variability in length of life among Americans.

Because allocation and timing effects are positive for some causes and negative for others, we must examine the main contributing cause groups separately to understand how allocation and timing effects contribute to the difference in variance. We begin with cancer, where the allocation and timing effects are negative. Hence, our findings suggest that eliminating differences in the allocation and timing of cancer deaths in the United States and Sweden, all else remaining the same, would *increase* the difference in the variance in age at death for individuals in the two countries. Because the spread effect works in the opposite direction, the gross contribution of cancer to the overall difference is negligible (−0.4 %).

Next consider heart disease, the leading killer in the United States. Figure 3 compares the male and female distributions of the life table ages at death from heart disease for the United States and Sweden. American men and women are less likely to die of heart disease than their same-sex counterparts in Sweden. These results are visible in Fig. 3, and are reflected in the negative allocation effects for heart disease in Table 1. The age profiles of life table deaths in Fig. 3 are noteworthy: Swedish men and American women are about equally likely to be victims of heart disease, while Swedish women are the most likely and American men the least likely to be victims.[7]

The heart disease timing effect is essentially zero for men, indicating similarity in the U.S. and Swedish male populations in the timing of heart disease deaths relative to the timing of other deaths in their respective populations. For women, the timing effect is positive, indicating greater disparity in the timing of heart disease deaths versus other types of deaths among American women than among Swedish women. It turns out that the average age at death from heart disease is similar for American and Swedish women. However, because the overall average age at death is lower in the

[7] Population-based heart disease death rates are higher for men than for women. This pattern is reversed after our life table nets out the age-compositional effects of the population.

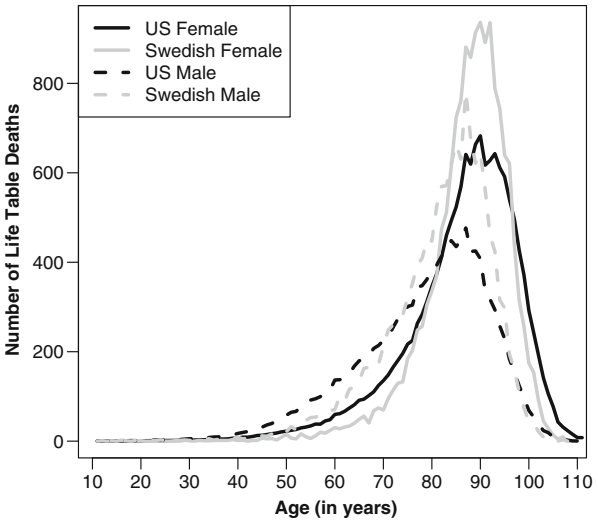


**Fig. 3** Sex-specific life table deaths occurring after age 10, United States versus Sweden: Heart disease (based on cause-age–specific rates from 2006)

United States, the U.S. age pattern is stretched out, and population heterogeneity is increased. Ironically, then, in the case of heart disease, the U.S.-Sweden difference in variance is elevated by the fact that American women who are heart disease victims die, on average, at about the same age as female Swedish victims.

Next consider traffic fatalities—a major overcontributor to the difference in variance. From the S-A-T decomposition, we immediately identify allocation effects, especially for men, as the source of that overcontribution. Altogether, Americans' greater risk of being traffic victims accounts for an estimated 15.8 % of the greater variance in the age at death among Americans, and this is entirely an allocation effect (see Table 1).

Allocation effects also underlie the contribution of homicides to the difference in variance. As shown by the S-A-T decomposition, the greater homicide victimization rate for males in the United States alone amounts to 6.2 % of the greater variance in age at death in the United States compared with Sweden. Along those lines, it is important to bear in mind that death from one cause precludes death from another. Because homicide and traffic accident victims, who tend to be young, are not at risk of dying later, we expect positive allocation effects for homicide and traffic accidents to be paired with negative allocation effects for "older-age causes," such as heart disease, cancer, and cerebrovascular disease—and that is exactly what we find, as illustrated in Table 1. Moreover, given that American men are more likely than American women to be homicide or traffic accident victims, on the basis of competing risks alone, we would expect the allocation effect for heart disease, cancer, and cerebrovascular disease to be more negative for men than for women. Again, that is what we find for heart disease and cancer, but not for cerebrovascular disease.

Finally, note the relatively large timing effect (8.4 %) for men for "other external causes." Figure 4 shows the dynamics underlying this effect. In the United States, 62 % of all deaths in this cause category occur before the age of 60, with the leading causes of deaths being accidental poisoning (mostly from narcotics) and suicide. In Sweden, by contrast, only 36 % of deaths in this cause category occur before age 60—most of which

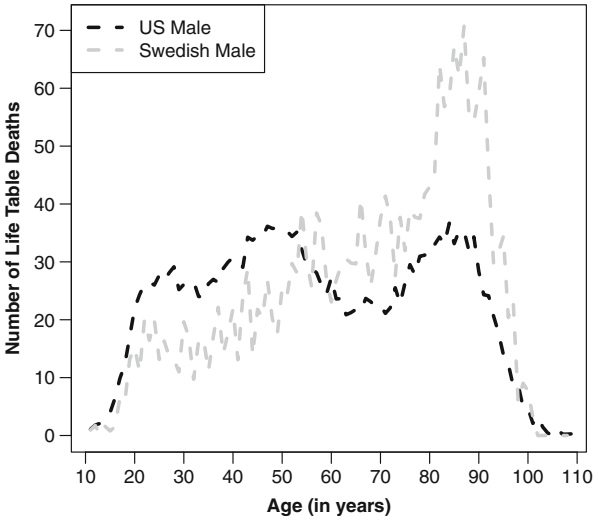


**Fig. 4** Life table deaths occurring after age 10, United States men versus Swedish men: Suicides and accidents other than traffic accidents (based on cause-age–specific rates from 2006)

are due to suicides and, to a lesser degree, to accidental poisoning. In both countries, mortality in this category for those older than 60 is due mostly to falls and suicides.

## Comparing the United States and Sweden at the Same Life Expectancy

A reviewer asked whether we would find the same patterns if we decomposed the difference between the United States in 2006 and Sweden in an earlier period when Swedes had the same life expectancy (at age 10) as Americans did in 2006. We had to go back to 1993 in Sweden to make the match ($e_{10}$ = 68.73 in Sweden in 1993 vs. 68.65 in the United States in 2006). This complicated the analysis in two ways. First, deaths in 1993 and 2006 are coded in different versions of the International Classification of Disease, which forced us to create cause groups that bridge the differences in the two coding schemes, resulting (among other things) in a much larger residual category of unclassified causes of death (see Online Resource 2).[8] Second, although our 2006 data consist of one-year age groups to age 110+, the 1993 data for Sweden are from the World Health Organization mortality database (World Health Organization n.d.), which provides only five-year age groups with the open-ended age interval being 85+ years.[9] The first step, then, was to recalculate the S-A-T decomposition on

[8] The residual category here—causes not elsewhere classified (n.e.c.)—is necessarily large because coding changes from 1993 to 2006 made it impossible to separate causes of death other than heart disease, cancers, and the others listed in Table 1.

[9] Because the WHO Mortality Database does not contain information for the United States for 2006, we had to aggregate the data from the multiple-cause mortality files for the United States in 2006 into the more crude age categories used in the WHO mortality data for Sweden in 1993. To gauge whether using the broader age categories matters for our results, we recalculated $S_{10}$ for Sweden in 2006, using the more crude age categories. The results are reassuring ($S_{10}$ = 12.68 years using the original age categories vs. 12.44 years using the cruder WHO format).

the 2006 data using the more crude cause and age grouping necessary to accommodate differences in the format of the data for the 1993–2006 comparison. Those results are shown in Table 2 under the "Same Year" heading. Then, to compare those 2006 results with the results when both countries had comparable levels of life expectancy at age 10, we decomposed the difference in variance for Sweden in 1993 versus the United States in 2006. Those results are also shown in Table 2, under the "Same $e_{10}$" heading.

Given the 13-year time gap from 1993 to 2006, two surprises appear in our results. The first surprise (not shown in Table 2) is that in Sweden, the SD in age at death after age 10 declined only from 12.90 to 12.44 (based on five-year age categories, not the one-year categories used in Table 1). This result suggests that the greater variance among Americans versus Swedes in 2006 is not linked to their differences in longevity in 2006 because the difference in variance was roughly the same when they had the same life expectancy.

The second surprise is the similarity of the same-year and same-life expectancy decompositions. Although the S-A-T results reveal some differences between the components for the same-year and same-$e_{10}$ comparisons, the essential story is the same. Regardless of whether we match the United States and Sweden on calendar year or on life expectancy, spread effects dominate, accounting for 63.1 % of the difference in same-year variance and 61.7 % of the difference in same-$e_{10}$ variance. In line with our earlier decomposition reported in Table 1, allocation effects are next, with timing effects contributing much less. Note, however, that allocation effects are significantly smaller for some causes in the 2006 comparison and significantly smaller for other causes in the 1993–2006 comparison, with the former outweighing

**Table 2** Comparison of S-A-T decomposition of the United States and Sweden at the same year (2006) and at the same level of life expectancy at age 10 (Sweden in 1993 and United States in 2006), both sexes jointly (percentages)

| | Gross Contribution (%) | | Spread (%) | | Allocation (%) | | Timing (%) | | Joint (%) | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Same Year | Same $e_{10}$ | Same Year | Same $e_{10}$ | Same Year | Same $e_{10}$ | Same Year | Same $e_{10}$ | Same Year | Same $e_{10}$ |
| Infectious | 3.02 | 2.75 | 0.81 | −0.32 | 1.47 | 3.61 | −0.01 | −0.06 | 0.75 | −0.48 |
| Cancer | 1.04 | 6.53 | 7.41 | 5.31 | −3.30 | 4.89 | −2.78 | −3.74 | −0.29 | 0.07 |
| Cerebro. | 0.79 | −3.45 | 6.19 | 6.89 | −4.13 | −8.16 | 1.16 | 1.31 | −2.43 | −3.49 |
| Heart | 20.81 | 12.45 | 23.64 | 30.67 | −3.10 | −13.66 | 2.00 | 2.46 | −1.73 | −7.02 |
| Diabetes | 3.87 | 4.65 | 1.80 | 1.00 | 1.30 | 2.98 | 0.06 | 0.00 | 0.71 | 0.67 |
| HIV | 4.08 | 3.93 | −0.01 | 0.03 | 3.64 | 4.87 | 0.05 | −0.25 | 0.40 | −0.72 |
| Transport | 16.68 | 16.38 | −0.13 | −0.14 | 17.48 | 15.69 | −0.14 | 0.48 | −0.52 | 0.35 |
| Other External | 8.49 | 10.05 | 4.57 | 4.51 | −3.61 | 0.18 | 8.94 | 5.22 | −1.41 | 0.14 |
| Homicide | 11.21 | 13.23 | 0.01 | 0.01 | 8.06 | 7.42 | 0.36 | 1.16 | 2.78 | 4.64 |
| n.e.c. | 29.99 | 33.48 | 18.81 | 13.72 | 6.59 | 15.51 | 1.85 | 0.49 | 2.75 | 3.76 |
| Total % | 100 | 100 | 63.10 | 61.67 | 24.40 | 33.33 | 11.50 | 7.08 | 0.99 | −2.08 |

*Notes*: *Cerebro*. refers to cerebrovascular diseases, *transport* refers to transportation-related deaths, and *n.e.c.* refers to causes not elsewhere classified.

the latter, suggesting that U.S.-Sweden differences in cause-specific death rates on the whole narrowed during the 13-year period.[10]

With regard to specific causes of death, transport accidents in Table 2 parallel the results for the cause category "traffic accidents" in Table 1, accounting in every instance for about 16 % of the difference in variance—a contribution driven by allocation effects. Homicides also have very similar effects whether we match the United States and Sweden on year or on life expectancy (see Table 2). In both cases, spread effects and timing effects are negligible, so the gross contribution of homicides owes mainly to allocation effects—and this contribution is not trivial because, as shown in Table 2, homicides account for 11 %–13 % of the overall difference in variance.

## Further Applications

By isolating cause-specific spread, allocation, and timing effects, S-A-T decomposition provides a useful tool for determining the specific issues that need to be researched to understand why length of life varies more in some societies than in others. For example, the large spread effects found in Tables 1 and 2 point future research to the question of why Americans are at risk of being heart disease victims over a wider swath of their lives than Swedes are. S-A-T decomposition, then, constitutes a strategic first step in comparing mortality regimes to identify the tractable sources of inequality for populations. In this section, we briefly discuss how the method can be applied to (1) determine whether results are being driven by particular groups that are present in only one of the populations and (2) handle the comparison of more than two populations.

### Accommodating Differences in Population Composition

America's population is more racially diverse than Sweden's, and that greater diversity could affect the U.S.-Sweden difference in variance. Subpopulations affect the difference in variance when they account for a disproportionate share of the difference. To determine whether a subpopulation accounts for a disproportionate share, we simply insert additional columns in the S-A-T table (as we did for men and women in Table 1) when the subpopulation is present in both populations.

Determining the contributions of a subpopulation is more complicated when the subpopulation is too small in one of the populations to permit the calculation of reliable multiple-decrement life tables. The small population of nonwhites in Sweden is a case in point, since one might wonder how much of the U.S.-Sweden difference is due to differences in the racial and ethnic compositions of the two populations (Tuljapurkar and Edwards 2011). Consider two questions:

(1) How much of the *observed* U.S.-Sweden difference in variance is due to nonwhites in the American population?
(2) How large would the U.S.-Sweden difference be if we removed nonwhites from the American population?

[10] The other possibility is a significant decline in the weighting term $\left(\sigma_{cB}^2 + \Delta\,\overline{x}_{cB}^2\right)\Big/ N\left(\sigma_A^2 - \sigma_B^2\right)$ for allocation effects (see Eq. (6)), but it is beyond the scope of our study to examine the 1993–2006 change in Sweden versus the United States.

To appreciate the difference between the two questions, note that if we remove nonwhites completely (as called for in question 2), we would then be comparing a different "America"—namely, an America with different cause-specific and overall means—so we would not be addressing the issue of how much the *actual* U.S.-Sweden difference in variance is accounted for by differences in the racial/ethnic composition of the two countries.

To address question 1, we recalculated spread, allocation, and timing effects, using the variance[11] of ages at death of the non-Hispanic white population around the actual overall and cause-specific means (based on both whites and nonwhites). Table 3 shows those results. Note first that the sum of spread effects for whites is smaller than the sum for the entire population. This difference in sums is accounted for by variance in age at death in the top four cause categories (heart disease, cancer, cerebrovascular disease, and minor causes not elsewhere classified), indicating that in each of the four categories, age at death for whites tends to cluster more tightly around the cause-specific age mean than does age at death for nonwhites. For other causes, variance in age at death around the cause-specific mean is about the same for whites as it is for the entire population.

The sum of allocation effects is also smaller for whites than it is for the entire population. In this case, the difference is due largely to higher nonwhite rates of deaths from traffic accidents and homicides. Overall, however, the central conclusion is the same: spread effects, along with allocation effects associated with external causes of death, contribute the most to the greater inequality in age at death among Americans as opposed to Swedes. In other words, despite nontrivial differences between non-Hispanic whites and nonwhites in the spread, allocation, and timing effects for some specific causes, the basic pattern of results that we found earlier is not driven by the existence of a more racially and ethnically heterogeneous population in the United States.

### Comparing Multiple Populations

In some instances, we might want to think of the United States and Sweden as members of a super-population of rich nations, or we might be interested in comparing populations on the North American continent or within the Organisation for Economic Development (OECD) to a given standard. The decomposition results would be based on multiple-decrement life tables calculated from cause-age-nation–specific life table decrement rates. As with any decomposition, however, there must be a reference population. One possibility is to use an exemplary population as the standard, as we have done here with Sweden. Another possibility is to use the mean of the populations as the standard so that results can be interpreted as deviations from the average for the super-population.

## S-A-T Decomposition and Future Research

S-A-T decomposition can be used to compare variances for any populations for which the number of cause-specific deaths by age can be computed via multidecrement life tables. The method can be used cross sectionally, to compare populations or subpopulations at a given point in time; and longitudinally, either to determine the sources of

[11] We use variance rather than sum of squares, since the latter would inherently be smaller because of the smaller *N*.

**Table 3** S-A-T effects based on entire population and on non-Hispanic whites only (2006 data)

| | Spread Effects (%) | | Allocation Effects (%) | | Timing Effects (%) | |
|---|---|---|---|---|---|---|
| | Table 1 Population | Based on Whites Only | Table 1 Population | Based on Whites Only | Table 1 Population | Based on Whites Only |
| Heart | 25.19 | 20.87 | −3.48 | −2.94 | 3.38 | 3.99 |
| Cancer | 6.22 | 3.32 | −3.58 | −3.57 | −2.72 | −4.54 |
| n.e.c. | 17.90 | 10.19 | −2.26 | −1.66 | 0.39 | 1.98 |
| Cerebro. | 6.34 | 3.73 | −4.08 | −4.17 | 1.20 | 2.48 |
| Respiratory | 1.20 | 0.91 | 2.55 | 3.07 | 0.04 | 0.02 |
| Alzheimer | 0.01 | −0.02 | 1.81 | 2.06 | 1.76 | 1.86 |
| Other External | 4.01 | 4.75 | −3.65 | −3.95 | 9.23 | 5.42 |
| Diabetes | 1.36 | 1.02 | 1.17 | 0.56 | 0.12 | −0.04 |
| Infl./Pneu. | 1.80 | 1.28 | 0.33 | 0.41 | 0.02 | 0.37 |
| Nephritis | 0.51 | 0.29 | 1.68 | 1.46 | 0.04 | 0.12 |
| Septicemia | 0.55 | 0.40 | 1.16 | 1.01 | 0.02 | 0.01 |
| Traffic | −0.03 | 0.16 | 17.29 | 14.16 | −0.47 | −1.37 |
| Other Infect. | 1.07 | 0.69 | −0.48 | −0.47 | −0.16 | −0.14 |
| STD/NTD | 0.03 | 0.02 | 3.51 | 1.54 | 0.16 | 0.34 |
| Homicide | 0.02 | 0.18 | 7.60 | 1.91 | 0.32 | −0.22 |
| Total % | 66.19 | 47.76 | 19.56 | 9.42 | 13.34 | 10.28 |

*Notes:* See notes to Table 1 for explanation of abbreviations. The equations for whites only are as follows:

$$\textit{White population spread effect} = \left(\sigma^2_{cAW} - \sigma^2_{cB}\right) n_{cB} \Big/ N\left(\sigma^2_A - \sigma^2_B\right).$$

$$\textit{White population allocation effect} = \left(n_{cAW} - n_{cB}\right)\left(\sigma^2_{cB} + \Delta\overline{x}^2_{cB}\right) \Big/ N\left(\sigma^2_A - \sigma^2_B\right).$$

$$\textit{White population timing effect} = \left(\Delta\overline{x}^2_{cAW} - \Delta\overline{x}^2_{cB}\right) n_{cB} \Big/ N\left(\sigma^2_A - \sigma^2_B\right).$$

The new terms are defined as follows (see text and appendix for definitions of other terms):

$$\sigma^2_{cAW} = \sum_{i=1}^{n_{cA}} \left(x_{icAW} - \overline{x}_A\right)^2 \Big/ n_{cAW}$$

where

$x_{icAW}$ is age at death for $i$th non-Hispanic white victim of cause $c$ in population A;

$n_{cAW}$ is number of non-Hispanic white deaths attributable to cause $c$ per 100,000 total white deaths;

$\Delta\overline{x}^2_{cAW} = \left(\overline{x}_{cAW} - \overline{x}_A\right)^2$, where $\overline{x}_{cAW}$ is mean age at death of non-Hispanic victims of cause $c$.

change in variance over time for a given population or to determine the sources of convergence or divergence of the mortality regimes of different populations.

A limitation of decomposition methods in general is their assumption that causes of death are independent, when in fact causes of death are not independent but compete against each other and therefore condition each other's presence in the mortality regime (Beltrán-Sánchez et al. 2008; Chiang 1991). Although the S-A-T scheme does not solve the problem, it provides traction on it by isolating the allocation component. Recall, for example, our

finding that positive allocation effects for homicide and traffic accidents are paired with negative allocation effects for "older-age causes," such as heart disease, cancer, and cerebrovascular disease—precisely as competing risks would suggest.

Like other cause-of-death analyses, the decompositions here are sensitive to the grouping of the causes of death. Cause groupings need to be etiologically sensible and not too broad if we want to detect differences in the cause-specific mechanisms that are underlying variability in length of life. At the same time, overly narrow cause groupings can lead to unwieldy results that are hard to interpret. Moreover, cause-of-death diagnosis and coding is not an exact science. The more specificity, the more disagreement there is likely to be over the underlying cause of death. The problem is likely to be exacerbated when comparing results from two populations with different health care systems and medical traditions; Swedish and American coders who agree at a broad level may nonetheless select different causes at a more disaggregated level.

Finally, in addressing the question of why longevity varies more in some populations than in others, there is some disagreement over whether investigations should focus on variability in age at death for individuals or on variability in average age at death for groups of individuals.[12] The S-A-T method readily accommodates either approach. On one hand, we might assume implicitly that two populations are equally robust (or frail) and thus have the same potential age-at-death variance—we call that the "population uniformity assumption"—and interpret any differences in that variance as reflecting varying exposure to risks for the two populations. When that assumption is dubious, it is usually a straightforward matter to partition the variance further to reach levels where the assumption is more plausible. For example, we separated men and women; others might wish to partition the variance along additional dimensions. An appealing feature of the S-A-T method is that given the requisite life tables, variance can be decomposed by nesting smaller populations within larger ones to accommodate alternative assumptions about population uniformity.

## Summary and Conclusion

Based on their analysis of over 9,000 national life tables, Smits and Monden (2009:1114) concluded that, "At similar levels of life expectancy, substantial differences in [age-at-death] inequality are observed, even among highly developed countries." In this study, we investigate the epidemiological underpinnings of such national differences. We begin with the leading causes of death because age differences for leading causes have greater potential to affect the amount of variability in age at death *within* a population. A leading cause of death also has more potential to age-compress (or age-dilate) the mortality experience of a population *relative to some other population* if that cause has a low (high) spread relative to its spread in the other population.

A good example of this is heart disease for Sweden versus the United States. Heart disease is the leading killer of both Swedes and Americans, so age differences for heart

[12] Studies in the fields of epidemiology and public health often focus on variability in life expectancies for meaningful aggregates of individuals, as determined by race, social class, or geography (e.g., Hahn and Eberhardt 1995; Lantz et al. 1998; Murray et al. 2006). Other researchers (e.g., Gakidou and King 2002) aver that there is more to be gained by examining variability across individuals rather than across categories of individuals.

disease victims greatly determine the variance in age at death *within* each country. Moreover, because deaths attributable to heart disease are more age-concentrated in Sweden than in the United States, heart disease also greatly affects the U.S.-Sweden *difference* in the variance in age at death within each country.

That said, another important finding of this study is that the leading causes of death are not necessarily the leading contributors to population differences in variance. Indeed, the four leading killers collectively account for more than three-fourths of all deaths in the United States but for only about 39 % of the U.S.-Sweden difference in variance. To offset that "undercontribution" of the four leading killers, some rarer causes of death must overcontribute to the difference in variance. One notable finding here is that, owing almost entirely to allocation effects, traffic fatalities and homicides combined account for as much of the difference in variance (26 %) as do heart disease and cancers combined (23 %), even though heart disease and cancers account for about 30 times more deaths in the United States.

These findings have practical implications. To the extent that Sweden can be taken as representing today's "best health practices" (Edwards and Tuljapurkar 2005), we can single out the most promising targets for policy interventions to reduce inequality in the length of life of Americans. For example, our findings suggest that reducing cancer deaths would have limited effect in narrowing the inequality gap between the United States and Sweden. On the other hand, effective policy interventions to reduce homicides and traffic fatalities would have immediate impact on inequality in the duration of life of Americans. That is not to say that effective policies would be easy to implement; for example, traffic fatalities increase with miles traveled, and the United States is a large country. But it is to say that external risk factors in the United States are readily identifiable and relatively limited in numbers compared with internal causes, such as heart disease and cancers, whose etiologies are complex and for which risk factors are numerous. Our results, then, provide reason for optimism that mortality inequality in the United States could be significantly reduced by addressing a limited set of risk factors responsible for premature deaths from external causes.

**Acknowledgments** Nau was supported by a National Institute of Child Health and Human Development Training Grant from the Population Research Institute at Pennsylvania State University. This research was partially supported through an NIH infrastructure grant to the Population Research Institute (R24HD041025). Versions of this article were presented at the annual meeting of the Population Association of America in March 2009, at Indiana University in September 2010, at Pennsylvania State University in February 2011, at the University of Texas in April 2011, and at Oxford University in November 2011. We thank Doug Anderton, Patrick Heuveline, and Bob Hummer for helpful and reassuring comments.

## Appendix

Here we present proof that the S-A-T equations provide an exact decomposition of the difference in the variances of two groups $\left(\sigma_A^2 - \sigma_B^2\right)$.

Notation for Group A (Notation for Group B is Analogous)

$n_{cA}$ = number of life table deaths in group $A$ contributed to the $c$th cause

$\sigma_{cA}^2$ = variance of the age distribution of life table deaths for the $c$th cause in group $A$
$n_{cA}\sigma_{cA}^2$ = sum of squares of the age distribution of life table deaths for the $c$th cause in group $A$

$$\Delta\bar{x}_{cA}^2 = \left(\bar{x}_{cA} - \bar{x}_A\right)^2$$

$\bar{x}_A$ is mean age of life table deaths for group $A$
$\bar{x}_{cA}$ is mean age of life table deaths in group $A$ attributed to the $c$th cause.

$n_{cA}\Delta\bar{x}_{cA}^2$ is the contribution of the $c$th cause to the between-cause sum of squares of the age distribution of life table deaths in group $A$:

$v_{cA}^2 = \sigma_{cA}^2 + \Delta\bar{x}_{cA}^2$.

Proof

We first show that the contribution of the $c$th cause of death to the difference in the total sum of squares equals

$$n_{cB}\left(\sigma_{cA}^2 - \sigma_{cB}^2\right) + \left(n_{cA} - n_{cB}\right)v_{cB}^2 + n_{cB}\left(\Delta\bar{x}_{cA}^2 - \Delta\bar{x}_{cB}^2\right) + \left(n_{cA} - n_{cB}\right)\left(v_{cA}^2 - v_{cB}^2\right).$$

That is, it equals the sum of the numerators for the spread, allocation, timing, and joint effects:

Contribution of cause $c$ to $N\left(\sigma_A^2 - \sigma_B^2\right) =$

$$\sum_{i=1}^{n_{cA}}\left(x_{icA} - \bar{x}_A\right)^2 - \sum_{i=1}^{n_{cB}}\left(x_{icB} - \bar{x}_B\right)^2 \text{ (from Eq. (4) in text)} =$$

$$n_{cA}\sigma_{cA}^2 + n_{cA}\Delta\bar{x}_{cA}^2 - \left(n_{cB}\sigma_{cB}^2 + n_{cB}\Delta\bar{x}_{cB}^2\right) =$$

$$n_{cA}\sigma_{cA}^2 + n_{cA}\Delta\bar{x}_{cA}^2 - \left(n_{cB}\sigma_{cB}^2 + n_{cB}\Delta\bar{x}_{cB}^2\right) + \left(n_{cA}\sigma_{cB}^2 - n_{cA}\sigma_{cB}^2\right) + \left(n_{cB}\sigma_{cA}^2 - n_{cB}\sigma_{cA}^2\right)$$
$$+ \left(n_{cB}\sigma_{cB}^2 - n_{cB}\sigma_{cB}^2\right) + \left(n_{cA}\Delta\bar{x}_{cB}^2 - n_{cA}\Delta\bar{x}_{cB}^2\right) + \left(n_{cB}\Delta\bar{x}_{cA}^2 - n_{cB}\Delta\bar{x}_{cA}^2\right)$$
$$+ \left(n_{cB}\Delta\bar{x}_{cB}^2 - n_{cB}\Delta\bar{x}_{cB}^2\right) =$$

$$n_{cB}\left\{\left(\sigma_{cA}^2 - \sigma_{cB}^2\right) - \left(\sigma_{cB}^2 + \Delta\bar{x}_{cB}^2\right) + \left(\Delta\bar{x}_{cA}^2 - \Delta\bar{x}_{cB}^2\right) - \left[\left(\sigma_{cA}^2 + \Delta\bar{x}_{cA}^2\right) - \left(\sigma_{cB}^2 + \Delta\bar{x}_{cB}^2\right)\right]\right\}$$
$$+ n_{cA}\left[\left(\sigma_{cB}^2 + \Delta\bar{x}_{cB}^2\right) + \left(\sigma_{cA}^2 + \Delta\bar{x}_{cA}^2\right) - \left(\sigma_{cB}^2 + \Delta\bar{x}_{cB}^2\right)\right] =$$
$$n_{cB}\left(\sigma_{cA}^2 - \sigma_{cB}^2\right) + \left(n_{cA} - n_{cB}\right)v_{cB}^2 + n_{cB}\left(\Delta\bar{x}_{cA}^2 - \Delta\bar{x}_{cB}^2\right) + \left(n_{cA} - n_{cB}\right)\left(v_{cA}^2 - v_{cB}^2\right).$$

As shown in the text, the term $\left[\sum_{i=1}^{n_{cA}}\left(x_{icA} - \bar{x}_A\right)^2 - \sum_{i=1}^{n_{cB}}\left(x_{icB} - \bar{x}_B\right)^2\right] \Big/ N\left(\sigma_A^2 - \sigma_B^2\right)$ is the proportion of the total difference in the variances that is due to cause $c$, so dividing the above expression by $N\left(\sigma_A^2 - \sigma_B^2\right)$ gives the cause-specific spread, allocation,

timing, and joint effects as a proportion of the total difference in the variances. Thus for the *c*th cause,

$$\left[\sum_{i=1}^{n_{cA}}\left(x_{icA}-\bar{x}_A\right)^2-\sum_{i=1}^{n_{cB}}\left(x_{icB}-\bar{x}_B\right)^2\right]\Big/N\left(\sigma_A^2-\sigma_B^2\right)=$$

$$\left(\sigma_{cA}^2-\sigma_{cB}^2\right)n_{cB}\Big/N\left(\sigma_A^2-\sigma_B^2\right)+\left(n_{cA}-n_{cB}\right)\left(\sigma_{cB}^2-\Delta\bar{x}_{cB}^2\right)\Big/N\left(\sigma_A^2-\sigma_B^2\right)$$

$$+\left(\Delta\bar{x}_{cA}^2-\Delta\bar{x}_{cB}^2\right)n_{cB}\Big/N\left(\sigma_{cA}^2-\sigma_{cB}^2\right)+\left(n_{cA}-n_{cB}\right)\left[\left(\sigma_{cA}^2-\sigma_{cB}^2\right)+\left(\Delta\bar{x}_{cA}^2-\Delta\bar{x}_{cB}^2\right)\right]\Big/N\left(\sigma_{cA}^2-\sigma_{cB}^2\right).$$

This completes the proof because the separate contributions of the $c = 1, 2, \ldots, C$ causes of death sum exactly to the difference in the variance of the age distribution of life table deaths between groups *A* and *B* (Eqs. (1)–(3) of text), and the final equation above shows that the cause-specific S-A-T components sum exactly to the contribution *of each cause* to the difference in the variance of the age distribution of life table deaths between groups *A* and *B*.

## References


Becker, G., Philipson, T., & Soares, R. (2005). The quantity and quality of life and the evolution of world inequality. *American Economic Review, 95,* 277–291.

Beltrán-Sánchez, H., Preston, S. H., & Canudas-Romo, V. (2008). An integrated approach to cause-of-death analysis: Cause-deleted life tables and decompositions of life expectancy. *Demographic Research, 19* (article 35), 1323–1350. doi:10.4054/DemRes.2008.19.35

Chevan, A., & Sutherland, M. (2009). Revisiting Das Gupta: Refinement and extension of standardization and decomposition. *Demography, 46,* 429–449.

Chiang, C. (1991). Competing risks in mortality analysis. *Annual Review of Public Health, 12*(1), 281–307.

Crimmins, E. M., Preston, S. H., & Cohen, B. (Eds.). (2010). *International differences in mortality at older ages: Dimensions and sources* (report of the National Research Council of the National Academies). Washington, DC: The National Academic Press.

Edwards, R. D., & Tuljapurkar, S. (2005). Inequality in life spans and a new perspective on mortality convergence across industrialized countries. *Population and Development Review, 31,* 645–674.

Firebaugh, G. (2003). *The new geography of global income inequality.* Cambridge, MA: Harvard University Press.

Gakidou, E., & King, G. (2002). Measuring total health inequality: Adding individual variation to group-level differences. *International Journal for Equity in Health, 1,* 3. doi:10.1186/1475-9276-1-3

Goesling, B., & Firebaugh, G. (2004). The trend in international health inequality. *Population and Development Review, 30,* 131–146.

Hahn, R., & Eberhardt, S. (1995). Life expectancy in four US racial/ethnic populations: 1990. *Epidemiology, 6,* 350–355.

Heron, M., Hoyert, D., Murphy, S., Xu, J., Kochanek, K., & Tejada-Vera, B. (2009). *Deaths: Final data for 2006*. National Vital Statistics Reports 57(14). Hyattsville, MD: National Center for Health Statistics.

Himes, C. L. (1994). Age patterns of mortality and cause-of-death structures in Sweden, Japan, and the United States. *Demography, 31,* 633–650.

Ho, J. Y., & Preston S. (2010 April). *U.S. mortality in an international context: Age variations*. Paper presented at the Annual Meeting of the Population Association of America, Dallas, TX.

*Human Mortality Database*. (n.d.). University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Retrieved from www.mortality.org or www.humanmortality.de (data downloaded on 01/15/2009).

Kallan, J. (1997). Effects of sociodemographic variables on adult mortality in the United States: Comparison by sex, age, and cause of death. *Social Biology, 44,* 136–147.

Kannisto, V. (2000). Measuring the compression of mortality. *Demographic Research, 3*(article 6), 1–24. doi:10.4054/DemRes.2000.3.6

Kitagawa, E. (1964). Standardized comparisons in population research. *Demography, 1,* 296–315.

Kitagawa, E., & Hauser, P. (1973). *Differential mortality in the United States: A study in socioeconomic epidemiology.* Cambridge, MA: Harvard University Press.

Lantz, P., House, J., Lepkowski, J., Williams, D., Mero, R., & Chen, J. (1998). Socioeconomic factors, health behaviors, and mortality: Results from a nationally representative prospective study of US adults. *Journal of the American Medical Association, 279,* 1703–1708.

Lynch, S., & Brown, J. (2001). Reconsidering mortality compression and deceleration: An alternative model of mortality rates. *Demography, 38,* 79–95.

Mackenbach, J. P., Stirbu, I., Roskam, A. J. R., Schaap, M. M., Menvielle, G., Leinsalu, M., & Kunst, A. E. (2008). Socioeconomic inequalities in health in 22 European countries. *The New England Journal of Medicine, 358,* 2468–2481.

Murray, C., Kulkarni, S., Michaud, C., Tomijima, N., Bulzacchelli, M., Iandiorio, T., & Ezzati, M. (2006). Eight Americas: Investigating mortality disparities across races, counties, and race-counties in the United States. *PLoS Medicine, 3*(9), e260. 1513–1524.

National Center for Health Statistics (NCHS) (2009). *Mortality multiple cause files*. Retrieved from http://www.cdc.gov/nchs/data_access/Vitalstatsonline.htm (Data downloaded on 12/06/2009).

Nau, C., & Beemer, J. (2004). *Pathocenosis, insult accumulation and ecological niche: Identifying mortality dynamics in a late nineteenth-century Massachusetts town* (working paper). Amherst: Social and Demographic Research Institute, University of Massachusetts Amherst.

Omran, A. (1971). The epidemiologic transition. *The Milbank Memorial Fund Quarterly, 49,* 509–538.

Peltzman, S. (2009). Mortality inequality. *Journal of Economic Perspectives, 23,* 175–190.

Rogers, R., & Crimmins, E. (2011). *International handbook of adult mortality.* New York: Springer.

Shkolnikov, V. (2004). Introduction to the special collection of papers on determinants of diverging trends in mortality. *Demographic Research, Special Collection, 2*(article 1), 1–10. doi:10.4054/DemRes.2004.S2.1

Shkolnikov, V., Andreev, E., & Begun, A. (2003). Gini coefficient as a life table function: Computations from discrete data, decomposition of differences and empirical examples. *Demographic Research, 8* (article 11), 305–358. doi:10.4054/DemRes.2003.8.11

Smith, G. D., Shipley, M. J., & Rose, G. (1990). Magnitude and causes of socioeconomic differentials in mortality: Further evidence from the Whitehall study. *Journal of Epidemiology and Community Health, 44,* 265–270.

Smits, J., & Monden, C. (2009). Length of life inequality around the globe. *Social Science & Medicine, 68,* 1114–1123.

Tuljapurkar, S., & Edwards, R. D. (2011). Variance in death and its implications for modeling and forecasting mortality. *Demographic Research, 24*(article 21), 497–526. doi:10.4054/DemRes.2011.24.21

van Raalte, A., Kunst A., & Mackenbach, J. P. (2010 April). *How do rising educational levels affect individual lifespan variation? Estonia and Lithuania examined over the 1990s*. Paper presented at the annual meeting of the Population Association of America, Dallas, TX.

White, K. M. (2002). Longevity advances in high-income countries 1955–1996. *Population and Development Review, 28,* 59–76.

Wilmoth, J., & Horiuchi, S. (1999). Rectangularization revisited: Variability of age at death within human populations. *Demography, 36,* 475–495.

Wilson, C. (2001). On the scale of global demographic convergence 1950–2000. *Population and Development Review, 27,* 155–171.

Wong, M., Shapiro, M., Boscardin, W., & Ettner, S. (2002). Contribution of major diseases to disparities in mortality. *The New England Journal of Medicine, 347,* 1585–1592.

World Health Organization. (2010). *WHO mortality database documentation*. Retrieved from http://www.who.int/whosis/mort/download/en/index.html

*World Health Organization Mortality Database*. (n.d.). World Health Organization. Retrieved from http://www.who.int/whosis/mort/download/en/index.html (Data downloaded on 7/9/2010).