Is life expectancy really falling for groups of low socio-economic status? Lagged selection bias and artefactual trends in mortality

Recent public health studies made headlines, reporting that for some subpopulations in the USA, mortality rates have been higher and life expectancies lower for recent compared with earlier time periods. These patterns have been described in both popular and academic discourse as a ‘rise’ in mortality or a ‘decline’ in life expectancy. We suggest that it is long past time to admit an alternative—and arguably more plausible—interpretation of these patterns. Imagine if researchers measured the average temperature for the whole of the USA a decade ago, and then for only Alaska this year, and found the former number to be lower than the latter. Would it be appropriate to say that average temperatures had ‘declined’ over the decade? We argue that it would not, and that it is likewise not appropriate to be describing many of the observed differences in subgroup life expectancy or mortality as ‘trends’. Nevertheless, scholars and journalists alike have quickly adopted this ‘trend’ conclusion and given short shrift to an alternative interpretation that we find far more plausible. Here, we make the case that this alternative explanation should be the subject of serious empirical investigation and discussion in scholarly and public discourse, rather than getting the treatment it gets now—relegated to the ‘limitations’ sections of academic papers. We take seriously the reality of health disparities in the USA and the research and policy attention that they demand; understanding and redressing these disparities requires that trends be accurately characterized.

Lagged selection bias

Our concern is that stable differences between noncomparable subgroups are being mistaken for time trends in a broader group—a phenomenon we term lagged selection bias (LSB). We use the term ‘selection bias’ in a spirit that is common in the population sciences, although in epidemiology the phenomenon might also be understood by some as a form of collider bias and by others as a form of confounding. Regardless of how it is labelled, LSB generally occurs when: (i) there is a temporal lag between when individuals are selected into their exposure group and when the outcome manifests; and (ii) the dynamics of selection into the exposure group were changing over the period spanned by the lag. Due to this temporal lag, the implications of changing selection dynamics only become manifest long after the changes have happened. When these conditions apply, it is impossible to tell from the data alone whether contemporary differences in the outcome indicate contemporary trends (like increasing social exclusion of high school non-completers), or changes from a long time ago in the exposure selection process, or some complicated combination of these two. A proper interpretation, therefore, must be explicitly based on the history of the selection process itself.

To take a very specific example, we focus on recent provocative papers that have analysed life expectancy or age-standardized mortality over time, stratified by high school completion. These papers have explicitly interpreted differences as ‘trends’. Other papers have built on this ‘trend’ interpretation, and jumped to conclusions about health impacts of contemporaneous trends in employment, health behaviours or social policies.

These ‘trend’ interpretations gloss over important differences between demographic constructs like period age-standardized mortality rates or life expectancy and real-world outcomes like the risk of death or expected length of life. They also give short shrift to two critical facts of the social history of the USA:

i. Lag between exposures and outcome. High school completion in the USA has historically been determined by the time a person is in his or her early 20s, long before the ages when most mortality is observed.
ii. Secular change in the dynamics of selection into the exposure group. Access to high school increased dramatically over the 20th century in the USA (Figure 1). The average White girl born at the end of World War I stood about a 50% chance of finishing high school by her 20th birthday; born at the end of the Lyndon Johnson administration, she had a 90% chance. Much of this expansion in access was driven by changes in the opportunity costs facing working class families who wanted to send their children to high school;15–17 as access expanded, it likely became more equitable.

When one computes mortality rates or life expectancy for the year 1990 for White female high school non-completers, one ascribes to those reaching age 70 the risk of dying that was faced by a woman who was excluded from high school around the end of the Great Depression—a normative experience for her time. By contrast, the same exercise for the year 2010 ascribes mortality risk at age 70 based on the experience of a woman excluded from high school in the late 1950s/early 1960s—which means she was left behind during a period of unprecedented expansion in access to secondary education.

The high school completion status of those dying in a given year is like light from a distant star—it reflects social conditions that prevailed decades before. In terms of mortality risk, those excluded from high school in the early part of the 20th century are not comparable with those excluded from high school a generation later, because those left behind by the high school expansions in mid century likely had childhoods that were more disadvantaged along many dimensions, and so were at higher mortality risk all along. Life expectancy among high school non-completers for the year 1990 will largely be determined by the mortality experience of the relatively lower-risk subgroup; for the year 2010, it will be determined by the mortality experience of the higher-risk subgroup. Describing differences between these two subgroups as a ‘decline’ in the life expectancy of high school non-completers simply because they were born at different times almost certainly reflects LSB.

In recent discourse, others have acknowledged the problem of LSB while retaining the ‘trend’ interpretation by noting that some or all of the observed ‘declines in life expectancy’ among high school non-completers may have been ‘caused’ or ‘explained’ by the fact that the non-completer group became ‘increasingly disadvantaged’ over the period of analysis.18 The words ‘decline’, ‘cause’ and ‘increasingly’ in this context are simultaneously true and misleading, in the same way that it would be simultaneously true and misleading to compare average temperatures in Alaska this year with the whole of the USA a decade ago, and then add a phrase acknowledging that the observed ‘decline’ in temperatures may have been ‘caused’ by the fact that the land area under consideration grew ‘increasingly’ distant from the Equator.

Figure 1. High School Completion Prevalence: White women, by birth cohort.

Notes: Data source: U.S. Census, Integrated Public Use Microsamples, 1% samples for 1940, 1950, 1960, 1970 censuses, and 5% samples for 1980, 1990, and 2000. Dotted lines indicate birth cohorts for which census data is not available; we used assumptions to fill in those probabilities.

Lagged selection bias in action: a specific example

Lagged selection bias is always a concern that should be in the background when two different period demographic constructs are compared; but, owing to the historical specifics in this case, it is likely to be especially important. To illustrate the dynamics more clearly, we have used US population data and forecasts to recreate the mortality experience of 141 birth cohorts of women from 1880 to 2020. The overall risk of dying at each age for women in each cohort is based on actual and forecasted data from the US Social Security Administration.19 We also capture the social gradients around these average risks by randomly assigning each simulated person to one of 1001 early life socioeconomic status (EL-SES) categories and assigning those who are more (less) disadvantaged in terms of EL-SES to a higher (lower) than average risk of dying at any age. These disparities are based on patterns observed in the USA.20,21 We also incorporate an EL-SES gradient in access to high school reflecting historical patterns.16 In cohorts for whom high school completion is rare, it is only the most advantaged young adults who achieve it; as access expands, it also becomes more equitable. Our example rules out an increase in actual mortality risk for anyone—every individual stands a lower chance of dying at every
age if they are born later, than they would have stood if they had been born earlier but had exactly the same characteristics. For example, in our simulation the average person in the most disadvantaged quintile in 1990 dies at age 69.5 years and in 2010 at age 72.2 years (Figure 2).

One key point of this exercise is to illustrate how a universal increase in real-world longevity like the one shown in Figure 2 does not preclude a ‘decline’ in period life expectancy and parallel ‘increase’ in period mortality. This is illustrated in Figures 3 and 4. Life expectancies at age 25 for the bottom quintile of the EL-SES distribution as well as for high school non-completers are shown in Figure 3. When we identify disadvantage based on the stable EL-SES characteristic, period life expectancies at age 25 are 43.8 years in 1980, 44.2 years in 1990, 44.4 in 2000 and 44.7 in 2010. When we rely on people’s access to high school to identify their exposure, the corresponding values are 48 years, 47.3 years, 45.3 years, and 43.4 years. Age-standardized mortality rates, shown in Figure 4, move in parallel with period life expectancies; this reflects the fact that the two demographic constructs are computed in very similar ways.

Even though, as Figure 2 illustrates, mortality disparities are actually large and stable in this example, Figures 3 and 4 depict facts from the same simulated history and seem to indicate that disparities are changing. This difference arises because of a policy success—namely, improvements in equity of access to education over the 20th century. As a result of that policy success, people who were already vulnerable to shorter lifespan because of conditions in their early lives nonetheless had access to high school, whereas those exposed to similar conditions from an earlier birth cohort had no such access. As a result, averages for high school non-completers in earlier years include the less vulnerable, as well as the very vulnerable; in the later years, the less vulnerable were simply reclassified to the high-school completer group, leaving behind only the very vulnerable. It is that reclassification that drives the artefactual ‘decline’ in life expectancy a half-century later, not any change in anyone’s actual risk of dying.

More than simply illustrating the important distinction between real-world health outcomes like longevity and demographic constructs like period life expectancy, the exercise underlying Figures 2–4 helps pin down the timing when we would expect to start seeing artefactual ‘trends’ in the demographic constructs, given the history of educational expansion in the USA over the 20th century. The exercise suggests that, if health disparities have been large and stable over the past century, and given trends in overall mortality risk and educational expansion over the past century, one would expect to see an artefactual ‘rise’ in period age-standardized mortality rates among the least educated, starting in the 1990s and continuing for about a generation. This matches the patterns that have been reported.

This example further allows us highlight two important forms of harm that LSB can do to policy.

- In this example, trends in disparities are mischaracterized. Claims that these patterns ‘indicate that increasing high
school graduation rates and redesigning work-family policies may improve longevity are inappropriate, in that they encourage advocates and policy makers to investigate the wrong time span in search of social or policy changes to be remedied. The period represented in Figures 3 and 4 might look like an age of worsening disparities, when in fact it is marked by the same large, stable disparities that have prevailed for generations.

• Perhaps even more troubling, as these dynamics play out, if analysts were to continue to stratify on high school completion, disparities would appear to solve themselves. In our simulated example, life expectancy is about 2 years shorter using high school non-completer mortality risks from 1999 compared with 1990; by the 2010–20 period, this ‘rate of decline’ is only half as steep. Even if underlying health disparities remain unaddressed as the cohorts that lived through the mid-20th century die out, the ‘light from a distant star’ will dim. This ‘dimming light’ reflects only the stabilization of schooling access in the last quarter of the 20th century; it should not be interpreted as an indication of improvements in health equity in the first quarter of the 21st century.

Conclusions and recommendations

In some circumstances where there is high risk of LSB, analysts have recognized the problem and tried to manage it. Meara et al.10 for example, aim to neutralize the problem by randomly assigning some individuals with more than high school education to the less than high school category, such that the fraction of the population labeled disadvantaged remains constant for all years where they compute life expectancies.26 We support this kind of approach, but as our discussion has highlighted, the risk of LSB arises not simply from changes in social patterning in the period when life expectancy is computed, but rather from changes over time that happened about a half-century before.

In order to neutralize the bias, it is not necessarily sufficient to equalize the prevalence of disadvantage from one period to the next in the analysis. Rather, one must ensure that the standard applied, for example to the cohort born in 1920, when analysing 1990 mortality is substantively the same as the one applied to the cohort born in 1940 when analysing 2010 mortality. This is a very difficult adjustment to make properly, in part because it requires assumptions about the very thing the analysis is meant to estimate—the relationship between disadvantage and mortality. This challenge relates closely to the long and controversial literature in demography on how to disentangle age, period and cohort effects.27–30 It also applies to health inequality measures such as the Slope Index of Inequality (SII), which aim to address changes in the distribution of socioeconomic indicators like educational attainment by average age-specific mortality risks over time for individual US counties and found ‘increases’ in that risk for some counties. Kindig et al. remark, ‘Though we are accustomed to seeing varying rates of mortality reduction in states and nations, it is striking and discouraging to find female mortality rates on the rise in 42.8 percent of US counties, despite increasing medical care expenditures and public health efforts’. Migration between urban and rural areas has ebbed and flowed—in both directions—over the past 60 years, with different social patterning for different birth cohorts.22 Departing from a rural county to move to the urban core in the late 1970s and early 1980s, for those born in the post 1955 baby boomer cohorts, may have been a sign of socioeconomic advantage; in the cohorts before and after, the opposite may be true.23–25 In that sense, comparing mortality risks among those who were born in rural areas in the 1930s and aged in place with risks among those who had the same experience but were born 20 years later may generate an artefactual ‘trend’. What it means to age in place would have been different for these different cohorts.

Lagged selection bias in other applications

There are many other contexts where LSB can generate artefactual ‘trends’. For example, recent work has compared

Figure 4. Simulated Age-Standardized Mortality among disadvantaged subgroups.

Notes: We decomposed the population of our simulation example in two different ways—once, by early life socioeconomic status (EL-SES), which in our example has a stable relationship with mortality risk, and once by high school completion status, the meaning of which changed as access to high school expanded over the mid-20th century, and which in our example has no independent relationship with mortality risk. Age-standardized period mortality rate is computed in the standard way—which is, by taking a weighted average of age-specific mortality rates, where weights are determined by a “standard” age distribution, which is held constant in every year for each subgroup. In this case, the “standard” distribution is the age distribution of the full 2010 population.
re-weighting social groups according to their population share.  

Ezzati et al. recognize that the population composition of counties has changed over time; they try to neutralize this effect by averaging computed life expectancy for a county together with life expectancies in the counties from which recent in-migrants arrived. This approach can address some important sources of migration-driven selection bias—especially for places that tend to receive people of relatively high mortality risk (e.g. places with tertiary care facilities or hospice services). It does not, however, help at all with the problem of lagged selection bias. In order to use an approach like theirs to manage LSB, it would be necessary to find those who moved away many decades ago, and aggregate their experience back with their county of origin.

These approaches are laudable steps toward proper accounting for the risk of LSB, but none is sufficient—in part, because none takes proper account of the social history of selection into the exposure groups. Nonetheless, public health research has largely ignored this important empirical issue. In the case of education especially, it is premature to close the case with a claim that ‘research has found that compositional changes contributed little to diverging trends in well-being and mortality’; rather, we agree that ‘more research is needed’ on this question. This need is urgent, given the field’s rush to judgment on the meaning of observed ‘trends’.

In the meantime, what can scholars interested in mortality trends do? We offer three recommendations:

i. Whenever possible, identify trends by tracking the same population. For example, Preston and Wang transformed data on mortality and smoking in order to track individual birth cohorts of men and women as they aged. In doing so, they were able to identify and compare real sex-specific trends in mortality, thereby shedding light on the long-term effects of smoking in early life.

ii. When different populations must be compared, stratify on characteristics that have been stable for at least a human lifespan, or that are not effectively fixed at an early point in the life course. It is methodologically possible to compute measures of life expectancy over time for subgroups like high school non-completers or residents of small geographical areas in the same way that one might compute them for men or women. However, when doing so, it is nearly impossible to interpret observed differences.

iii. Stratify with explicit reference to social history and with a clear, articulated grounding in both social and biological science. When in the life course were the stratification criteria most likely to have been determined? How does that translate to calendar time, and to relevant social history? Did the timing or social patterning shift across birth cohorts? Given what is known about the biology of the outcome, how might this affect estimation results, and in what calendar years would these effects be observed? Drawing on what is known in biology and social science, it may be possible to evaluate to what extent we might be observing ‘light from a distant star’—i.e. the long reach of social conditions that prevailed decades previously. If the relevant social and demographic dynamics are well documented (as in the case of education), then a simulation along the lines that we have used here to generate our example might help to shed light on the extent of the risk of LSB.

In general, our objective has been to urge caution in interpretation of analyses that are at high risk of LSB. If an outcome is measured on non-comparable populations, then the mere fact that the measures are computed at two different times does not make the difference between them a ‘trend’, and the blithe misuse of words like ‘fall’ or ‘decline’ is likely to do more harm than good for public health.

Jennifer B Dowd, Department of Epidemiology and Biostatistics, CUNY School of Public Health, Hunter College, 10010 NY, USA.
E-mail: jdowd@hunter.cuny.edu, and

Amar Hamoudi, Sanford School of Public Policy, Duke University, Durham, NC, USA.

References
6. Kindig DA, Cheng ER. Even as mortality fell in most US Counties, female mortality nonetheless rose in 42.8 percent


