Mortality and Longevity

Aging and Retirement

Article from

2020 Living to 100 Call for Essays

Mortality Improvements at Advanced Ages in the U.S.: Progress Masked by Age Misreporting

April 2020

Contents

Mortality Improvements at Advanced Ages in the U.S.: Progress Masked by Age Misreporting[1]

Kirill Andreev

ABSTRACT

Adjusted rates of mortality improvement for the United States have been computed for the period 1950–2014 by correcting death rates in the 1950s for age misreporting. The results suggest that conventionally constructed rates of mortality improvement based on direct estimates of U.S. death rates are biased downward, understating progress made against U.S. mortality over the last decades. The adjusted U.S. rates of mortality improvement were found comparable with the improvement rates observed in other high-longevity countries.

INTRODUCTION

There is circumstantial, yet strong, evidence that U.S. data on deaths at older ages and especially for earlier decades are severely affected by age misreporting (Coale and Kisker 1990; Elo and Preston 1994). There is also strong evidence that the quality of U.S. data significantly improved over time (Andreev, Gu and Dupre 2017). In such circumstances, directly computed rates of mortality improvement would be underestimated because death rates in earlier years, a starting point for computing mortality improvement rates, are understated because of age misreporting. Producing adjusted estimates of rates of mortality improvement thus requires adjusting direct estimates of death rates for bias introduced by age misreporting.

We applied two methods to adjust U.S. death rates in the 1950s. The first one was developed by Elo and Preston (1994) for producing adjusted estimates of African-American mortality. The second, which is new, is based on constrained extrapolation by the gamma-Makeham model and available information on mortality of supercentenarians.[2] The adjusted mortality rates are used subsequently for computing adjusted mortality improvement rates for United States for the period 1950–2014.[3] Finally, comparisons have been performed with mortality improvement rates in Denmark, England and Wales, Japan and Sweden for the same period of time.[4]

DIRECT ESTIMATES OF U.S. DEATH RATES

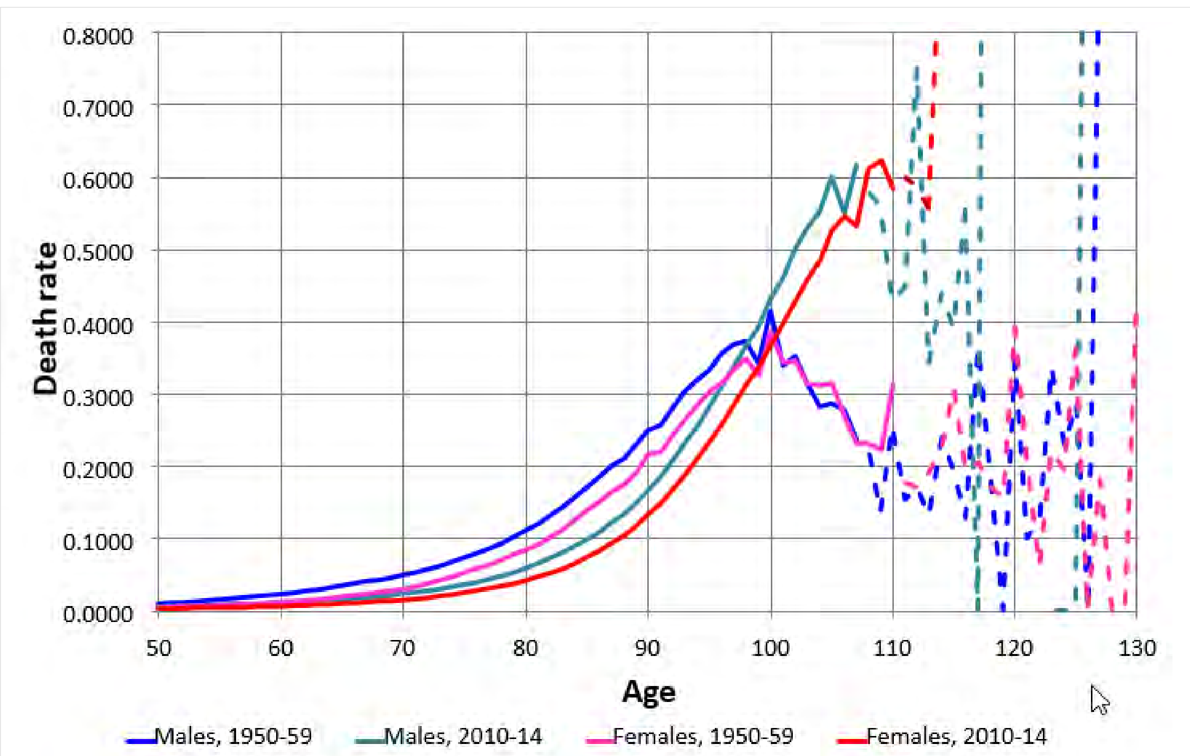
Direct mortality estimates for the United States have been produced by applying the method protocol used for construction of annual life tables for the Human Mortality Database (HMD) (Wilmoth et al. 2007). The HMD method protocol produces direct estimates of death rates and life tables in the conventional way, by estimating mortality by relating deaths in the numerator to estimates of the population at risk in the denominator. At ages 80 and older, it relies on population estimates computed by the almost-extinct cohort method rather than on official population estimates published by national statistical offices (Vincent 1951; Andreev et al. 2003). For closing life tables and for smoothing death rates at the highest ages, about age 95 and above, the Kannisto mortality model is used. For the

1 The views expressed in the paper do not imply the expression of any opinion on the part of the United Nations Secretariat.
2 People age 110 and older.
3 Data from the National Center for Health Statistics (NCHS) mortality files, vital registration deaths and census bureau population estimates were used.
4 Data from the countries' vital registration deaths and census information were used.

purpose of analysis here, we did not use the Kannisto model for smoothing and closing life tables; rather, the analysis performed here is based on direct estimates of death rates only. HMD method protocol does not provide any provisions to correct mortality estimates for age misreporting, and, at advanced ages, death rates will be biased downward if age misreporting is present in the data (Andreev, Gu and Dupre 2017; GBD 2016 Mortality Collaborators 2017). Whether the data are considered to be reliable or not is left to the analyst to decide.

Figure 1 shows direct U.S. mortality estimates for the periods 1950–59 and 2010–14. In the 1950s, U.S. death rates increase through age 95, then level off, and even decline above age 100. The age schedule of death rates also reveals age heaping at ages 90 and, more severely, at age 100—there is a spike in death rates at age 100, with corresponding dips at adjacent ages 99 and 101. The observed age pattern of death rates provides strong evidence of severe age misreporting in the United States in the 1950s. The resulting estimates reflect age misreporting present only in the registered deaths as denominators, for the rates are computed by the almost-extinct cohort method by cumulating registered deaths over cohorts; no information from censuses (commonly with worse age reporting) is incorporated in computing death rates. In the period 2010–14, conversely, U.S. death rates are increasing, at least through age 105. This suggests that the U.S. data for this period are of a reasonably good quality and can be used for estimating mortality improvement without adjustments.[5]

Figure 1
DIRECT U.S. MORTALITY ESTIMATES IN 1950–59 AND 2010–14



Source: Computed by author with data from the National Center for Health Statistics (NCHS) mortality files, vital registration deaths and census bureau population estimates.

Note: Direct estimates of deaths (no smoothing, adjustments or extrapolations). Death rates based on less than 100 deaths are shown with dotted lines.

5 For data quality analysis for the U.S. death rates at advanced ages by sex and by state, see also Andreev, Gu and Dupre (2017).

APPROACHES FOR ADJUSTING U.S. DEATH RATES FOR AGE MISREPORTING

How to adjust estimates of death rates for age misreporting at advanced ages? One of the widely used approaches for adjusting death rates for age misreporting is extrapolation of observed death rates at younger ages by means of a mortality model. The death rates at younger ages are assumed to be correct and the selected mortality model is assumed to produce a good extrapolation based on information available. This approach, for example, was used by Coale and Kisker (1990) for adjusting U.S. death rates for the year 1980. They proposed to accept direct estimates of death rates for age 65 and extrapolated death rates starting with age 65 by a quadratic model of mortality, constrained to mortality level at age 110. For males, level of mortality at age 110 was assumed to be 1 (q_x = 0.632), and for females, to avoid a crossover, it was selected to be 0.8 (q_x = 0.551). Their choice of mortality levels at age 110 is arbitrary as there is no rationale provided in the article for selecting such levels of mortality for age 110. This procedure relies only on death rates observed for ages below 65; no information is used at ages above it.

The approach of Elo and Preston (1994) is based on the assumption that the aggregate mortality rate for age *x+* and above, *m(x+)*, and age-specific population growth rates above *x* are correct and not affected by age misreporting. By selecting a model life table or mortality model, adjusted mortality estimates could be produced that are in agreement with the observed data on *m(x+)* and observed age-specific growth rates. This method uses additional information above age *x* but it also makes more assumptions about the accuracy of the observed data.

Both approaches require selecting an age below which the directly computed death rates are deemed to be accurate. This is not a trivial task as age misreporting is not observed. Elo and Preston (1994), for example, suggest that for earlier years, the directly computed African-American death rates may be seriously flawed as early as age 50. Undoubtably, one of the reasons for poor quality of the U.S. data on age of death is due to the fact that the birth registration area had not been completed until 1933 (U.S. Department of Health, Education and Welfare 1954). Age reporting for individuals with birth certificates is more accurate, leading to more accurate estimates of age patterns of mortality (Anderson 1999; Rosenwaike and Hill 1995; Kannisto 1988).

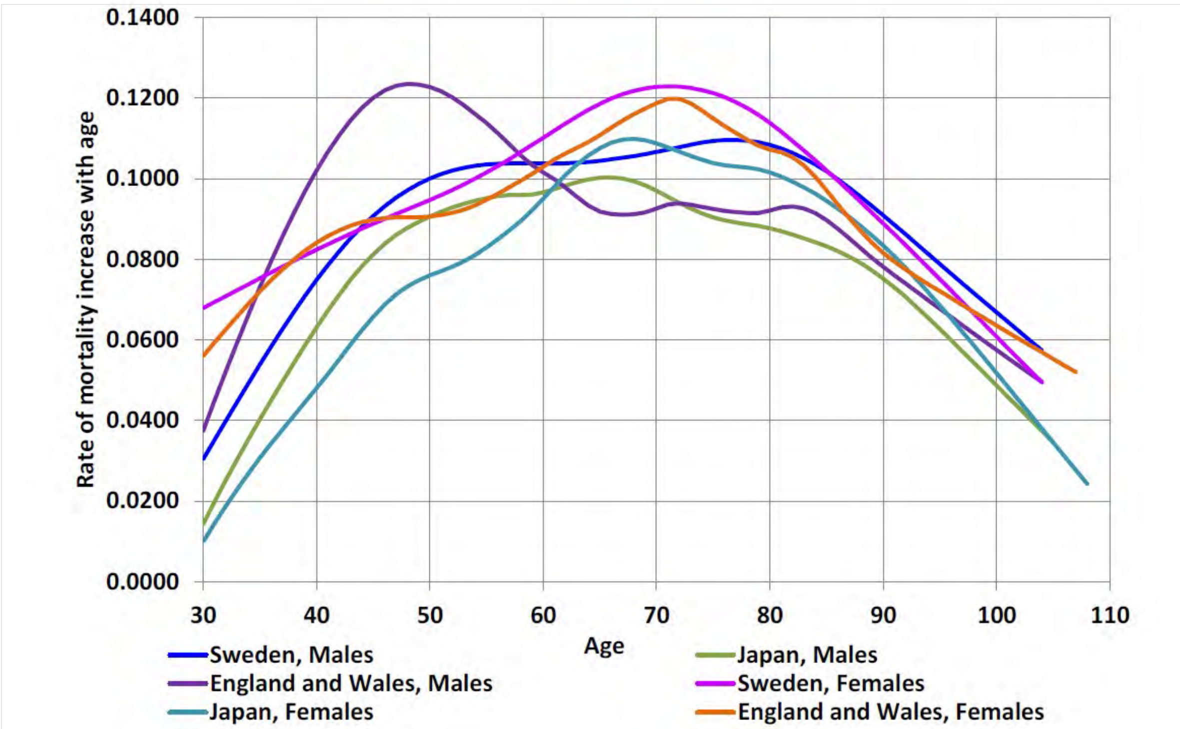
An individual born in 1933, for example, is only age 17 in 1950 and 27 in 1960, so the assumption that death rates are reliable only below age 50 in 1950–59 is not an unreasonable one. It would probably be too pessimistic to assume that we can rely on death rates only below age 20 or so, as expanding the birth registration system to the entire territory of the United States took several decades, and by 1933 most of the United States was already covered by birth registration. On the other hand, it would be reasonable to assume that the probability of possession of a birth certificate declines with age, and, in consequence, the quality of age reporting deteriorates with age. The highest ages are expected to be affected most as people who survive to such ages were born a long time ago.

By similar reasoning, we expect that the quality of age reporting on death certificates is improving over time, as more people who reach advanced ages will possess a birth certificate. As an individual born in 1933 is age 81 in 2014, and a person age 50 in 1950 is 114 in 2014, it would be reasonable to expect that death rates could be reliably estimated for all ages in 2010–14. Age schedule of death rates for the period 2010–14 shown in Figure 1 is consistent with this expectation.

CHOOSING MORTALITY MODEL FOR EXTRAPOLATION

A model for extrapolation of death rates is expected to reproduce an age pattern of mortality reasonably well starting as early as age 50. For selecting a model, we proceeded with exploration of empirical patterns of rates of mortality increase over age, $k_x = ln(m_{x+1}/m_x)$, observed for the period 1950–59 (Figure 2). For males, in general, k_x increases up to about age 50; it stays approximately constant until, at age about 85, k_x starts to decline. Speaking of death rates, male death rates are increasing at an accelerating rate until about age 50, then they continue to increase at an approximately constant rate.[6] After approximately age 85, male death rates are still increasing but at a progressively decelerating rate. The rate of increase over age k_x drops from about 10% at age 80 to about 6% at age 100. For females for this period, the pattern of k_x is somewhat different—k_x keeps increasing until about age 70 and then, virtually immediately, it starts to decline. There is a very short age interval in female populations (if any at all), around age 70, with constant k_x, where death rates are increasing at a constant rate, the age interval where the Gompertz model is applicable.

Figure 2
EMPIRICAL PATTERNS OF RATES OF MORTALITY INCREASE WITH AGE, $k_x = ln(m_{x+1}/m_x)$, 1950–59



Source: Computed by author with data from the countries' vital registration deaths and census information.

The models of mortality developed for older ages, for example, the Gompertz, gamma-Gompertz, Kannisto and quadratic models (Thatcher, Kannisto and Vaupel 1998), are able to capture a plateau at middle ages and decline in k_x but they are not able to capture an initial increase in k_x. The Gompertz model, for example, obviously, is not able to reproduce observed patterns in Figure 2 as k_x in this model is simply constant and independent of age. The Gompertz-Makeham model is flexible enough to reproduce an initial increase in k_x and the following plateau but it is not able to capture decline in k_x starting age 85. The gamma-Gompertz model, on the other hand, is able to fit the plateau in k_x and decline at the highest ages but not flexible enough to reproduce initial increase in k_x. The simplest model, perhaps, the one we settled on for the extrapolation task at hand, the one that is flexible enough to capture the bell-shaped pattern of k_x in Figure 2, is the gamma-Makeham model:

$$\mu(x) = \frac{ae^{bx}}{1 + \sigma^2 \frac{a}{b}(e^{bx} - 1)} + c.$$

6 Age interval with constant k_x is the interval where the Gompertz model of mortality is applicable.

We proceeded with testing extrapolative properties of the gamma-Makeham model on Swedish data, period 1950–59. The results turned out to be promising but unstable. On several occasions, extrapolation based on fit to death rates at ages 40–49 and, especially, 35–84, overshot the empirical death rates at ages 80 and older. To further ensure the robustness of extrapolation, additional constraints on parameter space were introduced. First, similar to Coale and Kisker (1990), it was decided to constrain level of mortality at age 110. Estimates by Gampe (2010) based on validated cases of supercentenarians suggest that mortality reaches a plateau at 110 with a level of about 0.7, and with no differences between males and females. Barbi et al. (2018), on analyzing data on Italian centenarians, reached a conclusion that human mortality reaches a plateau at age 105 with $m_x = 0.645$ for females and 0.678 for males. Barbi's results are generally consistent with mortality levels estimated by Gampe (2010). It was decided then to constrain the level of mortality at age 110 to be higher than 0.7, $m(110) > 0.7$, for both sexes. As shown with further experimentations with the gamma-Makeham model, an additional constraint on parameter space is still needed to ensure the robustness of extrapolation—a constraint on the slope of the mortality curve at age 110. It was further assumed that the rate of mortality increase with age must be less than 3% at ages 110 or above, $k(110) < 0.03$, more or less in line with the estimates of k_x given in Figure 2. We refer to this extrapolation procedure hereafter as the constrained gamma-Makeham model.

ELO-PRESTON APPROACH

A second set of adjusted U.S. mortality estimates for the 1950s was produced by a method from Elo and Preston (1994) slightly generalized to work with data by single years of age. The method relies on the assumption that observed U.S. population growth rates above age 50 are correct. For the age pattern of mortality required by this method, we used the extended Coale-Demeny North model life tables covering higher levels of life expectancy and wider age intervals.[7] The assumption of correct age-specific population growth rates might be a questionable one, but, to the best of our knowledge, there is no research available on how age misreporting might affect age-specific growth rates.

RESULTS

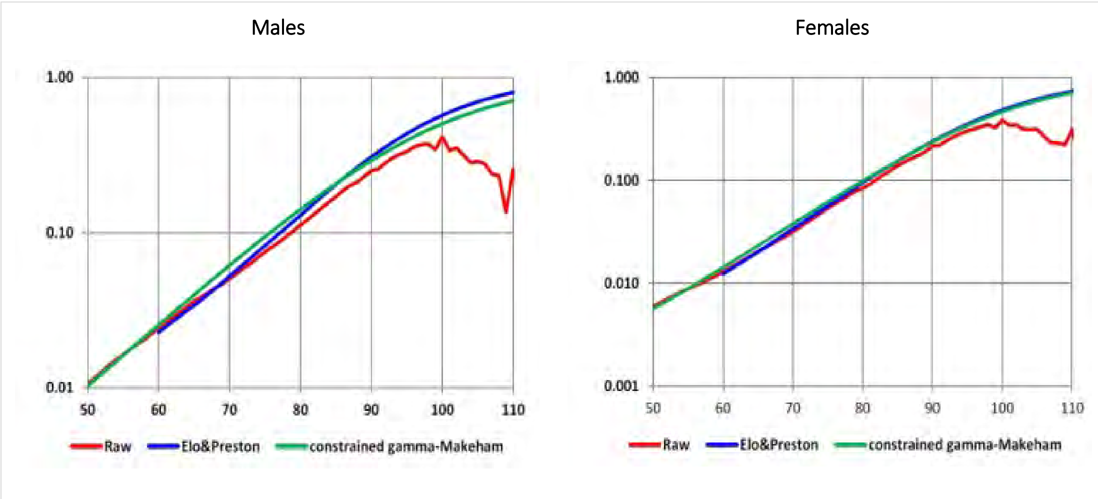
Figure 3 shows direct and adjusted U.S. mortality estimates for the period 1950–59. For both methods, the adjusted estimates are higher than the direct, with the difference between adjusted and direct rates widening progressively with age. The adjustments introduced for males are larger than for females, consistent with the commonly encountered observations that age misreporting more affects male populations than female.[8] For males, the estimates produced by the constrained gamma-Makeham model are higher than Elo-Preston estimates for males up to age 85, but lower above this age. For females, the estimates produced by both methods are quite close.

---

7 The extended model life tables have been produced by U.N. Population Division and available at *unpopulation.org*. At the higher ages, death rates in these life tables have been extrapolated by the Kannisto model of mortality to age 110, based on original death rates in the Coale-Demeny tables, which used the Gompertz model to smooth and close life tables at older ages. They are also available for higher levels of life expectancies needed for mortality projection purposes.
8 Age misreporting usually has a larger effect on male populations resulting in slower rates of mortality increase with age, often resulting in spurious crossovers between male and schedules of death rates at older ages.

Figure 3

DIRECT AND ADJUSTED U.S. MORTALITY ESTIMATES, 1950–59

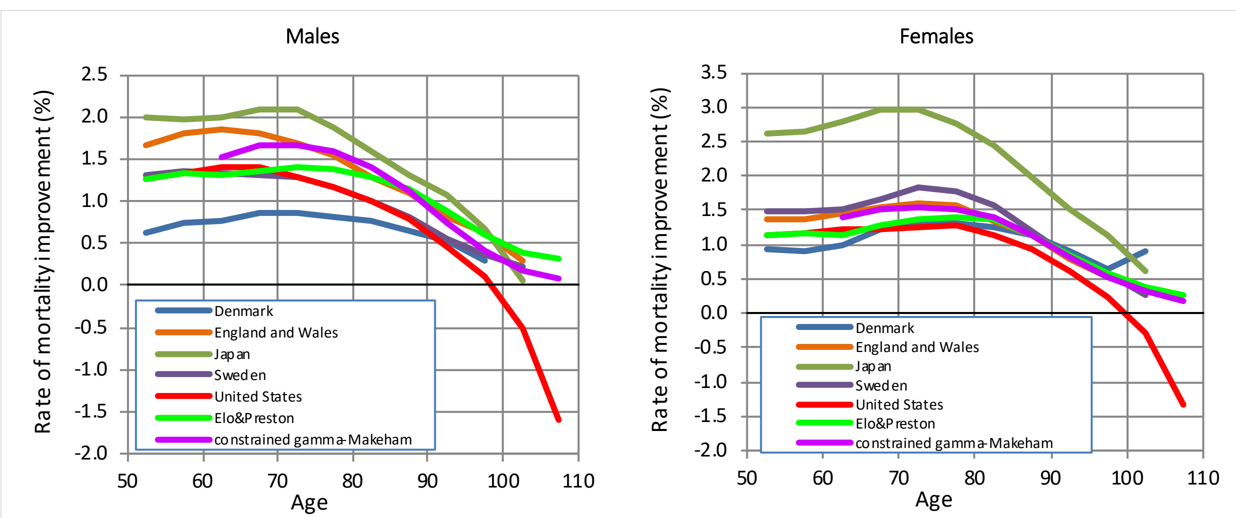


Source: Computed by author with data from the National Center for Health Statistics (NCHS) mortality files, vital registration deaths and census bureau population estimates.

Estimates of rates of mortality improvement over the period 1950–59 to 2010–14, for the United States, both direct and adjusted, and for Denmark, England and Wales, Japan and Sweden are given in Figure 4. The direct U.S. estimates are based on direct estimates of U.S. mortality in the period 1950–59 and the adjusted estimates on two sets of adjusted estimates of death rates for the period 1950–59 in Figure 3. In both cases, direct estimates of death rates in the period 2010–14 are used without adjustment. The direct estimates of rates of U.S. mortality improvement (red line, in Figure 4) are the lowest among all countries compared, for males above age 90 and for females above age 75. At the highest ages, negative values of the direct estimates of mortality improvement indicates that death rates were increasing over time. Adjustments of death rates in the 1950–59 set for age misreporting has a profound effect on the estimated rates of mortality improvement (Figure 4).

Figure 4

RATES OF MORTALITY IMPROVEMENT (%) FOR THE PERIOD FROM 1950–59 TO 2010–14



Source: Computed by author with data from the National Center for Health Statistics (NCHS) mortality files, vital registration deaths and census bureau population estimates.

Upward adjustments of mortality in the period 1950–59 result in higher estimated rates of U.S. mortality improvement for the period 1950–59 to 2010–14. For the entire age range, the adjusted U.S. rates are now positive while decline over age, approaching zero above 100, is similar to the patterns found in other countries. Compared with the unadjusted rates of mortality improvement, they are significantly higher for age 70 and above.

For males, the adjusted rates of mortality improvement are quite close to the rates observed in England and Wales, and higher than those in Sweden and Denmark. For females, the adjusted U.S. rates are close to those in England and Wales, Sweden and Denmark, and higher than Danish for ages below 80. Rates of mortality improvement in Japan are still the highest among all countries compared, especially for females.

CONCLUDING REMARKS

Adjusting U.S. death rates for the period 1950–59 for age misreporting with two methods resulted in higher estimates of rates of mortality improvement in the United States over time compared with the rates of mortality improvements based on unadjusted (direct) death rates. The U.S. adjusted mortality improvement rates turned out to be comparable to the rates of mortality improvements observed in other high-longevity countries (England and Wales, Sweden and Denmark) and even higher for some age intervals.

Quality of population data depends crucially on well-functioning systems of civil registration and statistics. For relatively recently established systems of vital registrations like the one in the United States, there is a period of time until quality of mortality data improves and reaches the level of quality in other countries with longer functioning systems of civil registration and vital statistics. Poor data quality in the earlier periods leads to underestimation of mortality levels and, consequently, to underestimations of rates of mortality improvement. In severe cases, even spurious increases in mortality could be observed that are simply manifestations of improvements in data quality over time.

Mortality projections that rely heavily on observed trends in death rates would be unnecessarily pessimistic as they are informed by underestimated rates of historical mortality improvements. Some work has been done to understand the effects of mortality adjustments on mortality projections and future growth of the elderly population (Bennett and Olshansky 1996) but no systemic approach to address this problem has been developed so far. One of the main resources on human mortality trends, the Human Mortality Database,[9] has no provisions for adjusting deaths rates at older ages for age misreporting. As a consequence, death rates for older ages in the United States for years about 2010 and before (and for many other countries) in this database should not be used for analytical and forecasting purposes as currently published.

Kirill Andreev, Ph.D., is the population affairs officer with the United Nations Department of Economic and Social Affairs, Population Division. He can be reached at *andreev@un.org*. His personal website is *www.kirillandreev.com*.

9 More about the HMD can be found at *www.mortality.org.*

REFERENCES

Anderson, R. N. 1999. Method for Constructing Complete Annual U.S. Life Tables. *Vital and Health Statistics* 2, no. 129. *https://www.cdc.gov/nchs/data/series/sr_02/sr02_129.pdf*.

Andreev, Kirill, D. Jdanov, E. Soroko and V. Shkolnikov. 2003. Methodology of Kannisto-Thatcher Database on Old Age Mortality at the Max Planck Institute for Demographic Research. *http://www.demogr.mpg.de/databases/ktdb/xservices/method.htm*.

Andreev, Kirill F., Danan Gu and Matthew E. Dupre. 2017. Regional Mortality in the United States at Ages 80 and Older: An Analysis of Direct Estimates, 1959–2011. *Living to 100 Monograph*, Society of Actuaries.

Barbi, Elisabetta, Francesco Lagona, Marco Marsili, James W. Vaupel and Kenneth W. Wachter. 2018. The Plateau of Human Mortality: Demography of Longevity Pioneers. *Science* 360, no. 6396:1459–61. *https://doi.org/10.1126/science.aat3119*.

Bennett, Neil G., and S. Jay Olshansky. 1996. Forecasting U.S. Age Structure and the Future of Social Security: The Impact of Adjustments to Official Mortality Schedules. *Population and Development Review* 22, no. 4:703–27.

Coale, Ansley J., and Ellen E. Kisker. 1990. Defects in Data on Old-Age Mortality in the United States: New Procedures for Calculating Mortality Schedules and Life Tables at the Highest Ages. *Asian and Pacific Population Forum* 4, no. 1:1–31.

Elo, Irma T., and Samuel H. Preston. 1994. Estimating African-American Mortality from Inaccurate Data. *Demography* 31, no. 3:427–58.

Gampe, Jutta. 2010. Human Mortality Beyond Age 110. In *Supercentenarians*, 219–30. New York: Springer.

GBD 2016 Mortality Collaborators. 2017. Global, Regional, and National Under-5 Mortality, Adult Mortality, Age-Specific Mortality, and Life Expectancy, 1970–2016: A Systematic Analysis for the Global Burden of Disease Study 2016. *Lancet* 390, no. 10100:1084–1150.

Kannisto, Vaino. 1988. On the Survival of Centenarians and the Span of Life. *Population Studies* 42, no. 3: 389–406.

Rosenwaike, Ira, and Hill, Mark E. 1995. The Accuracy of Age Reporting among Elderly African Americans: Evidence of a Birth Registration Effect. Working Paper 95-04. Population Studies Center. University of Pennsylvania.

Thatcher, A. Roger, Vaino Kannisto and James W. Vaupel. 1998. *The Force of Mortality at Ages 80 to 120*. Odense: Odense University Press. *http://www.demogr.mpg.de/Papers/Books/Monograph5/ForMort.htm*.

U.S. Department of Health, Education and Welfare. 1954. History and Organization of the Vital Statistics System. *Vital Statistics of the United States: 1950 1:1–19.* Reproduced as Appendix II in Hetzel, Alice M. 1997. U.S. Vital Statistics System: Major Activities and Developments, 1950–95. U.S. Department of Health and Human Services report.

Vincent, Paul. 1951. La Mortalité des Vieillards. *Population* 6, no. 2:181–204.

Wilmoth, J.R., K. Andreev, D. Jdanov and D.A. Glei. 2007. Methods Protocol for the Human Mortality Database. *http://www.mortality.org*.