Chapter 6

Mortality of Centenarians in the United States

Bert Kestenbaum

6.1 Introduction

The recent government publication (Xu 2016) of death rates for the U.S. centenarian population – persons age 100 and over – that are clearly incorrect should serve as a reminder to the research community of the folly of using data for the extreme-aged without questioning and evaluating them. It could also serve as an impetus for demographers, actuaries, and others to whom the subject is important to determine which data are best for measuring the mortality of the extreme aged and whether there are creative ways to improve them. In this essay, after describing the incorrect recent estimates of the National Center for Health Statistics' (NCHS) Mortality Branch and how they came about, I present ways to make published Medicare data more useful for measuring the mortality of the very old.

6.2 The NCHS Estimate

In January 2016 the National Center for Health Statistics' Mortality Branch published Data Brief no. 233 on the subject of annual mortality rates of the centenarian population taken as a whole, over the 15-year period from 2000 to 2014. Data Brief no. 233 features the finding that for both males and females the death rate for centenarians taken as a whole increased substantially from the year 2000 – the beginning of the observation period – to the year 2008, then decreased substantially

The views expressed here are those of the author, and no endorsement of those views by the U.S. Security Administration should be inferred.

B. Kestenbaum (✉)
Baltimore, MD, USS
e-mail: bmkestenbaum@outlook.com

through 2014, the end of the period. For females the increase was a significant 10%, from 385 per thousand in 2000 to 424 per thousand in 2008, and the decrease from 2008 to 2014 was a significant 14%. For males the increase to 2008 was a huge 41%, from 292 per thousand to 413 per thousand, and the decrease from 2008 was 20%, to 332 per thousand. It will be noted furthermore that in 2000 the male mortality rate for centenarians was about one-quarter less than the female rate – certainly very surprising. These rates are shown in Fig. 6.1.

How were the death rates for centenarians in Data Brief no. 233 computed? The numerators are counts of registered deaths at ages 100 and over (with no assessment of the accuracy of the reported ages). The denominators for the years 2000 through 2010 come from the Census Bureau's *intercensal* mid-year population estimates, and for the years after 2010 from *postcensal* estimates produced by the Bureau. These estimates are anchored by the counts of centenarians for April 1, 2000 and April 1, 2010 obtained in the decennial censuses of 50 thousand and 54 thousand, respectively – *even though the Bureau acknowledges that the decennial census significantly overcounts the centenarian population* (Meyer 2012). In particular, the Bureau itself speculated that the true number of centenarians on census day, 2000 is about one-third less than the official count of 50 thousand. It seems likely that the overcount was smaller in the 2010 decennial census, creating the illusion that the centenarian population grew little from the beginning of the decade to its end.

In fact, the Bureau's intercensal estimates of the size of the centenarian population for the 2000–2010 decade fall in the first half of the decade to 46 thousand and then turn around and rise. In sharp contrast, for the period after 2010 the Bureau's postcensal estimates imply an estimated average annual growth rate of about 10%!

Fig. 6.1 Deaths per thousand, each year, 2000–2014, Data Brief No. 233

The Bureau's acknowledgement that the decennial census counts of the centenarian population are much too high appears only in the aforementioned special study it did on centenarians. Nowhere in the Bureau's intercensal population estimates publication – from which come the denominators for the NCHS' Data Brief no. 233 – is there a warning to users that the estimated number of centenarians in 2000 is much too high. When I asked "why not?", a Bureau staff person answered that the Bureau expects a certain degree of sophistication in its readership, to recognize that small estimates are subject to large relative errors. I did not pursue the matter further to argue that while 50 thousand may be a small number in an analysis of the total U.S. population, for readers interested in the oldest-old population 50 thousand is not a small number.

While the failure of the Census Bureau to alert its audience to shortcomings of the data might excuse others who use the data "as is", it doesn't excuse the National Center for Health Statistics' Mortality Statistics Branch, the nation's authority on population mortality. Indeed, this same organization declines to use Census Bureau population denominators beyond age 94 in the preparation of the U.S. Life Tables (Arias 2014).

It is abundantly clear that the rates computed in this manner, and the time series pattern which emerges, are casualties of the incorrect census estimates, and should never have been published. I urged NCHS executives to withdraw Data Brief no. 233, given that its primary finding is not correct and that any race-ethnicity detail and cause-of-death detail that the report provides cannot be relied on.

The NCHS was unwilling to do so. Instead, on July 28, 2016 – about 6 months after initial publication – the NCHS reissued the publication with the following addendum, which, I believe, fails to convey the gravity of the errors.

"The death rates for centenarians shown in this report should be interpreted with caution. An overestimation of the centenarian population in the 2000 Census may have led to a slight underestimation of the death rates from 2000 to 2009 shown in this report. The rate of increase in the death rates for these years was also somewhat overestimated."

6.3 The Medicare Experience

Where then can one find good-quality national data on the mortality of the centenarian population in the United States, and are there creative approaches to increase the utility of these data? One possibility: *if* the reports of extreme age at death in death registration are of acceptable quality over a substantial period of years (and if net migration is negligible), the decedents can be grouped by year of birth, and an approach known as the "extinct cohort" method can be employed to compute from the death registration data alone the death probability at each age x , using as the denominator the number of deaths occurring at or after age x . However, it has not yet been demonstrated that extreme age reporting in death registration is of acceptable quality. Indeed, U.S. death probabilities at extreme ages for the 1900 birth

cohort computed exactly in this manner for the Human Mortality Database project, and shown in Table 6.1, appear to be too large.

Table 6.1 $q(x)$ for ages 95–109, males and females: various sources

Study	This study	SSA Actuary (2005)	NAAJ journal (2002)	NCHS	Human Mortality Database
Data	Medicare tabulations	Medicare tabulations	Medicare microdata	Medicare tabulations	NCHS deaths
Type of table	Cohort	Cohort (from period)	Period	Period	Cohort
Target population	Born 1893–1902	Born 1900	1990–1999 period	1999–2001 period	Born 1900
Males					
95	0.272	0.273	0.274	0.260	0.313
96	0.293	0.292	0.294	0.279	0.351
97	0.315	0.313	0.313	0.299	0.370
98	0.335	0.330	0.330	0.320	0.401
99	0.356	0.354	0.348	0.342	0.447
100	0.381	0.376	0.368	0.364	0.457
101	0.402	0.392	0.384	0.387	0.518
102	0.418	0.420	0.405	0.410	0.496
103	0.430	0.444	0.433	0.434	0.607
104	0.438	0.470	0.425	0.458	0.533
105	0.452	0.497	0.436	0.482	0.591
106	0.423	0.524	0.423	0.506	0.877
107	0.376	0.552	0.450	0.530	0.605
108	0.471	0.580	0.494	0.554	0.524
109	0.555	0.608	0.489	0.578	0.546
Females					
95	0.222	0.220	0.223	0.216	0.254
96	0.241	0.239	0.241	0.234	0.279
97	0.263	0.260	0.260	0.254	0.308
98	0.287	0.279	0.281	0.275	0.340
99	0.309	0.307	0.298	0.297	0.372
100	0.333	0.329	0.319	0.320	0.406
101	0.360	0.345	0.342	0.343	0.440
102	0.380	0.373	0.362	0.368	0.495
103	0.401	0.398	0.373	0.393	0.509
104	0.427	0.425	0.394	0.419	0.530
105	0.453	0.453	0.415	0.445	0.534
106	0.469	0.483	0.429	0.472	0.536
107	0.482	0.513	0.462	0.498	0.579
108	0.504	0.544	0.469	0.525	0.629
109	0.515	0.576	0.492	0.552	0.584

Another source of information for extreme-age mortality is the Medicare (national governmental health care program for the elderly and the disabled) experience, captured in the administrative records of both the Centers for Medicare and Medicaid Services (CMS) and the Social Security Administration (SSA). Because only a very small fraction of the very old does not participate in Medicare – probably less than 5%, the Medicare experience is a quite satisfactory representation of the U.S. experience. Enrollment and death data for the Medicare population are tabulated routinely and consistently each year by the Centers for Medicare and Medicaid Services and provided to the Social Security Administration’s Office of the Chief Actuary and the National Center for Health Statistics, both of whom use the Medicare experience in their preparation of life tables. However, both use the Medicare experience only through age 94 and close the life tables with mathematical models.¹ Published estimates of extreme-age mortality from the Social Security Administration’s Office of the Chief Actuary (Bell and Miller 2005) and the National Center for Health Statistics Mortality Statistics Branch (Arias 2014), respectively, are shown in Table 6.1.

The quality of the Medicare data at extreme ages is affected by two problems. One is unique to the Medicare experience: the inclusion in the experience of “immortals” – deceased persons whose death was either not reported or not recorded. The other is the usual problem of age misstatement – intentional or unintentional – and misrecording. These problems have the greatest effect at the oldest ages, where there is not a predominance of accurate records. Both of these problems can be mitigated somewhat by excluding more error-prone records from the experience study. For this reason, both the SSA Actuary and the NCHS life-table program use tabulations of the Medicare experience from which are excluded records of persons who are not beneficiaries of the Social Security or Railroad Retirement programs, an exclusion of less than 5% of the total experience. I use these same tabulations in the new investigation reported on here.

I took a different approach in earlier studies on the mortality of the extreme-aged based on the Medicare experience, using person-level records of the Social Security Administration on participation in and termination by death from Medicare Part B – medical insurance (Kestenbaum 1992; Kestenbaum and Ferguson 2002). Unlike participation in Medicare Part A – hospital insurance, which generally is free of charge, participation in Part B requires the payment of monthly premiums to continue enrollment – either by the insured or by a third-party payor on behalf of the insured, such as a State government for its low-income population. If premium payments are discontinued, enrollment is terminated; hence there should be few immortals in the Medicare Part B experience.

This different approach with the Medicare experience – using person-level records in the SSA’s master file – confers two advantages. One is that SSA

¹ Strictly speaking, this description applies to the (smoothed) *period* life tables produced by the SSA Actuary. The death probabilities in the *cohort* life tables produced by the SSA Actuary are taken directly from the period tables, which explains why their progression is not as smooth as the progression of the probabilities in the period table.

records – although more difficult to use – are slightly more accurate than CMS records. This is because updates such as death reports are generally received by SSA, which then shares the information with CMS; if on a particular day the transmission of information fails, the update (death) will be in SSA records but not CMS records.

More importantly, with the person-level approach it is possible to do some data cleansing. The Social Security Administration has other large person-based databases, indexed by social security number,² in which the record contains information on date of birth and/or date of death; the dates in the master file of Medicare enrollment can be compared with these other dates and adjusted for discrepancies. These other collections of data include the file of almost 1 billion applications and reapplications for a social security number, the file of nearly a half-billion earnings histories, and the file of 75 million applicants to the Supplemental Security Income (SSI) welfare program.

This approach, that is, tabulating the person-level experience in Medicare Part B as reflected in SSA records and after editing dates of birth and death by comparing with other SSA master files, was used to produce (ungraduated) estimates of male and female mortality prevailing during the decade of the 1990s by single year of age from age 85 to age 109 and for age 110+. The study was published in the *North American Actuarial Journal* (Kestenbaum and Ferguson 2002) and the published estimates, beginning with age 95, are reproduced in Table 6.1.

The implementation of this approach entails a significant investment of resources, and, furthermore, requires access to the Medicare microdata. In contrast, the new study reported here uses a series of the aforementioned tabulations of the Medicare experience of enrollees (Parts A and/or B) who are entitled to Social Security or Railroad Retirement benefits, routinely prepared by CMS and shared with the SSA Actuary and the National Center for Health Statistics. Our objective is to develop adjustments to these grouped data to enable the estimation of single-age mortality at ages beyond the ages they are currently used for, that is, at ages 95 and over.

The CMS tabulates the number of enrollees in the Medicare program alive on January 1st by gender and single year of age (last birthday), through age 110. Persons tabulated at any age, x , are on average age $(x + 1/2)$ on January 1. The most recent counts I had are for January 1, 2013, and a historical series with data which are comparable from year to year goes back to January 1, 1988.

To address the problem of the “immortals”, I combined the annual tabulations of the number of enrollees on January 1st, and rearranged the counts by year-of-birth cohort to follow the cohort until extinction.³ Under the reasonable assumption that at extreme ages the net of entrants to and exits from the Medicare population is negligible, the decrease in the size of the cohort from one tabulated age, y , on

²The social security number is the nearly universal personal identifier used in the United States.

³It is perhaps worth noting the simple difference between this and the “extinct cohort” method mentioned earlier. Namely, in the “extinct cohort” method the researcher knows the number of deaths at each age up to extinction, from which he *constructs* the number of survivors to each age. Here I already know the number of survivors at each age.

January 1st of year y to the next tabulated age, $x + 1$, on January 1st of year $y + 1$ represents deaths occurring in calendar year y among persons with tabulated age x at the beginning of the calendar year; the mortality probability, $q(x)$, is easily obtained.

Few persons survive to age 110.5. Thus the count of persons at tabulated age 110 is the sum of a very small number of true “supercentenarians”⁴ and a larger number of immortals. If we assume, for example (other assumptions will work, too), that the probability of mortality at (tabulated) age 109 is 0.5, then half of those who reached tabulated age 109 reach tabulated age 110, and the others counted at age 110 are the immortals. We then proceed to subtract the number of immortals from the tabulated numbers at age 110 and all the extreme ages.⁵

To address the problem of age misstatement and misrecording, I borrow and extrapolate results published by Rosenwaike and Stone in 2003 in the journal *Demography*. These were results from a record-check study of the accuracy of ages of 110 and over in the Medicare experience. The authors searched for these purported supercentenarians in the records of the decennial censuses of 1880 and 1900 (the records of the 1890 census were destroyed in a fire), using well-defined, replicable matching rules.

In the Rosenwaike-Stone study, when the search was successful the years of birth in the Medicare record and the early census record were identical 73% of the time. The Medicare year of birth was the earlier of the two in the large majority of the other 27% of the matched records. The average discrepancy – the census year of birth minus the Medicare year of birth – among discrepant cases was about 3 years.

If these results are to be extrapolated to extreme ages less than 110, assumptions must be made. It seems reasonable to assume that the fraction among persons with recorded age x in the Medicare experience whose recorded age is their true age, call it $r(x)$, increases as the size of the experience increases (and thus as the age decreases). I assume that $r(x)$ increases from 73% at age 110 by almost 2 percentage points per year of (decreasing) age until it reaches 100% at age 96. I also assume that the average size of the age error among persons whose recorded age is not their true age is an overstatement of 3 years, the same as in the Rosenwaike-Stone study.

Under these assumptions I estimate $q(x)$, the probability of mortality at age x , from the following approximate relationship, where q_{obs} denotes an observed probability:⁶

$$q_{obs}(x) = r(x) \cdot q(x) + (1 - r(x)) \cdot q_{obs}(x - 3).$$

⁴A supercentenarian is a person who has attained age 110.

⁵In theory it would be better to distribute the deaths of these immortals across all ages in proportion to the observed number of deaths at each age, but the additional precision which is obtained is surely trivial.

⁶The relationship would be exact with $q(x - 3)$ in the equation, in place of $q_{obs}(x - 3)$. But, of course, $q(x - 3)$ is not known.

For example, by our assumption the fraction of persons with recorded age 100 who are truly age 100, $r(100)$, is 0.93. The equation expresses the approximate relationship that the observed probability at age 100 equals 93% of the true probability at age 100 plus 7% of the observed probability at age 97.

A single year-of-birth cohort has too few survivors at extreme ages to generate reliable results. Accordingly, I combine the results for 10 adjacent birth cohorts, with years of birth from 1893 to 1902. The 1902 birth cohort is the latest cohort for which we have counts through tabulated age 110, since our latest data are for January 1st of 2013.

The results of this investigation are displayed in Table 6.1 and Figs. 6.2a and 6.2b, together with other estimates mentioned earlier: (1) from the Social Security Administration's Actuary and the NCHS' life tables, both extrapolations by mathematical modelling of the tabulated Medicare experience through age 94; (2) from my study in the North American Actuarial Journal based on edited person-level records of the Medicare Part B experience; and (3) in the Human Mortality Database, based on death records from the U.S. Vital Statistics system, synthesized by extinct-cohort methods. The ungraduated results from our new investigation are in the first column of data. While the pattern is somewhat erratic for males, the results for females, for whom there is overall three times the amount of experience that there is for males and about six times as much at the oldest ages, seem credible up through age 105 – a significant improvement over the performance of unedited tabulations. For females through age 105 the new mortality probabilities are higher than previous estimates except for the Human Mortality Database estimates.

We began with the peculiar findings in the NCHS' Data Brief no. 233 regarding the mortality rate in the centenarian population taken as a whole. Its main finding is

Fig. 6.2a $q(x)$ for ages 95–109, males: various sources

Fig. 6.2b $q(x)$ for ages 95–109, females: various sources

Table 6.2 Overall probability of mortality of centenarians, by cohort and sex: edited Medicare tabulations

Birth cohort	Males	Females	Both sexes, combined
1893	.408	.366	.373
1894	.409	.365	.372
1895	.402	.365	.371
1896	.411	.373	.378
1897	.414	.374	.379
1898	.409	.377	.379
1899	.421	.379	.385
1900	.404	.384	.387
1901	.406	.383	.386
1902	.406	.380	.384

that the overall mortality rate among centenarians first increased significantly and then decreased significantly over the most recent 15-year period, particularly for males. It also found that mortality rates for centenarians were much smaller for males than for females for part of this period. In contrast to the NCHS findings, the mortality probabilities for the centenarian population taken as a whole in the edited Medicare data for ten birth cohorts from 1893 to 1902 – shown in Table 6.2 – (a) exhibit little variation from cohort to cohort⁷ and (b) show a consistent mortality advantage for females over males.

⁷ There *appears* to be a slight increase in mortality over time for females and both sexes, combined, but this might be due to a changing age distribution of the centenarian population.

6.4 In Conclusion

We have not yet reached the point where unedited numerators and denominators for death rates or probabilities are of satisfactory quality for measuring the mortality of the extreme aged. Ignoring this reality can lead to an estimation of extreme-age mortality which is unacceptable, as is the case for the NCHS' Data Brief no. 233. On the other hand, there are steps that can be taken to improve the quality of the underlying data, so that the threshold age at which a mathematical model replaces actual data and closes out the life table is pushed further out.

References

- Arias, E. (2014). *United States Life Tables, 2010* (National vital statistics report, 63 (7)). Hyattsville, MD: National Center for Health Statistics.
- Bell, F. C., & Miller, M. L. (2005). *Life tables for the United States social security area 1900–2100* (Actuarial study no. 120). Baltimore: Social Security Administration, Office of the Actuary.
- Human Mortality Database. University of California, Berkeley (USA) and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org or www.humanmortality.de.
- Kestenbaum, B. (1992). A description of the extreme aged population based on improved medicare enrollment data. *Demography*, 29(4).
- Kestenbaum, B., & Ferguson, B. R. (2002). Mortality of the extreme aged in the United States, based on improved medicare data. *North American Actuarial Journal*, 6(3).
- Meyer, J. (2012). *Centenarians: 2010* (2010 special census report. C2010SR-03). Washington, DC.
- Rosenwaike, I., & Stone, L. F. (2003). Verification of the ages of supercentenarians in the United States: Results of a matching study. *Demography*, 40(4), 727–739.
- U.S. Bureau of the Census. *Population estimates: National intercensal estimates (2000–2010)*. Intercensal estimates of the resident population by single year of age, sex, race, and Hispanic origin for the United States: April 1, 2000 to July 1, 2010. Available from: <http://www.census.gov/popest/data/intercensal/national/nat2010.html>
- U.S. Census Bureau. *Population estimates: Postcensal estimates for 2010–2014*. Available from: <http://www.census.gov/popest/data/national/asrh/2014/index.html>
- Xu, J. Q. (2016). *Mortality among centenarians in the United States, 2000–2014* (NCHS data brief no. 233). Hyattsville: National Center for Health Statistics.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.