National Vital Statistics Reports

Volume 57, Number 4 August 27, 2008

U.S. Decennial Life Tables for 1999–2001: Methodology of the United States Life Tables

By Rong Wei, Ph.D.; Lester R. Curtin, Ph.D.; Elizabeth Arias, Ph.D.; and Robert N. Anderson, Ph.D.

Abstract

Objectives—This report describes the methodology used in the preparation of the 1999–2001 decennial life tables for the United States.

Methods—Data used to prepare these life tables include population data by age on the census date April 1, 2000; deaths occurring in the 3-year period 1999–2001 classified by age at death; births for each of the years 1997–2001; and Medicare data for ages 66–100 years from the years 1999–2001. Methods that were kept the same as those of previous decennial tables (1) include data sources used in constructing tables, the preliminary adjustment for misreported ages, the smoothing techniques for vital statistics and census data, and the calculations of death rates in different age groups.

Two significant changes were made to the methodology used to estimate mortality for the populations aged 66 years and over. First, Medicare data were used to supplement vital statistics (number of deaths) and census data (population estimates) starting at age 66 years instead of age 85 years as was done in the estimation of previous U.S. decennial life tables (1,2). Second, smoothing and extrapolation of death rates for ages 66–109 was performed using a mathematical model given by Heligman and Pollard (3,4), instead of a Whittaker-Henderson Type B formula (1) or modified Gompertz method (2,5).

Keywords: decennial life tables • methodology • vital statistics • Medicare data • Heligman-Pollard model

Introduction

This report describes the methodology employed in preparing the 1999–2001 decennial life tables for the United States. The following tables were produced for the U.S. total population and for the male, female, white, white male, white female, black, black male, and black female populations. These tables are based on the 2000 U.S. census populations and deaths occurring during the 1999–2001 period and are presented in the report "U.S. Decennial Life Tables for 1999–2001, United States Life Tables." The methodology involved in developing the 1999–2001 life tables for each of the 50 individual states and the District of Columbia and for the United States by cause of death is described in the report containing those tables.

Mortality rates for a specific period may be summarized by the life table method to obtain measures of comparative longevity. Two types of life tables exist: the cohort-based (or generation) life table and the period-based (or current) life table. The cohort-based life table presents the mortality experience of a particular birth cohort from the moment of birth through consecutive ages in successive calendar years. For example, a cohort-based life table could be based on the mortality experience of all persons born in the year 1900 and followed until the cohort is extinct. On the basis of age-specific death rates observed through consecutive calendar years, the cohort-based life table reflects the mortality experience of an actual cohort from birth until no lives remain in the group.

The better known current life table may, by contrast, be characterized as "cross-sectional." Unlike the cohort life table, the current life table considers a hypothetical (or synthetic) cohort that is subject throughout its existence to the age-specific mortality rates observed for an actual population during a particular period of relatively short duration (often 1 to 3 years). The current life table may thus be characterized as rendering a "snapshot" of current mortality experience and shows the long-range implications of a set of age-specific death rates that prevailed in a given year(s). The life tables represented in this report are current life tables based on age-specific annual death rates for the period 1999–2001.

The decennial life tables differ in one main respect from the life tables prepared and published annually in the *National Vital Statistics Reports* and the *Vital Statistics of the United States* (6). The annual tables are based on deaths in any single year and, except for the census year, on postcensal population estimates rather than on the enumerated population from a decennial census. The decennial tables, although calculated as yearly rates, are based on deaths in a period of 3 years, including years 1999, 2000, and 2001, and on population counts from the 2000 decennial census.

Changes in the methodology of these tables from those of 1989–1991 involve the technique used to smooth estimates for the population aged 66 years and over. Two major changes were made in this part of the methodology. First, Medicare data were used to supplement vital statistics (number of deaths) and census data (population estimates) starting at age 66 years instead of age 85 years as in the previous method (1,2). Second, the third part of a mathematical model given by Heligman and Pollard (3,4), instead of a Whittaker-Henderson Type B formula (1) or a modified Gompertz method (2,5), was used for smoothing and extrapolation of death rates from ages 66–100 years and from ages 101–109 years. In the previous method, vital statistics and census data were used for ages under 85 years and Medicare data were used for ages 85 years and over. With the new methodology, these two sources of data were both used for the age range 66–94 years, where the Medicare data contribution increases with age throughout the age range 66–94 years. Medicare data are used exclusively for ages 95 to 100 years.

Methods

Data used for calculating life table values

The data used in preparing the 1999–2001 decennial life tables consisted of a) deaths occurring in a 3-year (1999–2001) period classified by age at death and collected from death certificates filed in state vital statistics offices and reported to the National Center for Health Statistics (NCHS) as part of the National Vital Statistics System (NVSS); b) population data by age on the census date April 1, 2000; c) Medicare data for deaths and population for ages 66 years and over during a 3-year (1999–2001) period; and d) births for each of the years from 1997–2001. These data were treated separately by sex and race (white and black).

Populations and deaths were available by single years of age from 0 through 100 years and over. In each case, the referenced age is in completed years—that is, the exact age on the person's last birthday. In addition, deaths occurring at under 1 year of age were available for four subdivisions of the first year of life: under 1 day, 1–6 days, 7–27 days, and 28–364 days. Life table values were calculated for these subdivisions of the first year of life and for single years of age throughout the remainder of the life span.

In preparing these decennial life tables, no specific allowance was made for possible incompleteness in the enumeration of the population or in the registration of births or deaths. In calculating previous decennial life tables, the use of birth statistics rather than population data in calculating the denominators of the mortality rates at ages under 2 years was justified largely on the basis that the census populations under 2 years of age were believed to be underenumerated. In addition, using the methodology based on birth data had other advantages because it might be expected to produce a more accurate estimate of the average population during the 3-year period than is provided by the population enumerated on the census date. Accordingly, its use was continued in the 1999–2001 life tables.

With regard to the census data, actuarial theory suggests that the populations to be used in the calculations should be those of the central date of the 3-year period, that is, July 1, 2000. However, the enumerated populations as of April 1, 2000, were used as if they were July 1 populations. This was done because the percentage differences between the two sets of population figures were very small at the national level. Producing new estimates to reflect the lapse of time between April 1 and July 1, 2000, was not considered necessary.

Preliminary adjustment of the data

Some preliminary adjustments were made to the data before the life tables were constructed. The census populations used in the denominators of the death rates are based on special estimation procedures and are not true counts of the 2000 decennial census. These estimates were produced under a collaborative agreement between NCHS and the U.S. Census Bureau. Reflecting the new guidelines issued in 1997 by the Office of Management and Budget (OMB), the 2000 census included an option for individuals to report more than one race as appropriate for themselves and household members (7). The 1997 OMB guidelines also provided for the reporting of Asian persons separately from Native Hawaiians or Other Pacific Islanders (NHOPI). Under the prior OMB standards (issued in 1977), data for Asian or Pacific Islander (API) persons were collected as a single group (8). Death certificates filed in state vital statistics offices for the years 1999 to 2001 included only one race for the decedent in the same categories as specified in the 1977 OMB guidelines, which also called for the grouping of Asians and NHOPI. Death certificate data by race were therefore incompatible with the population data collected in the 2000 census. To produce death rates for the years 1999–2001, the reported population data for multiple-race persons had to be "bridged" back to single-race categories. In addition, the 2000 census counts were modified to be consistent with the 1977 OMB race categories—that is, to report the data for Asian persons and NHOPI as a combined category, API, and to reflect age as of the census reference data (9). The procedures used to produce the "bridged" populations are described in separate publications (9).

A further relatively minor adjustment was made to the rates for the small number of deaths during the 3-year period for which age was not stated. There were 1,134 deaths with age not reported out of a total of 7,211,175 (0.016 percent) deaths during years 1999–2001. The assumption was made that these deaths were distributed among the various age groups in the same proportions as the deaths for which age was reported. To this end, adjustment factors were computed for each population category for which a life table was to be constructed. This factor F was obtained by dividing the total number of deaths reported for the given category for the 3-year period 1999–2001 by the total number of deaths with age stated. That is,

$$F = T/T_a,$$

where T is the total number of deaths and T_a is the total number of deaths for which age is stated. Death records $D_x^{original}$ with missing age information in each given race and sex category were then proportionally distributed among age categories as

$$D_x^{new} = F\, D_x^{original}.$$

Calculation of the probability of dying, ${}_nq_x$

The life table function q_x is the fraction or proportion of a group of persons at exact age x who are expected to die before attaining age $x + n$. The ${}_nq_x$ is also called the probability of dying. Other functions in the complete life table are derived from q_x, which depends on the number of deaths D_x and the midyear population P_x for each age interval x to $x + n$ observed during the calendar year of interest.

In constructing the decennial life tables, D_x for years 1999–2001 and P_x for decennial year 2000 were smoothed using Beer's graduation technique (10). Beer's multipliers were applied to observed deaths in 5-year age intervals to obtain smoothed single-year-of-age data. Similarly, Beer's multipliers were applied to population estimates in 5-year age intervals to obtain smoothed estimates for single-year age intervals.

For ages under 2 years, birth data B instead of P_x were used to calculate ${}_nq_x$. Life table deaths ${}_nd_x$ and surviving populations at the beginning of each age l_x were derived from ${}_nq_x$ as described in the following section.

Probabilities of dying at 2 years of age and under

At ages under 2 years, the first life table quantities to be calculated were the values of ${}_nd_x$, which is the number of deaths occurring between exact age x and $x + n$ in the life table cohort commencing with l_0 live births. This was calculated by the formula

$${}_nd_x = l_0 \, {}_nD_x \, / \, {}_nE_x \, ,$$

where ${}_nd_x$ is the estimated number of deaths per 100,000 birth population occurring in 1999–2001 between exact ages x and $x + n$, ${}_nD_x$ is the actual death counts (adjusted for nonreporting of age) in this 3-year period, and ${}_nE_x$ is the appropriate denominator as indicated in Table A. These denominators are based on the assumption of uniform distribution over the year of the births of 1997, 1998, 1999, 2000, and 2001. In each case, the initial life table population l_0 is taken as 100,000. Age intervals under 2 years were classified as 0 days (under 1 day), 1–6 days, 7–27 days, 28–364 days, and 0–1 year. The number of deaths at age 0 is the sum of deaths occurring at ages 0 to 364 days. The number of deaths at age 1 year includes all deaths occurring from 365 days to the day just before the second birthday.

The unrounded values of ${}_nd_x$ were then used to calculate the number of survivors l_x up to age 2 years by successive applications of the formula

$$l_{x+n} = l_x - {}_nd_x \, .$$

Table A. Denominators ${}_tE_x$ used in calculating ${}_td_x$ for ages under 2 years: U.S. Decennial Life Tables, 1999–2001

Age interval x to $x + n$	${}_nE_x$
Under 1 day	$\frac{1}{730}\,(B_{1998} + 730B_{1999} + 730B_{2000} + 729B_{2001})$
1–6 days	$\frac{1}{730}\,(8B_{1998} + 730B_{1999} + 730B_{2000} + 722B_{2001})$
7–27 days.	$\frac{1}{730}\,(35B_{1998} + 730B_{1999} + 730B_{2000} + 695B_{2001})$
28–364 days	$\frac{1}{730}\,(393B_{1998} + 730B_{1999} + 730B_{2000} + 337B_{2001})$
1–2 years	$\frac{1}{2}\,(B_{1997} + 2B_{1998} + 2B_{1999} + B_{2000})$

NOTE: For explanation of equations and variables, see *National Vital Statistics Reports*, Volume 56, Number 4, "U.S. Decennial Life Tables for 1999–2001: Methodology of the United States Life Tables."

The probability of dying within each age category for those under 2 years of age was estimated as

$${}_nq_x = {}_nd_x \, / \, l_x \, .$$

Probabilities of dying at ages 2–94 years from vital statistics data

Interpolation of P_x and D_x at ages 2–94 years

Anomalies, both random and those associated with age reporting, can be problematic when using vital statistics and census data by single years of age (10). Graduation techniques are often used to eliminate these anomalies and to derive a smooth curve by age. Beer's ordinary minimized fifth difference formula is such a technique and has been used in the construction of the previous U.S. decennial life tables (1,10) as well as annual life tables (2).

Population data were aggregated into 5-year age intervals except for those aged 100 years and over, which are allocated into a single category. Values of P_x by single years of age were obtained by interpolation using Beer's formula; Beer's general formula adapted to calculate P_x is

$$P_{x+k} = C_{k,\,x-10}\,{}_5P_{x-10} + C_{k,x-5}\,{}_5P_{x-5} + C_{k,x}\,{}_5P_x + C_{k,\,x+5}\,{}_5P_{x+5} + C_{k,\,x+10}\,{}_5P_{x+10} \, ,$$

where P_{x+k} is the population aged $x + k$ ($k = 0,1,2,3,4$), ${}_5P_x$ is the total population aged x to $x + k$, and $C_{k,x}$ is Beer's interpolation coefficient for the kth fifth of the age interval x to $x + 5$ applied to ${}_5P_x$. To interpolate single-year values from ${}_5P_0$ and ${}_5P_5$, the formula is slightly different. To obtain single-year values for these 5-year intervals, the formulas below are used.

$$P_{0+k} = C_{k,0}\,{}_5P_0 + C_{k,5}\,{}_5P_5 + C_{k,10}\,{}_5P_{10} + C_{k,\,15}\,{}_5P_{15} + C_{k,20}\,{}_5P_{20}$$

$$P_{5+k} = C_{k,0}\,{}_5P_0 + C_{k,5}\,{}_5P_5 + C_{k,10}\,{}_5P_{10} + C_{k,\,15}\,{}_5P_{15} + C_{k,20}\,{}_5P_{20}$$

Values for $C_{k,x}$ are shown in Table B.

Interpolating single-year values for D_x was conducted in a similar fashion as that for P_x. The Beer's coefficients $C_{k,x}$ for deaths are the same as for ${}_5P_x$ as shown in Table B. The difference is that when interpolating ages 5–9 and 10–14 years, a fictitious value for ${}_5D_0$ was used. Because of the mortality peak in infancy, the use of the observed ${}_5D_0$ does not yield values for D_x for ages 5–14 years that join smoothly with the numbers reported at ages 0–4 years. The fictitious value for ${}_5D_0$ was calculated such that

$$V = .4072\,{}_5D_0^* + .2146\,{}_5D_5 + .0080\,{}_5D_{10} - .0896\,{}_5D_{15} + .0328\,{}_5D_{20} \, ,$$

where V is the sum of the deaths occurring at ages 2, 3, and 4 years of age and ${}_5D_0^*$ is the fictitious value for ${}_5D_0$. Solving for ${}_5D_0^*$ gives

$${}_5D_0^* = 2.45580\,V - .59332\,{}_5D_5 - .01965\,{}_5D_{10} + .22004\,{}_5D_{15} - .08055\,{}_5D_{20}.$$

${}_5D_0^*$ is then treated in the interpolation formula as if it were the actual number of deaths at ages 0–4 years. This modification produces a smooth transition from the observed values at ages under 5 years to the interpolated values at ages 5 years and over.

Table B. Interpolation coefficients based on Beer's ordinary formula for the subdivision of grouped data into fifths: U.S. Decennial Life Tables, 1999–2001

Interpolated subgroup ($x + k$)	5-year age interval beginning with age				
2 to 4 years	0	5	10	15	20
2 years	0.1924	0.0064	0.0184	–0.0256	0.0084
3 years	0.1329	0.0844	0.0054	–0.0356	0.0129
4 years	0.0819	0.1508	–0.0158	–0.0284	0.0115
5 to 9 years	0	5	10	15	20
5 years	0.0404	0.2000	–0.0344	–0.0128	0.0068
6 years	0.0093	0.2268	–0.0402	0.0028	0.0013
7 years	–0.0108	0.2272	–0.0248	0.0112	–0.0028
8 years	–0.0198	0.1992	0.0172	0.0072	–0.0038
9 years	–0.0191	0.1468	0.0822	–0.0084	–0.0015
$5x$ to $5x+5$ years ($x = 2, 3, 4, ..., 18$)	$5x-10$	$5x-5$	$5x$	$5x+5$	$5x+10$
$5x$ years	–0.0117	0.0804	0.1570	–0.0284	0.0027
$5x + 1$ years	–0.0020	0.0160	0.2200	–0.0400	0.0060
$5x + 2$ years	0.0050	–0.0280	0.2460	–0.0280	0.0050
$5x + 3$ years	0.0060	–0.0400	0.2200	0.0160	–0.0020
$5x + 4$ years	0.0027	–0.0284	0.1570	0.0804	–0.0117

NOTE: An additional panel of interpolation coefficients is usually shown for the final age interval. This panel is not shown because the final age interval in the life table is open-ended and is not based on interpolated data.

Central death rates at ages 2–4 years

The life table function q_x, denoting the ratio d_x / l_x, is the fraction or proportion of a group of persons at exact age x who are expected to die before attaining age $x + 1$. If m_x denotes the ratio d_x / L_x, commonly called the central death rate, then on the assumption of uniform distribution of deaths over the year at age x,

1. $$q_x = \frac{2m_x}{2 + m_x}.$$

This approximation is appropriate when the life table is by single years of age. This formula was the basis of the calculation of probabilities of dying at ages 2–94 years. Completion of the calculations depends, therefore, on the ability to calculate the central death rate m_x at different ages. For this purpose, different methods were used at ages 2–4 years and at ages 5–94 years, as will now be described.

If D_x denotes the adjusted number of deaths in a population category at age x (in completed years) occurring in 1999–2001 and P_x denotes the population at age x in the middle of the period, then

2. $$m_x = \frac{D_x}{3P_x}.$$

As previously noted, the populations actually used were those of April 1, 2000.

However, because the deaths occurring in a single year of age during 1999–2001 were drawn from three consecutive annual cohorts of the population, the accuracy of these m_x values was considered to be improved by replacing $3P_x$ in the denominator of formula 2 by the sum of the populations at ages $x - 1$, x, and $x + 1$. Thus the formula becomes

3. $$m_x = \frac{D_x}{P_{x-1} + P_x + P_{x+1}}.$$

The combination of formulas 1 and 3 is equivalent to the single formula

4. $$q_x = \frac{D_x}{P_{x-1} + P_x + P_{x+1} + \frac{1}{2} D_x},$$

which was used for $x = 2$, 3, and 4, with values of D_x and P_x obtained by interpolation from data by 5-year age intervals.

Probabilities of dying at ages 5–94 years

The combination of formulas 1 and 2 is equivalent to

5. $$q_x = \frac{D_x}{3P_x + \frac{1}{2} D_x},$$

which was used for ages 5–94 years, with values of D_x and P_x obtained by interpolation from data by 5-year age intervals.

Probabilities of dying at ages 66–114 based on Medicare data

Medicare data were employed for the estimation of q_x at ages 66 years and over because age-specific death rates at the older ages derived from Medicare data are considered more accurate than those derived from vital statistics and census data (1,5). The prevalence of age misreporting at the oldest ages in census data has been found to be nontrivial, leading to the underestimation of death rates in the oldest ages. This may be especially the case for the African American population (5). Medicare data are believed to be superior because beneficiaries must prove their date of birth in order to enroll (11). Its coverage is also extensive. Approximately 98 percent of Americans aged 65 years and over are enrolled in Medicare Part A, and 96 percent of these are enrolled in Part B. In addition, 99 percent of deaths from Americans aged 65 years and over in the United States are accounted for in the Medicare program (12). Medicare data suffer, however, from a greater prevalence of people in the oldest age groups. The number of people aged 90 years and over is greater in

Medicare data than in census data (12). This is a result of the presence of "phantom records" that arise as a result of persons being registered more than once or because a person's death was not reported (11). To reduce the number of "phantom records," the Medicare death rates are based on the records of Medicare beneficiaries who were also eligible for Social Security or Railroad Retirement income benefits. Approximately 3 percent of Medicare records were eliminated as a result (11).

Estimates of age-specific number of deaths and populations by sex and race for the population aged 65 years and over were generated from the 1999, 2000, and 2001 Medicare file created by the Centers for Medicare and Medicaid (CMS) for the Social Security Administration, which then shares the files with NCHS. The files include population counts $P_{y,x-1}$ for age $x-1$ on January 1 of year y and $P_{y+1,x}$ for age x on January 1 of year $y+1$ as well as death counts $D_{y,x}$ at age x that occurred throughout year y by race and sex. Thus, the probability of death for ages 66 years and over derived from Medicare data is estimated as

$$q_x^M = D_{y,x} / [(P_{y,x-1} + P_{y+1,x} + D_{y,x}) / 2] ,$$

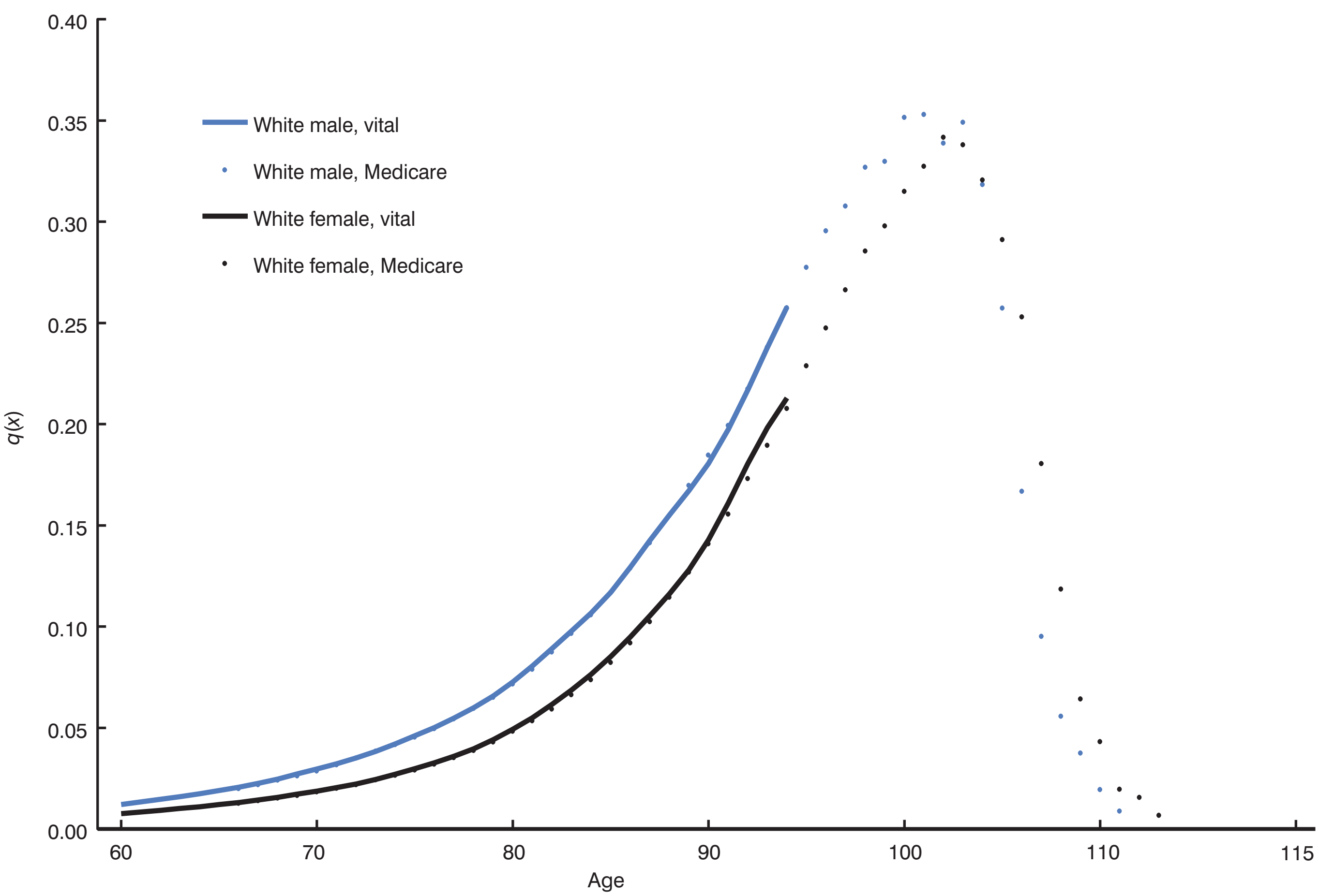
where $D_{y,x}$, $P_{y,x-1}$, and $P_{y+1,x}$ are the sum of the corresponding values for years 1999, 2000, and 2001.

Probabilities of dying derived from Medicare data ranged from ages 66–114 years, whereas the ones from vital statistics were estimated up to age 94. Figures 1 and 2 show q_x^M (Medicare) and q_x^v (vital statistics) by age for the white male and female populations and the black male and female populations. The figures show that q_x^M rises monotonically up to approximately age 100 years but then declines dramatically until the last age (about 114) among each of the four groups. The q_x^M is about the same as q_x^v through age 66 years to about 80 years, and then gradually departs from q_x^v for all four populations. Medicare q_x^M was used to adjust vital q_x^v for ages 66–94 years and to model mortality patterns for ages 95 years and over.

Final probabilities of dying at ages 66–109

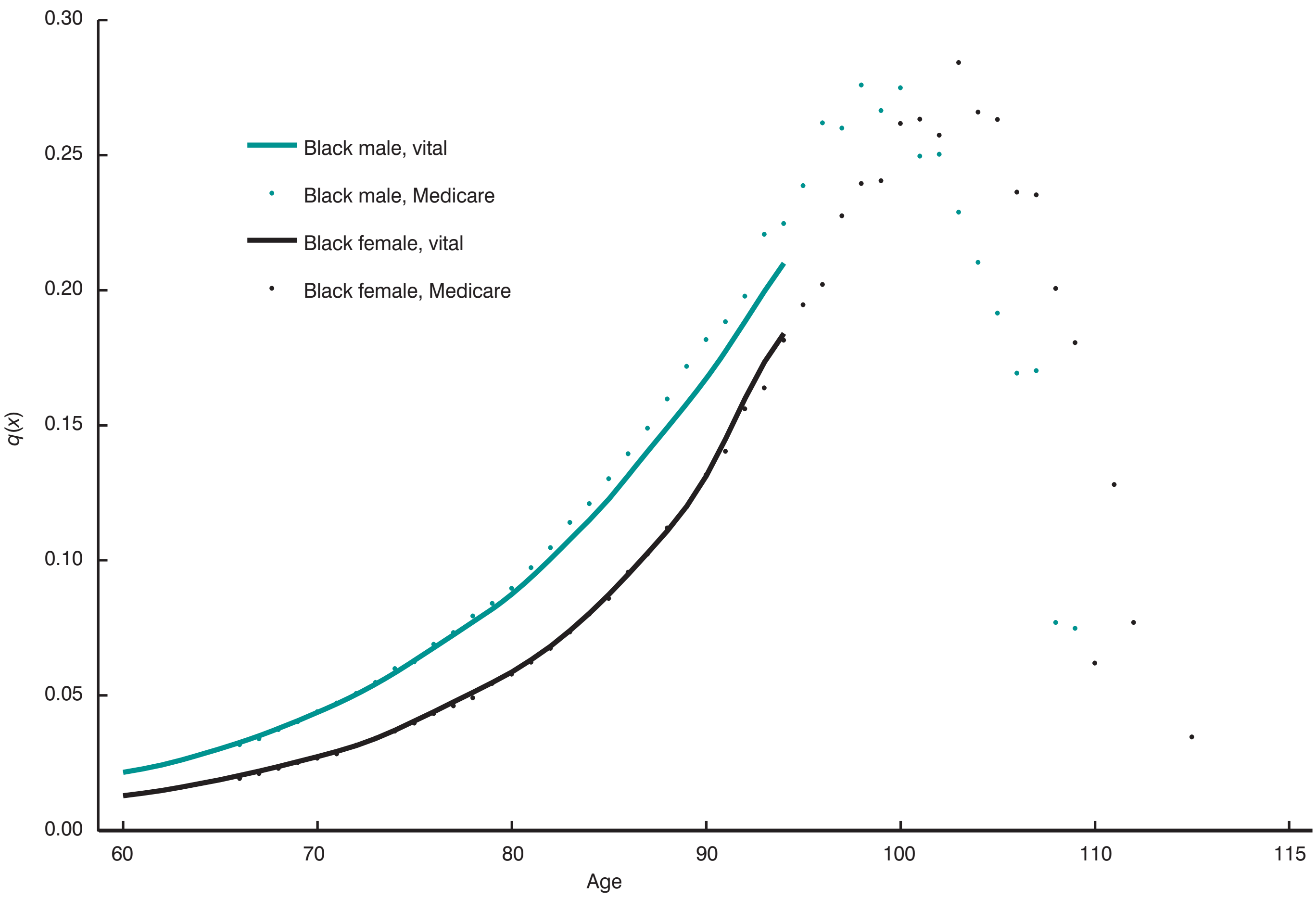
Combining vital statistics and Medicare data at ages 66–94 years

At ages 66–94 years, the probabilities of dying were obtained by blending vital statistics q_x^v with Medicare q_x^M through a weighting process giving gradually declining weight to vital statistics data and gradually increasing weight to Medicare data as follows:



SOURCE: CDC/NCHS, *National Vital Statistics Reports*, Volume 57, Number 2, "U.S. Decennial Life Tables for 1999–2001: Methodology of the United States Life Tables."

Figure 1. Comparison of probabilities of dying ($q(_x)$) from vital statistics data (1999–2001) and Medicare data (1999–2001) in the white population



SOURCE:CDC/NCHS, *National Vital Statistics Reports*, Volume 57, Number 2, "U.S. Decennial Life Tables for 1999–2001: Methodology of the United States Life Tables."

Figure 2. Comparison of probabilities of dying ($q(_x)$) from vital statistics data (1999–2001) and Medicare data (1999–2001) in the black population

6. $$q_x = \frac{1}{30}[(95-x)q_x^v + (x-65)q_x^M], \quad \text{when } x = 66, \ldots, 94,$$

and

$$q_x = q_x^M, \quad \text{when } x = 95, \ldots, 100,$$

where q_x is a combination of q_x^v and q_x^M, q_x^v is the death rate calculated with formula 5, and q_x^M is the corresponding rate based on Medicare data. Death rates q_x^M above age 100 were not used because their declining values were considered to be highly unrealistic.

Smoothing combined death rates at ages 66–100 years with the third part of the Heligman-Pollard equation

In a study using vital statistics 1999–2001 mortality data, patterns of age-specific death rates were found to be well fitted with the Heligman-Pollard equation (13). The Heligman-Pollard equation is a nonlinear model consisting of three components and eight parameters:

$$\frac{q_x}{1-q_x} = A^{(x+B)^C} + D\exp[-E(\log x - \log F)^2] + GH^x,$$

where, in the first term of the equation, parameter A measures mortality in the first year of life, parameter B measures the rate of change in mortality from birth to the first year of life, and parameter C measures the rate of mortality decline in childhood. The second term of the equation describes mortality from ages 10–40, in which an "accident hump" appears, and parameters D, E, and F measure the location, width, and height of this "accident hump." The parameters G and H in the third term measure mortality levels and changes in the older population (approximately 40 years and over) (3,14).

The third term of the Heligman-Pollard equation actually follows the well-known "Gompertz law" of exponentially increasing mortality with increasing age. The Gompertz law was published in 1825 and has been widely accepted and used (4,15,16). There have been several studies (2,5,16–18) applying a modified version of the Gompertz law using U.S. elder populations. In one such modified Gompertz's application, changes in mortality rates above age 85 years are linearly fitted and predicted. For example, the U.S. annual life tables since 1997 have used linear indexes of mortality change, k_xs, to model mortality from the ages of 85–100 years (2). In a recent study (4), the third component of the Heligman-Pollard model was found to provide a better estimate of mortality patterns at the older ages than the linearized k_x. In comparison with a straight line, the Heligman-Pollard nonlinear smoothing procedure provides a more flexible and robust fit of observed data, and predicted mortality rates appear more reasonable. Furthermore, fitting a nonlinear equation has become feasible with new computing techniques such as the SAS nonlinear procedure.

The q_x was fitted with the third component of the Heligman-Pollard equation for ages 66–100 years:

7. $$\frac{\hat{q}_x}{1-\hat{q}_x} = GH^x,$$

where the G and H are two fitted parameters specific for each population and used to interpolate for ages 66–100 years and extrapolate q_x for ages 101–109 years. The model was estimated with the SAS nonlinear weighted least squares procedure, where weights $w_x = 1 / q_x^2$ (3,9). The iterated estimation procedure requires that initial values of G and H be predetermined. These initial values were chosen by referring to estimated parameters in Table 2 of Hartmann's paper (14).

To ensure a smooth transition from vital q_x^v to the estimated $\hat{q}_x$ at ages 65–66 years, two treatments were applied. The first was to anchor the fitting curve in equation 7 at age 65 years by forcing the fitted curve to pass through age 65 years with a fitting option in the SAS nonlinear procedure. The second was to merge the estimated $\hat{q}_x$ and vital q_x^v from ages 66–74 years with a graduating process by taking

8. $$q_x = \frac{1}{10}[(75-x)q_x^v + (x-65)\hat{q}_x], \quad \text{when } x = 66, \ldots, 74.$$

Extrapolating q_x up to age 130

To estimate the population l_x at the last survival age, q_x was estimated for all ages up to 130 years by extrapolation even though the oldest observed survivor was aged 115 years and the final life tables were truncated at age 109 years. Also, estimating q_x up to age 130 years was done in order to be consistent with the previous decennial life tables, and allow for ease of comparison:

9. $$\hat{q}_x = \frac{\hat{G}\hat{H}^x}{1+\hat{G}\hat{H}^x},$$

where the extrapolated $\hat{q}_x$ was calculated by using estimated $\hat{G}$ and $\hat{H}$ from equation 7, and the estimated $\hat{G}$ and $\hat{H}$ are given in Table C.

Calculating remaining life table values: d_x, l_x, L_x T_x and $\mathring{e}_x$

Once q_x for each age is determined from age 0–130 years, the rest of the life table functions are derived from q_x sequentially.

The initial life table population l_0 at birth was set to 100,000 as a standard population. The remaining life table values were then estimated with the following equations.

Number dying between age x and $x+1$ is

10. $$d_x = l_x q_x.$$

Number alive at the beginning of each age interval is

11. $$l_{x+1} = l_x - l_x q_x.$$

Number of person years lived between age x to $x+1$ is

12. $$L_x = 0.5\,(l_x + l_{x+1}).$$

Note for ages less than 1 year: the L_x should be expressed as a quantity in a fractional year; that is, L_x should be calculated by multiplying the equation by 1/365 for age 0–1 day, by 6/365 for 1–7 days, by 21/365 for 7–28 days, and by 337/365 for 28–365 days.

Number of person-years lived after age x is

13. $$T_x = L_x + T_{x+1}, \text{ with } T_{end} = L_{end}.$$

The average remaining lifetime (or life expectancy) at age x is

14. $$\mathring{e}_x = T_x / l_x.$$

Finally, all life table functions were estimated for ages up through 130 years but truncated at age 109 years, even though last survival ages exceeded 109 years for all populations (Table D).

Decimal places and rounding

The published decennial life tables are based on a hypothetical cohort of 100,000 live births. Traditionally, published life tables show life table functions such as l_x, d_x, L_x, and T_x rounded to whole integers. However, because the U.S. total population is nearly 300 million, the accuracy of the life table functions extends beyond the hypothetical population of 100,000. Therefore, all life table calculations are carried out using floating point precision, allowing for fractional deaths and fractional years of life lived. This creates a problem for users who want to reproduce the life table estimates from the rounded numbers in the publication.

For example, in Table E, the actual life table values are shown using a large number of significant digits. For age 107 years in the table of the male population, the fractional values substituted in the equation d_x / l_x, specifically 6.180413227/11.653673864, reproduces the estimated q_x value of 0.53034. However, the published life table shows a rounded l_x as 12 and a rounded number of deaths d_x as 6. This seems to give a probability of dying as 0.5 rather that the published probability of dying of 0.53034. Furthermore, the rounded numbers of 12 minus 6 would give the next l_{108} as 6 instead of the published number of 5, whereas the unrounded numbers 11.653673864 minus 6.180413227 give the number 5.473260638. Thus, users of the decennial life tables are cautioned that the life table calculations were based on more significant digits than shown and that back-calculation using the rounded numbers cannot be expected to reproduce the exact published results.

Calculation of standard errors of the life table functions

The standard errors for the decennial life table functions, specifically for the probabilities of dying and for life expectancies, were calculated based on an assumption that the age-specific deaths follow a binomial distribution. One should consider that these standard errors reflect only stochastic variation. Stochastic variation is not the only source of error for life table functions; measurement error, such as age misstatement on death certificates or on census reports, also affects the accuracy of the life table functions. Although the extent of measurement error on life table functions has not been quantified, measurement errors are generally thought to be larger than stochastic errors (1). The standard errors presented are rather small because the life tables for the United States and for each individual state are based on relatively large numbers of deaths.

Table C. Estimated parameters *G* and *H* used for extrapolating q_x from ages 101–109 years: U.S. Decennial Life Tables, 1999–2001

	Population								
	Total	Male	Female	White	White male	White female	Black	Black male	Black female
G	0.0000224	0.0000343	0.0000106	0.0000190	0.0000286	0.0000089	0.0001427	0.0002436	0.0000694
H	1.1049	1.1021	1.1129	1.1071	1.1045	1.1152	1.0817	1.0779	1.0892

NOTE: For explanation of equations and variables, see *National Vital Statistics Reports*, Volume 56, Number 4, "U.S. Decennial Life Tables for 1999–2001: Methodology of the United States Life Tables."

Table D. The last age of survival (l_x) for each population: U.S. Decennial Life Tables, 1999–2001

	Population								
	Total	Male	Female	White	White male	White female	Black	Black male	Black female
Age at the end ($l_x > 0.50$)	112	110	112	112	112	114	115	112	115

Table E. The effect of rounding on life table estimates for the male population: U.S. Decennial Life Tables, 1999–2001

Age interval x to $x+1$	Probability of dying between ages x to $x+n$ $_nq_x$	Number surviving to age x l_x	Number dying between ages x to $x+n$ $_nd_x$	Person-years lived between ages x to $x+n$ $_nL_x$	Total number of person-years lived above age x T_x	Expectation of life at age x e_x
	Unrounded numbers (unpublished)					
106–107	0.50607	23.59401399	11.94034013	17.62384393	33.03351645	1.400080396
107–108	0.53034	11.65367386	6.180413227	8.563467251	15.40967253	1.32230168
108–109	0.55446	5.473260638	3.034720932	3.955900171	6.846205276	1.250845836
109–110	0.57833	2.438539705	1.410285401	1.733397005	2.890305104	1.185260629
	Rounded numbers (published)					
106–107	0.50607	24	12	18	33	1.4
107–108	0.53034	12	6	9	15	1.32
108–109	0.55446	5	3	4	7	1.25
109–110	0.57833	2	1	2	3	1.19

A binomial distribution assumption yields the following estimate for the variance of q_x (19):

15. $$S^2(q_x) = \frac{q_x^2(1-q_x)}{D_x},$$

where D_x is the age-specific number of deaths. For ages less than 65 years, D_x are the deaths from vital statistics data, smoothed by interpolation and adjusted for the number of deaths with age not stated. For ages 65 years and over, the D_x was obtained by treating the population as a cohort population and calculated from q_x because mixed vital statistics and Medicare data were used for estimation:

16. $$P_x = \frac{(P_{x-1} - 0.5D_{x-1}/3)(2-q_x)}{2}.$$

17. $$D_x = \frac{3q_xP_x}{1-0.5q_x}.$$

Note that D_x is the number of deaths in a 3-year data collection period (1999–2001) and P_x is the population at age x in the middle of this period.

For the variances of the life expectancies at ages 0–109 years, an equation from Chiang (19) with a slight modification was used (1):

18. $$S^2(\mathring{e}_x) = \frac{l_{end}^2 S^2(\mathring{e}_{end}) + \sum_{y=x}^{end-1} l_y^2 (\mathring{e}_{y-1} + 0.5)^2 S^2_{(q_y)}}{l_x^2},$$

where the "end" age is the age 1 year before the life expectancy becomes 0 for the population. For all populations, this "end" age fell into a range between 109 and 115 years. An approximate estimation for the variances of life expectancy at the "end" is

19. $$S^2(\mathring{e}_x) = \frac{S^2_{(q_{end})}}{q_{end}^4}.$$

This is because $\mathring{e}_{end} = 1/m_{end} = (2-q_{end})/2q_{end}$ and $S^2(\mathring{e}_{end}) = ([(2-q_{end})/2q_{end}]')^2 S^2(q_{end})$. The $S^2(q_x)$ and $S^2(\mathring{e}_x)$ were calculated until the "end" age, but they were cut back to the age of 109 years for the published life tables.

Conclusions

The methodology described in this report provides all details in producing estimates of indices in the United States Decennial Life Tables for 1999–2001, including the probability of death, survivor population, number of deaths, stationary population, and life expectancy at single-year age intervals from age 0–109 years. The tables include total population as well as the male, female, white, white male, white female, black, black male, and black female populations.

References

1. Armstrong RH. Methodology of the national and state life tables, U.S. decennial life tables for 1989–91; vol 1, no 2. Hyattsville, MD: National Center for Health Statistics. 1998.
2. Anderson RN. A method for constructing complete annual U.S. life tables. National Center for Health Statistics. Vital Health Stat 2(129). 1999.
3. Heligman P, Pollard JH. The age pattern of mortality. Journal of the Institute of Actuaries 107(1):49–80. 1980.
4. Wei R, Curtin RL, Anderson RN, Arias E. Estimation of United States year 2000 decennial life tables: Which model is most appropriate for old age mortality? Hyattsville, MD: National Center for Health Statistics. 2005.
5. Coale AJ, Kisker EE. 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(1):1–31. 1990.
6. National Center for Health Statistics. United States life tables, 2000. National vital statistics reports; vol 51, no 3. Available from: http://www.cdc.gov/nchs/products/pubs/pubd/nvsr/nvsr.htm. 2003.
7. Office of Management and Budget. Revisions to the standards for the classification of federal data on race and ethnicity. Federal Register 62FR58782–58790. October 30, 1997. Available from: http://www.whitehouse.gov/omb/fedreg/ombdir15.html. October 30, 1997.
8. Office of Management and Budget. Race and ethnic standards for federal statistics and administrative reporting. Statistical Policy Directive 15. 1977.
9. Ingram DD, Weed JA, Parker JD, Hamilton B, Schenker N, Arias E, et al. U.S. census 2000 with bridged race categories. Vital Health Stat 2(135). 2003.
10. Shryock HS, Siegel JS, et al. The methods and materials of demography, vol 2. U.S. Bureau of the Census. Washington, DC: U.S. Government Printing Office. 1971.
11. Bell FC, Miller ML. Life tables for the United States Social Security Area 1900–2100. Baltimore, MD: Social Security Administration, Office of the Chief Actuary. SSA Pub. No. 11–11536. 2005.
12. Research Data Assistance Center. Introduction to the use of Medicare data for research. Minneapolis, MN: University of Minnesota School of Public Health. 2004.
13. Wei R, Curtin LR, Anderson RN. Model U.S. mortality data for building life tables and further studies. In: 2003 Proceedings of the American Statistical Association, Biometrics Section [CD-ROM]. Alexandria, VA: American Statistical Association. p. 4458–64. 2004.
14. Hartmann M. Past and recent attempts to model mortality at all ages. Journal of Official Statistics 3(1):19–36. 1987.
15. Thatcher AR, Kannisto V, Vaupel JW. The force of mortality at ages 80 to 120. Odense, Denmark: Odense University Press. 1998.
16. Horiuchi S, Coale AJ. A simple equation for estimating the expectation of life at old ages. Popul Stud 36(2):317–26. 1982.
17. Wilmoth JR. Are mortality rates falling at extremely high ages? An investigation based on a model proposed by Coale and Kisker. Popul Stud 49:281–95. 1995.
18. Horiuchi S, Wilmoth JR. Deceleration in the age pattern of mortality at older ages. Demography 35(4):391–412. 1998.
19. Chiang CL. Introduction to stochastic processes in biostatistics. New York, NY: John Wiley and Sons. 1968.

Contents

Acknowledgments

The authors are grateful to Kirill F. Andreev, Population Division, Department of Economics and Social Affairs, United Nations; Robert Schoen, Department of Sociology, Pennsylvania State University; and John R. Wilmoth, Department of Demography, University of California-Berkeley for their insightful comments and suggestions. We are also grateful to Felicitie C. Bell, Stephen C. Goss, Bert Kestenbaum, Alice H. Wade, and Alvin K. Winters, Office of the Chief Actuary, Social Security Administration, for their helpful suggestions and for providing us with the Medicare data used to produce the life tables. This report was prepared under the general direction of Charles J. Rothwell, Director, Division of Vital Statistics, and Robert N. Anderson, Chief, Mortality Statistics Branch (MSB). Content review was provided by Betzaida Tejada-Vera of MSB. This review was edited by Megan Cox and Demarius V. Miller, CDC/CCHIS/NCHM/Division of Creative Services, Writer-Editor Services Branch, and typeset by Jacqueline M. Davis, CDC/CCHIS/NCHM/Division of Creative Services; graphics were produced by Sarah M. Hinkle CDC/CCHIS/NCHM/Division of Creative Services.

Suggested citation

Wei R, Curtin LR, Arias E, Anderson RN. United States decennial life tables for 1999–2001, methodology of the United States life tables. National vital statistics reports; vol 57, no 4. Hyattsville, MD: National Center for Health Statistics. 2008.

U.S. DEPARTMENT OF
HEALTH & HUMAN SERVICES

Centers for Disease Control and Prevention
National Center for Health Statistics
3311 Toledo Road
Hyattsville, MD 20782

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, $300

FIRST CLASS MAIL
POSTAGE & FEES PAID
CDC/NCHS
PERMIT NO. G-284

To receive this publication regularly, contact the National Center for Health Statistics by calling 1–800–232–4636
E-mail: cdcinfo@cdc.gov
Internet: www.cdc.gov/nchs

CS120144
T32231 (08/2008)
DHHS Publication No. (PHS) 2008–1120