

ABOUT MORTALITY DATA FOR UNITED STATES

By Mila Andreeva

Last Updated by Magali Barbieri: March 1st, 2023

Warning: the HMD mortality data for years 2010-2021 are preliminary as the US Census Bureau has yet to revise its inter-censal (2010-2019) and post-censal (2020-2021) population estimates using the results of the 2020 Population census. We will update the HMD data series as soon as the US Census Bureau has released new population estimates.

GENERAL

National censuses and population estimates for the United States are produced by the U.S. Census Bureau, U.S. Department of Commerce (www.census.gov). Vital statistics data are collected by the National Center for Health Statistics (NCHS) (www.cdc.gov/nchs/).

Sources of Data

Data on population originate from population censuses conducted every 10 years (U.S. Census Bureau, 1931, 1942, 1954, 1960–, 2010). Census counts serve as the basis for producing annual and monthly population estimates for intercensal and postcensal periods, which are published in *Current Population Reports: Population Estimates and Projections, Series P-25* starting with year 1970. Postcensal population estimates are frequently revised and intercensal population estimates are substituted to postcensal estimates as new census data become available. For recent years, population estimates are published on the Census Bureau web site (U.S. Census Bureau, Population Estimates Program). As mentioned on the Census Bureau website: "With each new release of annual estimates, the entire time series of estimates is revised for all years back to the last census."¹ So, for instance, with the release of the July 1st population estimates for 2019, the estimates for years 2010 through 2019 were also all revised. For years 1933-1939, annual population estimates were also available from the Census Bureau but for years 1940-1969, they were constructed from vital statistics and census data as indicated in Appendix 3.

Data on births and deaths are available from periodical vital statistics publications (U.S. Census Bureau, 1936–1938; National Center for Health Statistics, 1939–). Starting with 1959, data on deaths are also released by NCHS in the form of electronic public Detailed Mortality Files and Multiple Causes of Death Files (National Center for Health Statistics, 1959–, 1968–). These data include individual death records coded from death certificates.

Mortality data for years 1959-1967 were initially obtained from Duke University, who received them from the National Center for Health Statistics through a special agreement. Data for years 1968-1988 were accessed online through the Inter-University

¹ <https://www.census.gov/data/tables/2017/demo/popest/nation-detail.html>, accessed April 23rd, 2018.

Consortium for Political and Social Science Research (ICPSR, www.icpsr.umich.edu). Data for years since 1988 are directly available on the Centers for Disease Control and Prevention web site (CDC, http://www.cdc.gov/nchs/data_access/VitalStatsOnline.htm). However, due to issues of confidentiality, a number of variables are suppressed in the electronic mortality files publicly available (whether through CDC or through NCHS). In particular, the public mortality files do not include information about the date of birth of the deceased for all years starting with 1989. Consequently, we have requested and been granted access through the Berkeley Research Data Center to the restricted mortality files so as to tabulate death counts by Lexis triangles (i.e. by both age and year of birth) for all deaths. The publicly available data are released earlier than the protected data so we typically conduct two updates each year: a first, preliminary, update using the publicly available mortality data (deaths by sex and single year of age) and the mid-year population estimates for that same year, and, once the restricted mortality files have been released to our RDC account, a second update using these restricted data (from which we tabulate the deaths by sex, single year of age and birth cohort) as well as the mid-year population estimates for that same year and for the following year. The current update uses the publicly available mortality data (deaths by single year of age up to the last age of survival) for 2021.

TERRITORIAL COVERAGE

Until 1959, data on population and deaths refer to territory of the United States excluding Alaska and Hawaii. Since 1959, geographical coverage includes the 50 states and the District of Columbia.

DEATH COUNT DATA

Coverage and Completeness

In the early years of the 20th century, vital statistics for United States were based on data from those states admitted to the Death Registration Area, the number of which increased over time. This process was completed in 1933 with the admission of Texas. Given the legal requirements of death registration, the data on deaths are considered to be complete and of acceptable quality since 1933, explaining why the HMD mortality series for the United States start with this year.

Mortality data for the United States are confined to events registered in the United States (as defined in the previous section). The data do not cover vital events to U.S. residents occurring outside of the United States. If the death of a nonresident occurred in the United States, it was included in the U.S. vital statistics. For nonresidents, the usual practice was to attribute the place of death as the place of residence. Since 1970, it is possible to identify deaths of nonresidents and, consequently, to exclude them from tabulations. Therefore, for the years 1933–1969, deaths in the HMD cover both residents and nonresidents (i.e., the *de facto* population), and for the period starting in 1970, they only cover residents.

Specific Details

Tabulation of Deaths from Individual Death Certificate Data

As described above, deaths included in the mortality database have been tabulated from individual death records disseminated by NCHS in the form of Detailed Mortality Files (MDF) and Multiple Cause of Death Files (MCD) for all years 1959 and after. For all years when information on both the age at death and the date of birth are available (1989 and after), deaths were directly tabulated into Lexis triangles (see 1989-2019 for details). The detailed age at death was also used in cases where the age and birth cohort information were either inconsistent or missing as described below. For years before 1989, deaths by single-year-of-age were allocated to Lexis triangles using all the information available. More specifically, for deaths below age one, the detailed age at death (provided in hours, minutes, weeks or months) was used to exactly determine the corresponding upper or lower Lexis triangle. For all other deaths, information on the exact date of death (month and, where available, day) was used as described in Appendix 2.

1933–1958

For these years, death by age and sex tabulations were copied from published vital statistics volumes and processed according to standard HMD methods.

1959–1967

For the years 1959–1961, the day of death is not coded, so tabulations are based on the month of death only, while for the years 1962–1967, both the day and the month of death are available for most records.

According to the documentation, data for Hawaii are included in the 1959 file, even though they were not included in published U.S. figures until 1960. In addition, the annual report for the *Vital Statistics of the United States, 1959* was not produced using the 1959 tape file. The 1959 tape file was provided to the Division of Vital Statistics by a group commissioned by the American Public Health Association (APHA) to work on mortality in 1959–61.

1968–1978

The data for 1975 that are available through the ICPSR website appear to include several corrupt records, so the 1975 MCD file was obtained directly from NCHS by Duke University and the data from ICPSR were not used. The coding of variables used for tabulations and the information included in the data files are consistent throughout the whole period, so the same procedure has been applied to all files.

For the years 1968–1971, the day of death is not available; for years 1972–1978, the complete date of death is recorded.

1979–1980

The source and processing of the data are the same as for the years 1968–1978.

1981–1988

The source and processing of the data are the same as for the years 1968–1978.

In 1988, the residential status of 101 decedents was coded to unknown (code=9). According to the documentation, the State of Georgia did not disclose residency details for HIV/AIDS deaths, so U.S. residency was assumed in all such cases and those deaths were included in the database.

1989–2021

Death counts for all years starting with 1989 can be tabulated directly by Lexis triangles since both the age variable and the date of birth variable are available. When either the date of birth or the age is inconsistent or missing, the remaining information (including the exact date of death) is used to determine to which Lexis triangle a death should be allocated following the method described in Appendix 2.

In addition and per agreement with NCHS regulations, because of confidentiality issues, no count should be lower than 10 (and other than 0) in tabulations for national level data. Such counts are suppressed at the RDC and cannot be used for HMD calculations. Consequently, in these cases (which, so far, have only occurred at ages 100 or above), we used the publicly available mortality files from the NCHS (which do not include birth cohort information) to tabulate death counts by sex and single year of age. Death counts provided by single years of age were allocated to Lexis triangles following the standard HMD Methods Protocol.

POPULATION COUNT DATA

Coverage and Completeness

Data on population refer to the resident population of the United States. Every year, the US Census Bureau releases a new set of July 1st population estimates for all calendar years since the previous census, i.e. the so-called post-censal estimates. In addition, with each new census, a new (final) set of annual population estimates is produced for the period between the last two censuses, i.e. the inter-censal estimates. Whether inter-censal or post-censal, the Census Bureau annual estimates are calculated from adjusted census counts, birth and death counts, as well as estimates of net international migration counts.

The adjustments made to the census counts are carried out by the Census Bureau to account for coverage issues. For instance, the census typically undercounts children below the age of 17, especially those below age 4, and tends to overcount adults over the age of 50. The adjustments to Census counts are done based on the results of additional information. The primary sources of such information are the post-enumeration surveys (carried out following a census to verify which groups might have been missed or double counted) and demographic analysis relying on vital statistics, estimates of international migration, and other sources of population information such as data from the Social Security administration (more specifically data from the Centers for Medicare and Medicaid Services).

The most recent census conducted in 2020 was no exception and the Census Bureau concluded from its demographic analyses that, as in previous enumerations,

children had been undercounted (by as much as 5% under age 4) and that the population aged 50 years and over had been overcounted (by 0.55% for males and 2.63% for females). In addition, the census counts were found to exhibit evidence of age heaping starting at age 20 years² (Census Bureau, 2022). The results of these analyses are taken into account by the Census Bureau to produce reliable annual population estimates by sex and single year of age. The HMD series does not rely on the raw census data but on the adjusted annual population estimates as they become available each year from the Census Bureau.

Note that the currently available annual population estimates from the Census Bureau for the period 2010-2021 do not yet reflect the information from the 2020 Census. The 2010-2019 estimates are still post-censal estimates and the estimates for 2020 and 2021 are based on a blend of the Vintage 2020 Census Bureau estimates, demographic analysis, and the preliminary data from the 2020 Decennial census. The Census Bureau has announced the release of inter-censal estimates for 2010-2019 and a revised set of post-censal estimates for 2020 and 2021 (i.e. integrating the information from the final results of the 2020 Census) later on in 2023. These new estimates will be used for HMD mortality estimation when they become available.

Specific Details

For the years 1940-1969, custom population estimates have been computed based on published population estimates from various sources. Details are provided in Appendix 3.

BIRTH COUNT DATA

Coverage and Completeness

As for the mortality statistics, due to the legal requirements of birth registration, the data on births are considered to be virtually complete and of a good quality since 1959, with some limitations for prior years. For the years 1933–1959, the Human Mortality Database (HMD) includes births adjusted for under-registration. The adjusted births for the period from 1935 to 1950 were published by NCHS (see NCHS, 1961, for years 1935 and 1940-1959, and NCHS, 1954 for years 1936–1939). For the years 1933–1934, the adjusted births were estimated by applying published rates of under-report for both sexes combined to the unadjusted births by sex (NCHS, 1966). Both series of births are included in the raw data.

Specific Details

Birth data for the United States are confined to events registered in the United States. Births to legal residents of the U.S. that occurred in other parts of the world are excluded from published vital statistics. Prior to 1970, births to non-residents (that

² See in particular a plot produced by the Census Bureau that compares the age pyramids for the total US population by sex and single year of age using the 2020 Demographic Analysis study, the Vintage 2020 and Vintage 2021 Blended Base estimates, and the 2020 raw Census counts at <https://www.census.gov/content/dam/Census/newsroom/blogs/2022/rs-popestimates-covid-impacts/rsblog-popestimates-covid-impacts-figure1.jpg> [accessed April 19, 2022].

occurred in the U.S.) were included in the statistics (i.e., the *de facto* population), whereas for 1970 and thereafter, births to nonresidents are excluded (i.e., the data refer only to the *de jure* population).

The official vital statistics counts in 1959 exclude births that occurred in Hawaii. In order to be consistent with the geographic coverage for population and mortality statistics, the counts have been adjusted to include births in Hawaii.

REVISION HISTORY

Changes starting with the August 2015 revision:

Deaths: the HMD gained access to the NCHS restricted Detailed Mortality Files for years since 1989 through the University of California, Berkeley, Research Data Center. The restricted files provide more detailed information than the public file (in particular as regard the exact date of birth and the exact date of death for each record). These data have replaced the previous input mortality data (public files) and they have been used as described in this document.

Changes with the December 2017 revision:

Life tables: All life tables have been recalculated using a modified methods protocol. The revised protocol (Version 6) includes two changes: 1) a more precise way to calculate a_0 , the mean age at death for children dying during the first year of life and 2) the use of birth-by-month data (where and when available) to more accurately estimate population exposures. These changes have been implemented simultaneously for ALL HMD series/countries. For more details about these changes, see the revised Methods Protocol (at <http://v6.mortality.org/Public/Docs/MethodsProtocol.pdf>), particularly section 7.1 on Period life tables and section 6 and Appendix E, on death rates. The life tables calculated under the prior methods (Version 5) remain available at v5.mortality.org but will not be further updated.

ACKNOWLEDGEMENTS

We would like to thank Thomas Dunn and the entire staff at the Systems, Programming, and Statistical Resource Branch of the Division of Vital Statistics at NCHS for their assistance in compiling these data. Thanks as well to Robert Anderson and to Elizabeth Arias at NCHS for their help with earlier data collection for the precursor to the HMD—the Berkeley Mortality Database, and for their support to assemble more recent data. We also thank Cynthia Owens for the overall coordination of activities at Duke University.

REFERENCES

Census Bureau, Post-Enumeration Survey and Demographic Analysis help evaluate 2020 Census results, March 10, 2022. <https://www.census.gov/newsroom/press-releases/2022/2020-census-estimates-of-undercount-and->

overcount.html#:~:text=The%20PES%20found%20that%20the,not%20statistically%20different%20from%20zero.

National Center for Health Statistics. (1959–). Mortality Detail Files.

National Center for Health Statistics. (1939–). Vital Statistics of the United States, 1937–.

National Center for Health Statistics (1954). Vital Statistics of the United States, 1950. Volume 1. Table 6.06.

National Center for Health Statistics, (1961). Vital Statistics of the United States, 1959. Volume 1. Table 3-B.

National Center for Health Statistics, (1966). Natality Statistics Analysis, 1963. Series 21. No 8. Table 1 Live births, births rates, and fertility rates: United States, 1909 - 1965.

National Center for Health Statistics. (1968–). Multiple Cause of Death Public Use Tapes.

U.S. Census Bureau, U.S. Department of Commerce. Fifteenth Census of the United States: 1930. (1931). Population Volume I Number and Distribution of the Inhabitants.

U.S. Census Bureau, U.S. Department of Commerce. (1936–1938). Mortality Statistics. 1933–1936.

U.S. Census Bureau, U.S. Department of Commerce. Sixteenth Census of the United States : 1940. (1942). Population Volume I. Number Inhabitants.

U.S. Census Bureau, U.S. Department of Commerce. (1954). Series P-25, No. 98. Estimates of the Population of the United States and of the Components of change by Age, Color and Sex: 1940 to 1950.

U.S. Census Bureau, U.S. Department of Commerce. (1960–). Characteristics of the Population, United States: Summary.

U.S. Census Bureau, U.S. Department of Commerce. (1965). Current Population Reports. Population Estimates. Series P-25, No. 310. Estimates of the population of the United States, by Age, Sex, and Race: 1950 to 1960.

U.S. Census Bureau, U.S. Department of Commerce. (1974). Current Population Reports. Population Estimates. Series P-25, No. 519. Estimates of the population of the United States, by Age, Sex, and Race: April 1, 1960 to July 1, 1973.

U.S. Census Bureau, U. S. Department of Commerce. (2001). Summary File 1, 2000 Census of Population and Housing, Technical Documentation. U.S. Census Bureau (<http://www.census.gov>).

U.S. Census Bureau, U. S. Department of Commerce. POPULATION - U. S. POPULATION ESTIMATES BY AGE, SEX, AND RACE: 1900-1979 (PE-11). DSK.POP00/79-PE11

U.S. Census Bureau. Population Estimates Program (<http://www.census.gov/popest/estimates.php>).

APPENDIX 1:

Description of the original data used for HMD calculations

DEATHS

Period	Type of Data	Age groups	Comments	RefCode(s) [†]
1933– 1950	Annual death counts for the <i>de facto</i> population by 5-year age group	0,1...5-9 ... 100+		1, 5, 7, 9, 11, 13, 15, 25, 27, 29, 31, 33, 35, 37, 39, 41, 46, 49
1951– 1952	Annual death counts for the <i>de facto</i> population by 5-year age group to age 85 and by single year of age for ages 85 and older	0,1...5-9 ...80-84, 85, 86,...100+		53, 55
1953– 1958	Annual death counts for the <i>de facto</i> population by 5-year age group to age 85 and by single year of age for ages 85 and older	0,1...5-9 ...80-84, 85, 86,...maximum age attained		57, 59, 63, 65, 67, 69
1959– 1969	Annual death counts for the <i>de facto</i> population by single year of age and year of birth (Lexis triangles)	0, 1,...maximum age attained	Deaths have been tabulated from individual death certificates data. [‡]	79, 89, 93, 97, 101, 105, 109, 113, 117, 123, 129
1970– 1988	Annual death counts for U.S. residents by single year of age and year of birth (Lexis triangles)	0, 1,...maximum age attained	Deaths to U.S. residents occurring in outlying territories (e.g., Puerto Rico, U.S. Virgin Islands) or a foreign country (including Canada) are excluded, as are deaths to non-residents. Deaths have been tabulated from individual death certificates	133, 137, 141, 145, 149, 153, 157, 161, 165, 169, 175, 179, 183, 187, 191, 195, 199, 203, 207

Period	Type of Data	Age groups	Comments	RefCode(s) [†]
			data. [‡]	
1989–2021	Annual death counts for U.S. residents by single year of age and year of birth (Lexis triangles) or by single year of age [‡]	0, 1,...maximum age attained with 10+ deaths; complemented with deaths by single year of age from the publicly available mortality files for higher ages as described in this document	Deaths to U.S. residents occurring in outlying territories (e.g., Puerto Rico, U.S. Virgin Islands) or a foreign country (including Canada) are excluded, as are deaths to non-residents. Deaths have been tabulated from individual death certificates data. [‡]	212, 214, 216, 218, 220, 222, 224, 226, 228, 230, 236, 238, 240, 242, 244, 246, 248, 250, 252, 253, 254, 256, 260, 264, 266, 270, 277, 278, 280, 281, 285, 286, 292, 293, 296, 298, 299, 301, 303, 306

† The reference code is used in the raw data files (Input Database) to link data with sources.

‡ For details, see Appendix 2.

POPULATION

Period	Type of Data	Age groups	Comments	RefCode(s)
1933–1939	Annual population estimates for the resident population	0, 1 ... 75+		19
1940–1949	Annual population estimates for the resident population	0, 1 ... 85+	Custom computations based on population available from different sources [†]	45
1950–1958	Annual population estimates for the resident population	0, 1 ... 85+	Custom computations based on population available from different sources [†]	71
1959	Annual population estimates for the resident population	0, 1 ... 85+	Custom computations based on population available from different sources [†]	85

Period	Type of Data	Age groups	Comments	RefCode(s)
1960–1969	Annual population estimates for the resident population	0, 1 ... 85+	Custom computations based on population available from different sources [†]	125
1970–1979	Intercensal population estimates for the resident population	0, 1 ... 85+		171
1980–1999	Intercensal population estimates for the resident population [‡]	0, 1 ... 99, 100+		287, 288
2000–2009	Intercensal population estimates for the resident population [‡]	0, 1 ... 85+		289
2010–2019	Inter-censal population estimates for the resident population ^{‡‡}	0, 1 ... 100+		302
2020–2021	Post-censal population estimates for the resident population ^{‡‡}	0, 1 ... 100+		305

† For details, see Appendix 3.

‡ Data are available on the Census Bureau web site (<http://www.census.gov>). For the specific URLs and download dates, see the reference file for the raw data.

‡‡ Data are available on the Census Bureau web site:

<https://www.census.gov/programs-surveys/popest/data/tables.html>

BIRTHS

Period	Type of Data	Comments	RefCode(s) [†]
1933–1969	Annual births for the <i>de facto</i> population by sex	Counts for 1959 have been adjusted to include births that occurred in Hawaii (see section “Births Count Data”)	3, 23, 75, 87, 91, 95, 99, 103, 107, 111, 115, 121, 127
1970–2021	Annual births for the resident population by sex	Births to U.S. residents that occurred abroad are excluded as are births to non-residents.	131, 135, 139, 143, 147, 151, 155, 159, 163, 167, 173, 177, 181, 185, 189, 193, 197, 201, 205, 209, 211, 213, 215, 217, 219, 221, 223, 225, 227, 229, 235, 237, 239, 241, 243, 245, 247, 249, 251, 255, 259, 263, 269, 275, 282, 284, 291, 295, 300, 304, 307

BIRTHS BY MONTH

Period	Type of Data	Comments	RefCode(s) [†]
1933-2010	Annual number of live births by month	Downloaded from HFD (http://www.humanfertility.org/cgi-bin/getfile.plx?f=USA\20130206\INP UT\USAmohtly.txt&c=USA&tab=in db)	17, 21, 43, 47, 51, 61, 73, 77, 119, 257
2011-2021	Annual number of live births by month	Births to U.S. residents that occurred abroad are excluded as are births to non-residents	267, 275, 282, 284, 291, 295, 300, 304, 307

Appendix 2:

Tabulation of 1959-2020 Deaths from the Mortality Detailed Files by Lexis triangle

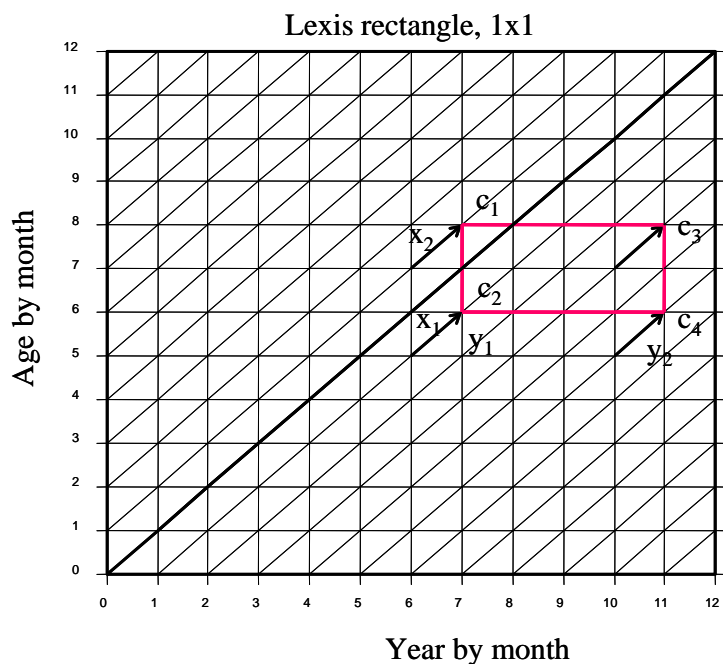
The method described here was used to derive death counts by Lexis triangle from the NCHS Detailed Mortality Files. As noted in the main text, for years 1989 to 2019, tabulation of deaths by Lexis triangle (age and birth cohort) is possible using the information available in the restricted files made available to us by NCHS through the Berkeley Research Data Center. When the information is inconsistent or missing, it is also possible to use the exact date of death when available to derive death counts by Lexis triangle. The methods used in such cases are described below. They also apply to all years before 1989 when the exact age at death (day/month) or when the month of death are included in the data file. When neither the year of birth, nor the exact date of death (in months or in months and days) are available, the standard HMD methods are applied to deaths tabulated by single year of age.

More specifically, if the date of birth is identified as erroneous because it takes place later than the date of death (as was the case for nearly 300 records out of a total 2,150,466 records of legal residents in 1989), then we use the completed age and the exact date of death to estimate the Lexis triangle as described below. If the completed age and the exact date of birth are inconsistent, then we assume that age is erroneous and we use the date of birth. If either the completed age or the date of birth is missing, we use the remaining information using the method described below. In the very few instances when age is missing and the date of birth and the date of death coincide, the records are split with equal (0.5) weight between the two triangles of the possible cohort-period parallelogram. To summarize, in all cases when the date of birth, the date of death and the age are inconsistent, lacking in details, or missing altogether, all remaining information (including the exact date of death) is used to determine to which Lexis triangle a death is allocated using the following method.

GENERAL APPROACH

Figure 1 illustrates a 1x1 Lexis square (by age and year of death), divided into two Lexis triangles (by age, birth cohort, and year), each of which is further divided into 144 smaller Lexis triangles (by age in months, birth month, and month of death). Suppose that we know a death occurred between July and November (months 7-11) and between age ($x + 6$ months) and ($x + 8$ months) (shown by the red rectangle in Figure 1). The red rectangle includes 16 of the smaller Lexis triangles (by month), of which 15 belong to the lower Lexis triangle (by year) and one to the upper triangle. If we assume that the probability of dying is the same in each of the smaller Lexis triangles, then the probability that such deaths occurred to someone from the older cohort is $1/16$. Therefore, among all individual death records that fall within this red rectangle, we assign $1/16^{\text{th}}$ of such deaths to the upper triangle and $15/16^{\text{th}}$ to lower triangle. This simple example captures the gist of the method.

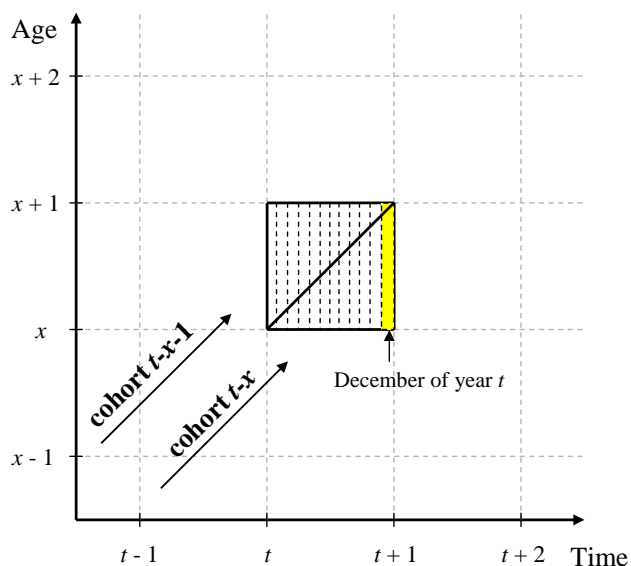
Figure 1. 1x1 Lexis Square by age (in months) and month/year of death



MONTH OF DEATH AVAILABLE BUT NO DATE OF BIRTH

For some years (1959–1961 and 1968–1971), only the month and year of death are included in the mortality files (not the day of death or the date of birth). Therefore, within a given 1x1 Lexis square (age by calendar year), we can further split the deaths into 12 rectangles representing the month of death (age by month/year of death), as shown on Figure 2. The proportion of deaths falling within the upper and lower triangles of each rectangle can be computed assuming a uniform distribution of deaths. For example, for deaths occurring in December of year t (shown in yellow on Figure 2), $23/24^{\text{th}}$ would fall into the lower triangle and $1/24^{\text{th}}$ in the upper triangle.

Figure 2. Illustration of Lexis Triangles and Rectangles



DAY AND MONTH OF DEATH AVAILABLE BUT NO DATE OF BIRTH

For some other years (1962-1967 and 1972-1988), the files include the complete date of death information but not the date of birth. Therefore, a procedure similar to that described in the previous section is applied. That is, deaths within each 1x1 Lexis square are split into 365 rectangles representing each possible day of death. Again, assuming that deaths are distributed uniformly within each of these rectangles, the proportion of deaths falling within the upper and lower triangles can be computed.

The method is also used to tabulate deaths by Lexis triangles in all years 1989 and later when the date of birth is either inconsistent (with age and year of death) or missing.

APPENDIX 3:

Custom population estimates for 1940-1969

One of the guiding principles of the HMD is to provide mortality estimates with as much age detail as possible. U.S. data on deaths by single year of age are available starting with 1959. To compute death rates by single year of age, they must be combined with population counts by single year of age. Such data are available from the U.S. Census Bureau, but for the period 1940–1969, these population counts (from report PE-11) include the Armed Forces overseas whereas the death counts exclude deaths that occurred overseas. Thus, the population estimates for 1940-1969 have been adjusted to exclude the Armed Forces overseas using a three-step process described below.

For the period 1940–1949, resident population estimates by 5-year age groups are available from report P25-98 (U.S. Census Bureau, 1954). In the first step, population estimates for the Armed Forces overseas (by 5-year age groups) were computed by aggregating the counts including the Armed Forces overseas (from PE-11) into 5-year age groups and subtracting population estimates of residents (from P25-98). Second, the estimates obtained in step one were distributed into single years of age using the cubic spline interpolation procedure described in the HMD Methods Protocol. Third, these estimates were subtracted from the population counts in PE-11 to derive the resident population by single year of age.

For the periods 1950–1959 and 1960–1969, estimates are adjusted in a similar way using data for the resident population from reports P25-310 (U.S. Census Bureau, 1965) and P25-519 (U.S. Census Bureau, 1974), respectively. In order to be consistent with the territorial coverage of death counts, the population estimates for 1950-58 were also adjusted to exclude Alaska and Hawaii. The population counts from PE-11 (by single year of age) include the Armed Forces overseas, Alaska, and Hawaii, whereas P25-310 provides two sets of counts for the resident population (by 5-year age groups): one *including* Alaska and Hawaii, and the other *excluding* these two states. Therefore, the three step-process described above was implemented, except that for 1950-1958 the first set of counts from P25-310 was used in step one; this procedure resulted in estimates for the resident population excluding Alaska and Hawaii during 1950-1958. For 1959, the second set of counts from P25-310 was used in step one, thereby producing estimates for the resident population including these two states.