REPORTED DEATHS OF CENTENARIANS AND NEAR-CENTENARIANS IN THE U.S. SOCIAL SECURITY ADMINISTRATION'S DEATH MASTER FILE

K. Faig

And the Lord said,
My spirit shall not always strive with man,
For that he also is flesh:
Yet his days shall be an hundred and twenty years.
--Genesis 6:3 (King James Bible)

The sum of a man's days is great
If it reaches a hundred years:
Like a drop of sea water, like a grain of sand,
So are these few years among the days of eternity.
--Sirach (Ecclesiasticus) 18:7-8 (New American Bible)

The days of our years are threescore years and ten;
And if by reason of strength they be fourscore years,
Yet is their strength labour and sorrow;
For it is soon cut off, and we fly away.
-- Psalm 90:10 (King James Bible)

Abstract

The Death Master File (DMF) maintained by the Social Security Administration (SSA) is the largest collection of publicly-accessible death records in the United States. As of December 2000, the public release version of the DMF (DMF-PR) contains 65,089,493 death records for deaths occurring between 1900 and 2000. Of these records, 1,658,559 relate to decedents aged 95-99; 319,975, to decedents aged 100-104; 29,800 to decedents aged 105-109; 2,439 to decedents aged 110-114; 773, to decedents aged 115-119; 265, to decedents aged 120-124; 124, to decedents aged 125-129. There are scattered records for even higher reported ages at death.

This article discusses the sources of the information in the DMF-PR and the potential utility of this information for demographic studies of centenarians and near-centenarians. While the United States will not celebrate the centennial of the completion of its birth registration area until 2033, geographically specific centenarian samples drawn from the DMF-PR can be validated using early twentieth-century censuses and other resources.

Table of Contents

Abbreviations Utilized In This Paper

BLF, Black Lung File
BY, Birth Year
CDC, [United States] Center for Disease Control and Prevention
DACUS, Death Alert, Control and Update System
DHEW, [United States] Department of Health, Education and Welfare
DMF, Death Master File
DMF-PR, Death Master File-Public Release Version
DY, Death Year
HCFA, [United States] Health Care Financing Administration
MBR, Master Beneficiary Record
NAPHSIS, National Association for Public Health Statistics and Information Systems
NDI, National Death Index
NCHS, [United States] National Center for Health Statistics
NTIS, National Technical Information Service (U.S. Department of Commerce)
OIG, Office of the Inspector General (U.S. Social Security Administration)
SER, Summary Earnings Record
SSA, [United States] Social Security Administration
SSN, Social Security Number

History of DMF-PR

From its inception in 1935, SSA has had as its primary mission accurate and timely payment of the benefits provided under the Social Security Act. These include not only primary old age benefits but also disability benefits and survivors' benefits to widows and widowers and minor children of decedent insureds. A secondary mission has been the protection of the program and its insureds from fraudulent claims made by non-qualified individuals. For example, a representative payee might attempt to continue to receive a deceased insured's benefits.[1] While secondary to the maintenance of accurate earnings records and benefit payment records, the maintenance of death records relating to insureds has been an important part of SSA's record-keeping mission from inception.

Aziz and Buckler (1992) provide a useful history of the maintenance of automated death information by SSA. Manual registers commenced in 1937 were automated to create the DMF in 1962. There is a scattering of records for deaths from the period 1900-1936 but they total only 1606 deaths, less than 0.0025% of the total deaths in the December 2000 DMF-PR. Since program inception the largest portion of death reports have been filed by funeral directors using form SSA-721. SSA also receives reports of deaths from relatives, friends, financial institutions and postal authorities; these sources together account for approximately 95 percent of deaths reported to SSA. However, Section 205(r) of the Social Security Amendments of 1983 authorized the Secretary of HHS to contract to receive death records from state bureaus of vital statistics. Death information which SSA receives solely from state bureaus of vital statistics is excluded from DMF-PR.

[1] SSA maintains a separate file of representative payee records, the Master Representative Payee File (MRPF). SSA OIG (Office of Inspector General) Audit Report A-01-98-61009 (September 1999) concerned the detection of deceased representative payees.

For many years death information obtained by SSA was posted both to the SER (Summary Earnings Record) and the MBR (Master Beneficiary Record). The posting of death information to the SER stopped in November 1977. However, death information derived from claims for the Social Security lump sum death benefit[2] began to be posted to MBR in 1977. In addition, SSA began matching Medicare death records obtained from HCFA in 1982. Since 1988, the NUMIDENT file (where all applications for Social Security Numbers [SSNs] are recorded) has been the official repository of all SSA death data. Death information obtained from all sources is consolidated in the NUMIDENT file; DACUS (Death Alert, Control and Update System) matches these records against the MBR and the SSR (Supplemental Security Record) to produce alerts for validation if a decedent is in current payment status on these files. As of December 2000, the DMF-PR also contains 544,050 death records supplied by the Railroad Retirement Board.

The first DMF-PR was created by collecting death records from the MBR, the SER, the SSR, and the BLF (Black Lung File). Since 1979, the National Center for Health Statistics (NCHS) of the Center for Disease Control and Prevention (CDC) has received death record information from state bureaus of vital statistics to create the National Death Index (NDI). NCHS has spearheaded a program for the automation of death records in cooperation with state bureaus of vital statistics. In 1999, SSA entered into a contract with NCHS and NAPHSIS (National Association for Public Health Statistics and Information Systems) to start developing a national electronic death registry. The ultimate objective is to receive death information within twenty-four hours of receipt by state bureaus of vital statistics.[3] The information in the NDI[4] is not available to organizations or to the general public for legal, administrative or genealogical work. It is available solely for statistical purposes in medical and health research. Names, SSNs, dates of birth and other related information can be submitted to NCHS in electronic form or on NCHS coding sheets once a research project is approved for death matching against the NDI. As of January 2000, the fees for routine NDI searches consist of a $350 service fee plus $0.30 per record per year searched. The NDI is updated annually with data received from state bureaus of vital statistics, about ten months after the close of each calendar year.

The DMF-PR is made available for private purchase through the NTIS (National Technical

---

[2] Research Note #2, "The History & Development of the Lump Sum Death Benefit," issued by the SSA Historian's Office in June 1996. An electronic edition is available at http://www.ssa.gov.history/lumpsum.html. Myers (1975) is an excellent reference work for the historical development of all SSA programs.

[3] Progress of EDRS (electronic death registration system) may be monitored on the NAPHSIS Internet site: http://www.naphsis.org/main/EDRS.

[4] Information concerning NDI may be found on the NCHS Internet site at http:// www.cdc.gov/nchs/r&d/ndi.htm.

Information Service) of the U.S. Department of Commerce.[5] The current subscription price for the DMF-PR with quarterly updates as of January 2001 is $6900. Subscribers who intend to maintain current data are also required to subscribe to a monthly update service which costs an additional $2760 per year. However, a number of private purchasers of the DMF-PR have made the data available to the public without charge through the Internet. While the primary usage of the DMF-PR is for probate and genealogical work, medical and health researchers also use it for deaths prior to the commencement of NDI in 1979.

The author used the Rootsweb SSDI Death Index[6] for his research; its advanced search function allowed ready surveying of the records by calendar year of birth (BY) and calendar year of death (DY). Had the necessary funding been available, purchasing a single issue of DMF-PR (current price $1725) would have enabled the author to import the death information into a SQL database and to write his own research queries.

The DMF-PR is an EBCDIC file with 60-byte records containing SSN, name, date of birth, date of death[7], state or country of residence (discontinued for deaths after February 1988), ZIP code of last residence, and ZIP code of lump sum payment. However, a significant number of records do not contain information in some of the fields. The Rootsweb SSDI site helpfully displays the locality, county and state corresponding to the ZIP code fields when they are present. It also displays the state of issuance of the decedent's SSN based on its prefix. Users attempting to obtain death certificates from state bureaus of vital statistics have found that there are sometimes inaccuracies in the DMF-PR; for example, the ZIP code where a representative payee resided may be shown instead of the ZIP code where a decedent insured resided.

---

[5] The subscription terms for DMF-PR may be found on the NTIS Internet site at: http://www.ntis.gov/product/ssa-death-master.htm. There is a link to a pdf which displays the layout of the DMF-PR file.

[6] Internet address: http://ssdi.genealogy.rootsweb.com/cgi-bin/ssdi.cgi.

[7] Many records contain only month and year of death and lack day of death. Thus it is impossible to determine completed years of life (age last birthday at death) exactly for some decedents who die in their birth month.

Table 1 compares SSA DMF-PR deaths with NCHS NDI deaths for each year from the inception of the NDI in 1979 through 1999.[8] The comparison is not complete because DMF-PR contains some deaths of insureds in foreign jurisdictions which are not included in NDI. However, during the period 1979-1999, DMF-PR has on average included about nine of every ten deaths included in NDI.

The SSA Office of the Inspector General (OIG) conducts periodic reviews of SSA operations and issued Evaluation Report A-09-98-61011 "Improving the Usefulness of Social Security Administration's Death Master File" [SSA-OIG (2000)] on July 28, 2000. This report discusses SSA's 1998 comparison of death information in the MBR and NUMIDENT files. As a result of this comparison, information relating to 0.9 million deaths identified in MBR was added to NUMIDENT. However, 1.3 million deaths found in MBR were not added to NUMIDENT because of data discrepancies (62% name, 32% date of birth, 5% gender).

In its report OIG recommended that SSA undertake the reconciliation necessary to add these 1.3 million unmatched deaths to DMF. It also recommended that SSA annotate the DMF to indicate which death reports have been verified by SSA prior to awarding or terminating benefits. SSA did not agree with OIG's recommendation that the remaining unmatched MBR deaths be reconciled with DMF but did accept the inclusion of the verification identifier in the DMF as part of its development plan.

Program integrity is the primary mission of SSA. Creating records which are useful to private individuals for genealogical, probate or demographic work is only a by-product of SSA's pursuit of its primary mission.

Centenarian and Near-Centenarian Deaths in DMF-PR

From 300,524 deaths reported in 1962 (the first year of automated record-keeping), death reports in DMF-PR have grown steadily, first exceeding one million deaths in 1966. Manually maintained records converted in 1962 grew from 630 in 1937 to 98,992 in 1961. There is a small scattering of earlier deaths for the period 1900-1936 which constitute fewer than 0.0025% of the total records in the December 2000 DMF-PR. Figure 1 displays the growth of death reporting in the DMF-PR in graphical form.

---

[8] NDI total death counts for 1979-1998 were obtained from the NCHS Internet site at http://www.cdc.gov/nchs/data/gm292-1.pdf. A provisional death count for 1999 was obtained from the article "Births, Marriages, Divorces and Deaths: Provisional Data for 1999" in the Center for Disease Control and Prevention's *National Vital Statistics Reports* (vol. 48 no. 19) dated February 22, 2001.

For simplicity in using the Rootsweb advanced query function, the author computed age at death (Y) as calendar year of death (DY) less calendar year of birth (BY). For Y = DY-BY = N, the youngest decedents (date of birth December 31, BY; date of death January 1, BY+N) die at age N-1 years 1 day and the oldest decedents (date of birth January 1, BY; date of death December 31, BY+N) die at age N years 364 days. The completed years of life (CY) for all these decedents satisfies the relationship N-1 < CY < N+1.

For example, the tabulation for Y = 100 includes persons who died between ages 99 years 1 day and 100 years 364 days. These deaths will average age 100 if both births and deaths are distributed uniformly in their respective years. There is probably no reason to expect that births are not distributed uniformly within each BY cohort. However, at advanced ages, it may be expected the deaths will be heaped toward the beginning of DY for each BY cohort. Thus, the actual average age of persons born in BY dying in DY = BY + 100 is probably less than 100 years.

For convenience, the author will refer to deaths occurring at Y = 100-129 inclusive as centenarian deaths and to deaths occurring at Y = 95-99 as near-centenarian deaths. The author will refer to Y as "age" throughout this paper.

Table 2 displays total deaths and deaths in seven attained age groups (95-99, 100-104, 105-109, 110-114, 115-119, 120-124 and 125-129) organized by quinquennial DY cohorts from 1900-1904 through 1995-1999. DY = 2000 is displayed separately because of the incompleteness of reporting in the December 2000 DMF-PR. The irregular pattern of deaths in the extreme age groups 115-119, 120-124 and 125-129 is notable. Deaths at 120-124 and 125-129 are heavily heaped in the period DY = 1970-1974 while deaths at 115-119 are heavily heaped in the period DY = 1975-1984.

Table 3 displays deaths in each of the seven attained age groups as a percentage of the total deaths observed at attained ages 95-129. For all years of death combined, deaths at 95-99 represented 82.44% of total deaths of centenarians and near-centenarians; deaths at 100-104, 15.90%; deaths at 105-109, 1.48%; deaths at 110-114, 0.12%; deaths at 115-119, 0.04%; deaths at 120-124, 0.01%; and deaths at 125-129, 0.01%.

Table 4 examines deaths in each centenarian attained age group as a percentage of deaths in the next lower group. That ages in the highest attained age groups are severely overstated is clear from the relationships displayed in this table: 31.69% of attained age 110-114 deaths at attained ages 115-119; 34.28% of attained age 115-119 deaths at attained ages 120-124; and 46.79% of attained age 120-124 deaths at attained ages 125-129. The ratios are affected only modestly if the DY = 1900-1964 cohorts largely antedating the keeping of automated death records beginning in 1962 are excluded.

Table 5 distributes the total deaths reported in each of the seven attained age groups over the quinquennial DY cohorts. For each attained age group, the reported deaths in each quinquennial DY cohort are displayed as a percentage of the total reported deaths for that attained age group. Note that

the reporting of deaths for DY = 2000 in the December 2000 DMF-PR is incomplete. As in Table 2, the heaping of deaths at 115-119 in the DY = 1975-1984 period and the heaping of deaths at 120-124 and at 125-129 in the DY = 1970-1974 period may be noted.

Figure 2 provides a graphical display of deaths recorded in the December 2000 DMF-PR by BY. The earliest BY in the DMF-PR is 1800, but deaths first exceed 100 per year in BY = 1857; 1,000 per year in BY = 1864; 10,000 per year in BY = 1869; 100,000 per year in BY = 1876; and 1,000,000 per year in BY = 1892. In the December 2000 DMF-PR, BY = 1908 has the largest number of deaths (1,646,671). Deaths fall back below 1,000,000 for BY = 1924; below 100,000 for BY = 1961; and below 10,000 for BY = 1983.

Death reporting in the DMF-PR for the pre-retirement ages is substantially less complete than for the post-retirement ages because deaths reported to NUMIDENT solely by state bureaus of vital statistics are excluded from DMF-PR. Exclusions include most juvenile deaths (except for juveniles receiving survivor benefits) and some pre-retirement adult deaths not qualifying for payment of the lump sum death benefit. Since 1981, payment of the $255 lump sum death benefit has been restricted to decedent insureds leaving surviving spouses or minor children. Some 121,335 deaths reported in the December 2000 DMF-PR are reported with BY = 0000 to indicate unknown BY and are not included in the graphical display.

Table 6 displays deaths in the seven centenarian and near-centenarian attained age groups by quinquennial BY cohorts from 1800-1804 through 1900-1904. BY = 1905 is displayed separately because of the incompleteness of attained age 95 death reports in DY = 2000 in the December 2000 DMF-PR.

Based upon the population increases and the mortality improvements experienced in the United States during the twentieth century, the December 2000 DMF-PR death count at attained ages 115-119 should be greatest for the BY = 1875-1880 cohort; at ages 120-124, greatest for the BY = 1870-1874 cohort; and at ages 125-129, greatest for the BY = 1865-1869 cohort. In actual fact, the respective maximum death counts actually occur in the BY = 1860-1864 cohort for attained ages 115-119; in the BY = 1850-1854 cohort for attained ages 120-124; and in the BY = 1845-1849 cohort for attained ages 125-129. This phenomenon indicates that the death counts in the highest attained age groups are suspect for the early BY cohorts.[9]

[9] The tiny population of the early BY cohorts in DMF-PR fails to support the alternative hypothesis that survival into the highest attained age groups has actually begun to decline in the United States.

Table 7 displays total centenarian and near-centenarian deaths by quinquennial BY cohort and the deaths in each attained age group as a percentage of total centenarian and near-centenarian deaths. The deaths in the BY = 1800-1874 cohorts are heaped much more markedly in the centenarian attained age groups than deaths in BY = 1875-1905 cohorts.

Table 8 displays the deaths for each quinquennial BY cohort in each of the centenarian attained age groups, as percentages of deaths in the next lower attained age group. The same phenomenon of over-contribution of the earliest BY cohorts to deaths at the highest attained groups observed in Table 6 is observed in this table.

Table 9 distributes the total deaths in each of the seven attained age groups over the quinquennial BY cohorts. For each attained age group, the reported deaths in each quinquennial BY cohort are displayed as a percentage of the total reported deaths for that attained age group. In every attained age group above 100-104, the maximum representation by quinquennial BY cohort occurs substantially earlier than expected.

It is fortunate that the early BY cohort component of the December 2000 DMF-PR is limited: 227,760 deaths (fewer than 0.35%) of the total DMF-PR deaths for BY = 1800-1874 cohorts and 5,610 deaths (or fewer than 0.001%) for the total DMF-PR deaths for BY = 1800-1864 cohorts. However, these sparsely populated early BY cohorts contribute disproportionately to the reported deaths in the highest attained age groups.

Outlier Records

The author did not search DMF-PR systematically for Y > 130. A sparse population of such records does exist in the December 2000 DMF-PR. Single instances of Y = 130 may be found in BY = 1867, 1854, 1852, 1851, 1847, 1846, 1841, 1839, 1838, 1837, 1835, 1833 and 1827, while BY = 1831 and 1829 offer two instances each, and BY = 1844, 1843 and 1828, three instances each. Altogether, the December 2000 DMF-PR contains 26 deaths at attained age Y = 130.

Since the sparsely populated early BY cohorts contribute a disproportionate number of deaths at the highest attained age groups, the author examined the BY = 1808 cohort in detail for all possible ages at death from Y = 92 (DY = 1900) to Y = 192 (DY = 2000). The results of the examination are shown in detail in Table 10.

It will be noted that 31 of the 53 deaths for BY = 1808 are concentrated in the twenty-five year period DY = 1948-1972 (135 <= Y <= 164). The single death for DY = 2000 and BY = 1808 [SSN 568-96-5690 (born October 28, 1808, died March 13, 2000)], a claim of 191 years 4 months 13 years, is certainly a remarkable outlier, exceeding even the 185 years claimed for the Hungarian centenarian Petracz Czartan (died 1724) in the eighteenth century.[10]

[10] For a discussion of Czartan's claimed 185 years, see L.-L.B. Petersen and B. Jeune, "Age

The 33 deaths for BY = 1800 (the lowest year of birth in DMF-PR) contain 10 deaths from the period 1900-1906. However, none of the remaining 23 deaths are earlier than DY = 1975. The oldest of these extreme records for BY = 1800 is for SSN 194-52-7259 (born December 11, 1800, died September 7, 1992), a claim of 191 years 10 months 27 days. It seems very likely that the 23 deaths with BY = 1800 and DY >= 1975 result from simple BY miscodes.

Historically many computerized systems specified 1800 as the low permissible value for calendar year fields and it is possible that 1800 was coded when 0000 ought to have been coded for unknown year of birth. Alternatively, the 23 extreme records with BY = 1800 and DY >= 1975 may result from an erroneous conversion from two-position to four-position years. The same possibility of errors interjected through miscoding or erroneous data conversions may explain many of the other extreme ages at death reported in the early nineteenth-century BY cohorts in DMF-PR.

The author examined decennial death years DY = 1940, 1950, 1960, 1970, 1980, 1990 and 2000 for all BY cohorts and found decreasing frequency of extreme records with Y > 130 with increasing DY from 1960 onward. The results of this examination are shown in detail in Table 11.

Validation in the Luxdorph Gallery" in Jeune and Vaupel (1999), pp. 45-47. Bolle Luxdorph (1716-1788) was a Danish civil servant who collected portraits of centenarians. Twentieth-century actuary W. G. Bowerman collected photographs of centenarians; it is to be regretted that he was unable to publish any of these photographs in his pioneering article [Bowerman (1939)] on centenarians.

The data quality problems found with the extremely sparse records for DY < 1937 and for BY < 1865 may cause many researchers to reject these records out of hand. Payroll deductions under the 1935 Social Security Act began as of January 1, 1937. Until the 1939 amendments, participation was limited to workers aged under 65 years. Therefore, DMF-PR records with DY < 1937 or BY < 1865 appear to be of problematic origin.[11]

Outliers will inevitably remain a fascination. There is valid demographic reason to pursue cases like Jeanne Calment (1875-1997) and Christian Mortensen (1882-1998)[12]; the validation of such cases leads demographers to seek other survivors of comparable age. While it may be true that it is harder to validate the age claim of a decedent centenarian than it is to validate the age claim of a living centenarian, nevertheless significant progress can be achieved based upon decedent records. The age claim of French-Canadian Pierre Joubert (died 1814), long considered a probable supercentenarian by authorities such as Bowerman (1939) and McWhirter (1977), was disproved in 1990 by the remarkable work of Charbonneau.[13]

How many true supercentenarians may be hiding in the ranks of the deaths reported at ages 110 and higher in DMF-PR as of December 2000? Is it possible that one or more exceeded even the proven 122 years 5 months and 14 days of the French supercentenarian Jeanne Calment? While no proof in the domain of human longevity can be regarded as "certain," one must posit a bizarre substitution of persons or falsification of vital records subject to numerous difficulties to deny Jeanne Calment's claimed lifespan. The author believes that her duration of life may safely be regarded as proven "beyond reasonable doubt"--the highest standard we can hope to achieve in such matters.[14]

---

[11] When SSA death data was automated in 1962, records for prior deaths in the paper registers were created only for deaths with active continuing survivors' benefits. All of the data in DMF-PR derive from SSA and RRB (Railroad Retirement Board) sources; there is no known component from outside sources (e.g., government pensioners or armed forces veterans). The oldest deaths in DMF-PR derive from the MBR. The author is indebted to Bert Kestenbaum of SSA for this information.

[12] For Calment, see J.-M. Robine and M. Allard, "Jeanne Calment: Validation of the Duration of Her Life" in Jeune and Vaupel (1999). For Mortensen, see A. Skytthe, B. Jeune and J. Wilmoth, "Age Validation of the Oldest Man" in Jeune and Vaupel (1999).

[13] H. Charbonneau's discoveries relating to Pierre Joubert, published in French-language genealogical publications in 1990-91, are cited by B. Desjardins, "Validation of Extreme Longevity in the Past: The French-Canadian Experience" in Jeune and Vaupel (1999), pp. 66-67.

[14] One extreme long-liver for whom the author has failed to find any comprehensive discussion of age validation is U.S. government pensioner Mark Thrash, cited by Robert J. Myers in his published discussion of Bowerman (1939). Thrash claimed a date of birth of December 25, 1820 when he retired in 1922. He was employed by the United States government between 1894 and 1922 and was living at Chickamauga National Park, Georgia when Myers published his discussion. He gave his date of birth as

Gould (1929), who gave some credence to 152 years claimed for Thomas Parr (died 1635)[15], provided an interesting analogy between extreme longevity and extreme stature. He compared the claimed height (9 feet 3 inches) of the Russian giant Fyodr Machnov (1880-1905)[16] with a "high" adult stature of 6 feet and applied this ratio (111 inches to 72 inches) to the 100-year maximum lifespan referred to in Sirach (Ecclesiasticus) 18:7 to obtain an outlier longevity estimate of 154 years. A fairer comparison would probably have been to the 80-year exceptional lifespan cited in Psalms 90:10 which would have resulted in a extreme lifespan estimation of 120 years [(9/6) x 80 years], close to the maximum lifespan which has to date been scientifically proven in modern times. Using instead the ordinary 70-year lifespan cited in the same psalm verse, we obtain an extreme lifespan estimation of 105 years [(9/6) x 70 years], which may represent quite a reasonable estimation of maximum human lifespan for the centuries prior to the nineteenth.[17]

Adapting Gould's arguments, we may compare a normal adult woman's height (63.7 inches) with the extreme observation (ca. 8 feet or 96 inches); applying this ratio (96.0/63.7) to a woman's current life expectancy at birth in the United States (79.5 years), we obtain 119.8 years. Performing the same computation for the opposite sex, we may compare a normal man's height (69.1 inches) with the extreme observation (ca. 9 feet or 108 inches); applying this ratio (108.0/69.1) to a normal man's life

---

December 1822 and his state of birth as Georgia when he was enumerated in Walker County, Georgia in the 1900 federal census. Thrash died on December 17, 1943 near Chattanooga, TN. One must probably favor 1822 over 1820 as his true year of birth based on the 1900 census schedule; in which case he died just short of his 121st birthday, missing the years of Jeanne Calment by nearly a year and a half but exceeding the years of Christian Mortensen by more than six years. Although both Myers and Gould were impressed by evidence in support of Mark Thrash's claimed age, the 1900 census record for Thrash is a late-life record; some earlier record of his life must be sought before his case can be considered validated. A search for Thrash in census records before 1900 is certainly indicated since he claimed birth in Georgia

[15] An appendix provides some discussion of Gould's remarkable outlier Thomas Parr.

[16] McWhirter (1977) (p. 15) debunked Machnov's claim and estimated his true stature at 7 feet 9.7 inches but giants very close to 9 feet in height have been verified. Robert Pershing Wadlow (born February 22, 1918, Alton, IL; died July 15, 1940, Manistee, MI; buried, Oakwood Cemetery, Alton, IL) was scientifically measured at 8 feet 11.1 inches by Dr. C. M. Charles of the Washington University School of Medicine in St. Louis, MO on June 27, 1940 (McWhirter (1977), p. 13). A series of remarkable photographs of Wadlow may be found in Polly Jae Lee, *Giant: The Pictorial History of the Human Colossus* (New York: A. S. Barnes & Company, 1970), pp. 129-133.

[17] Jeune and Vaupel (1995) gather several excellent discussions of claimed centenarians before 1800.

expectancy at birth in the United States (73.8 years), we obtain 115.3 years.[18] While numerous objections might be offered to this crude comparison of extreme stature with extreme longevity (e.g., the fact that some giantism is caused by disease), it is nevertheless notable that the comparison produces an estimate of extreme longevity in each sex quite comparable to what has so far been proven.[19]

---

[18] The average adult heights (ages 18-74) were obtained from NCHS Vital Health Statistics (Series 11 No. 238), "Anthropometric Reference Data and Prevalence of Overweight, United States 1976-80." Life expectancies at birth were obtained from Richard N. Anderson, Ph.D., CDC/NVSS *National Vital Statistics Reports* (Vol. 48 No. 18) (February 7, 2001), "United States Life Tables, 1998," Table A. Average adult heights in the United States have probably increased somewhat since the 1976-80 data was published.

[19] By way of contrast a comparison of the extremes of longevity with the extremes of human weight fails completely. While there may be genetic predisposition to extreme overweight, verified extreme weights exceeding 1,000 pounds for persons of normal stature easily exceed normal weight for the same stature by factors of 5:1 or greater. Fortunately for those who live many years, extreme longevity has never become the subject of exhibition to the extent of extreme stature and extreme weight. Thomas Parr was one of the few long-livers to be subjected to exhibition, with fatal result.

Outside the medical domain, most of the literature concening very tall or very heavy human beings relates to their exhibition.

Gould (1929) and Clair (1968) discuss very tall human beings. Discussions of very heavy human beings are more difficult to find; Gould and Pyle (1896) and Clare (1968) have essays on this subject. One of the best-authenticated cases of extreme weight was Robert Earl Hughes (born June 4, 1926, Monticello, MO; died July 10, 1958, Bremen, IN; buried July 13, 1958, Benville Cemetery, Benville, Buckhorn Township, Brown County, IL) of Fishhook, Fairmount Township, Pike County, IL (height: 6 feet 0.5 inches) who attained a maximum weight of 1,069 pounds and weighed 1,041 pounds at death. McWhirter (1977) has excellent discussions of record-holders among very tall, very heavy and very long-lived human beings. Recent editions of the Guinness record book contain far less discussion on these topics.

Carrying this argument one step further, if we peg the maximum verified human height at approximately 9 feet and the maximum verified human age at approximately 120 years, then a 133 1/3-year-old human being ought to be approximately as rare as a 10-foot human being. The tremendous increase in world population in the last 100 years ought to have produced such outliers if they have ever existed. The author remains skeptical that either 10-foot or 133 1/3-year-old human beings have been observed or will be observed within the near-term future, barring dramatic changes in the patterns of human growth and mortality rates.[20]

Using their Oldest Old Database, Thatcher, Kannisto and Vaupel (1998), estimated mortality rates in the 0.500-0.650 range at age 120. This is in good agreement with Kestenbaum (1992) who estimated a 0.520 mortality rate for ages 110 and over using his best data. Using 0.500 as an estimate for annual mortality at all ages 110 and over, one derives a 1 in 1,024 (or roughly 1 in 1,000) chance of surviving from 120 to 130. Using this method of estimation, a single 120-year-old would have a 1 in 1,000 chance of surviving to 130. Ten 120-year-olds would have a 1 in 100 chance of producing a survivor to 130; one hundred such 120-year-olds, a 1 in 10 chance of producing such a survivor.

[20] Refer to J. W. Vaupel, "The Average French Baby May Live 95 or 100 Years" and J. Vallin and G. Caselli, "Towards a New Horizon in Demographic Trends: The Combined Effects of 150 Years Life Expectancy and New Fertility Models" in Robine, Vaupel, Jeune and Allard (1997) for the viewpoint of demographers who believe that significant future improvements in mortality at ages 85 and over are possible. Stipp (1999) provides a lively discussion of advanced age mortality "bulls" and "bears" among demographers and gerontologists.

The problem is the extreme scarcity of verified 120-year-olds. Given their extreme scarcity in modern populations, it seems doubtful that there have ever lived as many as 100 persons who have actually attained age 120.[21] Even if there have been as many such persons as 100, they would still have only a 1 in 10 chance of producing a 130-year-old.[22]

---

[21] One of the earliest such claims is for the Hebrew patriarch Moses, who is said to have lived 120 years, even though Psalm 90 is traditionally described as a prayer of Moses. The years of life claimed for patriarchs like Moses and Abraham (175 years) pale, however, when compared with the lifespans claimed for antedeluvian patriarchs like Adam (930 years), Methusaleh (969 years), and Noah (950 years). Some nineteenth-century investigators of human longevity were offended by these assertions. The Jews, however, were not alone in claiming long lifespans *and* great heights for the progenitors of their nation.

Some Christian theologians thought that Noah and his family must have been giants, since only they survived the flood and the Bible records postdeluvian giants. On the other hand, some rabbinic traditions asserted that giants survived the flood on mountains. The French academician Henrion (1718) asserted that human lifespan and stature began to decline following Adam's and Eve's fall and that the decline was only arrested by the incarnation of Jesus Christ. Henrion set Adam's height at 123 feet 9 inches; Noah's, at 27 feet; Abraham's, at 20 feet; and Moses's, at 13 feet. The Hebrew Bible does not record Eve's lifespan, but Henrion set her height at 118 feet.

The long lifespans and great heights claimed for the Biblical patriarchs reflected their greatness as progenitors of the human race and were not intended to be taken literally in such a manner as to offend common sense. For ancient legends of giants (with some discussion of their longevity) refer to E. J. Wood, *Giants and Dwarfs* (London: Richard Bentley, 1868) and Walter Stephens, *Giants in Those Days: Folklore, Ancient History, and Nationalism* (Lincoln NE: University of Nebraska Press, 1989). Young (1899) provides an extensive discussion of the lifespans claimed for the antedeluvian Biblical patriarchs including an interesting citation (pp. 84-85) of E. J. Fripp's theory that the lifespans of Noah's ancestors were constructed artificially so that all of them, including Methusaleh, were deceased by the time of the flood.

[22] Livi-Bacci (1992) (pp. 32-33) estimated that 80 billion total human lives had been lived (or were still in existence) through the end of the twentieth century. Many demographers are skeptical that any true supercentenarians (ages 110 and over) emerged before the twentieth century. While some demographers are skeptical of any centenarian claims before 1800, the author believes that Sirach 18:7 offers some support to the thesis that centenarians were known even in ancient times. Smith (1993), an excellent interdisciplinary study of human longevity which presumes significant knowledge of medicine and genetics, provides a worthwhile discussion of the evolution of human longevity.

The abundance of exaggerated centenarian age claims have made it difficult for some scientific investigators of the subject to avoid the trap of hyper-skepticism. One of the pioneering figures in the scientific investigation of centenarian claims was William J. Thoms (1803-1885), editor of *Notes and Queries* and author of *Human Longevity: Its Facts and Its Fictions* (1873; revised edition, 1879). Many of Thoms's associates, including most notably Sir George Cornewall Lewis (1806-1863), were

skeptical that any human being had ever survived 100 years, only to be refuted by well-proven cases which emerged by the latter part of the nineteenth century. Gould (1929) (pp. 47-49) provides an illuminating discussion of Thoms and his associates. Jeune and Vaupel (1995) gather several excellent papers on historical observations of centenarians and supercentenarians.

However, there are good indications that there is no fixed limit to human life[23]; closing mortality tables using a final mortality rate equal to unity is a relic of the classical extinct cohort method of construction which probably does not provide a good model for future mortality of the oldest old even when the limiting age is set to a value significantly higher than any proven human lifespan.[24]

The search for lifespans clustered around the proven spans of Calment and other verified supercentenarians will continue. The author believes that decedent research can contribute to this task. In the United States the DMF-PR provides an accessible starting point for the investigation of decedent supercentenarian age claims.

Potential Use of DMF-PR for Demographic Work

---

[23] Refer to Brown (1990) and John R. Wilmoth, "In Search of Limits," in Wachter and Finch (1997). For an opposite perspective, note Job 14:5: "Seeing his days are determined, the numbers of his months are with thee, thou hast appointed his bounds that he cannot pass" [King James Bible]. It is interesting to note in this regard that the 120-year maximum lifespan cited in Genesis 6:3 has been exceeded only by small margins in modern times.

[24] One alternative proposal for closing mortality tables intended for use in financial applications may be found in Faig (1995). A more elegant formulation could undoubtedly be obtained by using the logistic model to close mortality tables. If the force of mortality has a finite limit as age approaches infinity, then the annual mortality rate will have a finite limit < 1 as age approaches infinity. On the other hand, if the force of mortality becomes infinite as age approaches infinity, then the limit of the annual mortality rate as age approaches infinity will be 1. Neither of these scenarios involves a finite age (omega) above which survival is impossible; a finite omega can only occur if the force of mortality becomes infinite at the immediately preceding age.

The DMF-PR cannot be used directly to construct reliable mortality tables. If it is required that each BY cohort be observed through Y = 125 in order to utilize the extinct cohort methodology of constructing mortality tables, there is a 125-year lag from birth before a complete cohort mortality table can be constructed. One may justifiably ask what interest other than an historical one may reside in the compilation of a mortality table for a birth cohort which is completely extinct. Table 12 illustrates the construction of such an extinct cohort mortality table from the DMF-PR deaths based on BY = 1865-1874 cohorts, which reached Y = 125 during the period 1990-1999 and were therefore presumably completely extinct by 2000. Some indication of the unreliability of the data at the oldest ages may be observed in the heaping of deaths at ages 115 and 120. Quinquennial age heaping is a classic measure of unreliability of mortality statistics. Graduated mortality rates for attained ages 100 and higher for the BY = 1865-1874 cohorts obtained from the reported DMF-PR deaths would probably not differ much from 300-350 per thousand.

Next to the lag necessary for extinct cohort basis mortality studies, the reliability of the underlying data is the next largest concern in attempting to construct mortality tables from deaths-only records. The BY = 1865-1874 cohorts comprise only 222,751 of the 65,089,043 deaths in the December 2000 DMF-PR.

While within the last twenty years (1979-1999) approximately ninety percent of total deaths of U.S. residents (as measured against NDI) have been reported in DMF-PR, completeness is much lower for earlier DY cohorts, especially before the keeping of automated death records by SSA beginning in 1962. Since the completeness of death reporting in DMF-PR has improved over time, the natural tendency for each BY cohort is for deaths at the highest ages to be more complete than deaths at the lower ages. This bias affects any extinct BY cohort mortality study constructed from DMF-PR.

The data in the DMF-PR may also be compared with the decennial enumerations of the population of the United States as compiled by the Census Bureau. In July 1999 the Census Bureau issued its report "Centenarians in the United States" [Krach and Velkoff (1999)] based largely upon the data compiled for the 1990 census. Table 13 compares the enumerated centenarians for decennial census years 1950 through 1990 (and the Census Bureau's best estimates of the actual centenarian populations in those years) with the total centenarian deaths for the same years reported in the DMF-PR.

The reliability of both the DMF-PR numerator and the population census denominator used in the computation of the crude centenarian death rate in Table 13 is open to considerable question. The extremely sparse death reporting in the December 2000 DMF-PR for DY = 1950 (only 18,406 deaths reported at all ages) and for DY = 1960 (only 90,310 deaths reported at all ages) and the known incompleteness of the DMF-PR for subsequent years (93.06% of NDI for DY = 1980 and 86.41% of NDI for DY = 1990) makes the use of DMF-PR reported centenarian deaths in the numerator untenable for 1950-1960 and questionable for 1970-1990.

The very large increase in the reported centenarian population in the 1970 census has been attributed to misunderstanding of some respondents in completing age information. The 1970 U.S. census was the first primarily self-reported census. The census schedules for prior decennial censuses were completed by employees of the U.S. government. Krach and Velkoff (1999) did divide the centenarians enumerated in the 1990 census into two broad categories: ages 100-104 (30,947 raw count) and ages 105 and over (6,359 raw count). The crude death rates derived from the 1990 DMF-PR deaths (12,095 at ages 100-104 and 1,160 at ages 105 and over) for these two categories are 391 and 182 per thousand, respectively.

Females outnumber males very significantly in all observed centenarian populations. In the case of the 1990 U.S. census, the 37,306 persons enumerated at ages 100 and older included 29,405 females and 7,901 males. The DMF-PR does not identify individuals by sex; however, with the use of a good dictionary of given names, it should be possible to make reasonably reliable sex assignments for all persons except those identified only by surname and initial rather than by surname and given name. Sampling studies performed in cooperation with NCHS or state bureaus of vital statistics or both will usually have access to sex information through the subject's death registration.

It is important, however, not to underestimate the reliability of the decedent data in the DMF-PR. Elo, Preston, Rosenwaike, Hill and Cheney (1995) of the Population Aging Research Center of the University of Pennsylvania compared age as stated in state death records, SSA death records, and the U.S. censuses of 1900, 1910 and 1920 for a sample of elderly Afro-Americans. They found that the age match with early-life census records (an excellent proxy for missing birth registration records) was better for SSA death records than for state death records.

Kestenbaum (1997) analyzed the completeness of SSA death records utilizing the 1993 National Mortality Followback Study. Of 2,184,387 decedents at ages 15 and over in 1993, 2,174,163 (99.5%) were matched to SSA records and 2,143,400 (98.1%) were recorded as deceased in SSA records. For 2,063,220 (94.5%) of the decedents, age in the SSA records agreed with age in the state records. The completeness of reporting for 1993 decedents aged 100 and over was even better: of 15,055 such decedents, 15,013 (99.7%) were matched to SSA records and 14,826 (98.5%) were recorded as deceased in SSA records. Age agreement, however, was poorer for decedents aged 100 and over: of the 15,055 such decedents, only 12,661 (84.1%) had the same age in SSA records and state records.

In a prior paper, Kestenbaum (1992) demonstrated the superior reliability of SSA death records even as compared to the HCFA Medicare deaths utilized to complete the decennial U.S. life tables from 1969-71 onward.[25] By restricting his analysis of 1987 deaths to insured status Medicare

[25] For the methodology used for the 1969-1971 U.S. Decennial Life Tables, refer to *Methodology of the National and State Life Tables for the United States: 1969-1971* in the series U.S. Decennial Life Tables For 1969-71 (vol. 1 no. 3) (DHEW Publication No. (HRA) 75-1150) (Rockville MD: National Center for Health Statistics, May 1975). Table 5 of this publication displays

Part B enrollees, excluding Railroad Retirement Board data, and combining death information from MBR and NUMIDENT, Kestenbaum (1992) (Table 1, column 8) was able to construct a very credible table of raw death rates covering individual attained ages 85 through 109 and closing with grouped ages 110 and over. While the DMF-PR (which excludes deaths deriving solely from state bureaus of vital statistics) is undoubtedly less reliable than Kestenbaum's carefully selected data, there is nevertheless strong evidence that SSA death data is generally very accurate especially when additional requirements such as insured status and Medicare Part B enrollment are imposed.

State death certificate information is only as reliable as the information provided by the informant; there is generally no cross-checking with birth registration records before state death records are filed. On the other hand those SSA death records which relate to insureds will generally have been subjected to an age verification process. Aziz and Buckler (1992) provide a brief history of the development of these age verification processes. If a death verification flag were to be added to the DMF as proposed in SSA-OIG (2000), it would doubtless help to identify the DMF records which may be expected to exhibit the highest data quality. However, it is probable that such a death verification flag will be treated by SSA as confidential data and not included in DMF-PR even on a prospective basis.

The greatest value of the DMF-PR for demographic studies is its unrestricted public availability. The author believes that carefully constructed sampling from the DMF-PR holds the best prospect for valuable demographic work.

---

the actual graduated Medicare death rates at ages 85-109 which were utilized for ages 95 and over in the 1969-71 U.S. Decennial Life Tables. Beginning with 1996, the highest age classification in the Complete Annual U.S. Life Tables was changed from 85 and over to 100 and over. The details concerning the construction of the Complete Annual U.S. Life Tables at the highest ages may be found in Anderson (1999). Spiegelman (1968) (p. 128) describes the methodologies used to close the U.S. Decennial Life Tables for 1939-41, 1949-51 and 1959-61. The 1949-51 and 1959-61 tables were closed with Union Civil War veteran mortality described by Myers and Shudde (1955).

Hetzel (1997) provides a history of the gathering of vital statistics in the United States. While the birth registration area in the United States was gradually completed between 1915 and 1933, birth reporting in some states remained incomplete for years afterward. A number of eastern states and municipalities, however, do have substantial earlier birth records. In addition, the U.S. censuses of 1900, 1910 and 1920[26] provide an excellent age-verification tool for persons living on their enumeration dates[27], which will include many of the persons who live to celebrate their one hundredth birthday during the next two decades.

---

[26] Szucs and Luebking (1988) provide excellent coverage of federal resources for age-identity validations. Eakle and Czerny (1984) is an excellent survey of all age-identity validation resources in the United States, including federal records.

[27] The enumeration date for the 1790-1820 U.S. censuses was the first Monday in August; for 1830-1900, June 1; for 1910, April 15; for 1920, January 1; from 1930 onward, April 1. The enumeration date must be distinguished from the date the enumerator actually conducted his or her survey to complete the census schedules. The enumerators were instructed to include all persons living in a household on the specified enumeration date; thus, persons who died subsequent to the enumeration date were to be included while persons born subsequent to the enumeration date were to be excluded. The instructions for the self-reported schedules utilized in the 1970 and later censuses have included the instruction to reflect household status as of the April 1 census enumeration date. Possible misinterpretation of the instructions regarding the enumeration date both by government enumerators (1960 and earlier censuses) and by self-enumerators (1970 and later censuses) should be anticipated when consulting census records. The accurate enumeration of transient populations as of a fixed enumeration date remains a challenging problem for the Census Bureau.

Both the 1900 and 1920 census have a full surname soundex index organized by state and by surname soundex code and given name. The 1900 census is almost ideal for age and identity verification of persons living in the United States on June 1, 1900 since it includes for each person enumerated the following critical age-identity data: name, sex, age, race, relationship to head of household, marital status, years married, children borne, children living, birthplace, month and year of birth, birthplace of mother and father, occupation, citizenship and year of immigration if foreign-born.

While the soundex index of the 1910 census was only compiled for a minority of states, it is nevertheless true that the censuses of 1900, 1910 and 1920 will provide an invaluable tool for age validation of centenarians during the next several decades.[28] Regrettably, the 1890 census schedules, with the exception of a few veterans' records, had to be destroyed after suffering severe damage in a fire at the Commerce Department in January 1921.[29] For the most extreme age claims, the 1880 census does have a soundex surname index which extends to all households containing children aged ten or less. One limiting factor for the usefulness of the census for early age-identity verification of centenarians is that 17% of the centenarian population recorded in the 1990 census was foreign-born [Krach and Velkoff (1999)]. For foreign-born centenarians, such as Christian Mortensen, other records must be utilized for early age-identity verification.

The downside of census-based age-identity validation is that the census schedules and soundex indexes do not exist in electronic form, but must be accessed using microfilm available in all the U.S. National Archives Field Branches and at major genealogical libraries. (Smaller genealogy collections will often own the census microfilm for their own and adjoining states.) A single-state soundex search for one census can usually be accomplished in fifteen minutes or less particularly if the soundex surname

---

[28] By law, the census schedules are released for public use seventy-two years after their compilation, in plenty of time for use for centenarian age-identity validations. The 1930 census (scheduled for release to the public in 2002) will be more difficult to use for age-identity validations than the 1900-1920 censuses since only ten complete states and two partial states have soundex indexes. In the 1940 census only eight counties in Mississippi are soundex-indexed; there is no name index of any kind for the 1950 and later censuses. In addition, no individual name-identified data from the 1930 and later census schedules has been recorded electronically. With the limited exceptions noted for 1930 and 1940, exact street addresses will therefore be almost mandatory to insure success for searches for individuals located in major municipalities in 1930 and later censuses. The completion of the birth registration area by 1933 should provide an alternative for census age-identity validations when state birth records are accessible to the researcher. NCHS receives only birth records stripped of identifying information and maintains no birth index equivalent to its NDI. The author is indebted to JoAnn Shepherd of the History Staff of the U.S. Census Bureau for information regarding the 1930 and later censuses.

[29] Ancestry, Inc. has created for its subscribers a searchable online 1890 "Census Substitute" which consists in large part of city directories clustering around 1890.

code is not a large one. However, a nationwide search for one census can easily consume twelve hours (48 x 0.25 hours/search), which could represent three days of research in four-hour sessions at a National Archives branch.

Clearly, state localization of birth or contemporary residence is an important element for most census searches for sampled populations except perhaps when the most extreme age claims are involved.[30] For persons residing in major municipalities, city directories and telephone directories can offer additional assistance in tracing the life-career of an individual. Older city directories will sometimes even provide the new residence city of a formerly listed individual in the first directory after his or her removal from the locality, thus providing a thread to continue the life-career development.[31] Newspaper obituaries (many posted on the Internet) provide another resource for localizing and validating centenarian age at death claims found in DMF-PR. For example, the March 2001 Associated Press reports of the death of John Painter (1888-2001), the oldest surviving United States veteran of World War I, reported that he spent his early life in Jackson County, TN.

For historically-focused studies, a number of private firms have compiled indexes of the 1790-1850 censuses, with more limited coverage of the 1860 and 1870 censuses. Many of these indexes are electronic and can be searched for a fee.[32] However, the 1790-1840 census schedules identify only the

---

[30] For example, a researcher wishing to study MA deaths might restrict selection from DMF-PR to records with: (1) SSN prefix indicating issuance in MA; (2) ZIPcodes of last residence in MA; and (3) ZIPcodes of last payment in MA.

[31] Eakle and Czerny (1984) provide an essay by Gordon Lewis Remington, "City Directories and Their Cousins," on this subject.

[32] Accelerated Indexing Systems (AIS) founded by the late Roger Vern Jackson (1946-1999) was a pioneer in indexing the pre-1880 censuses. Paper or microfiche editions of the AIS and other privately compiled census indexes can also be found in many major genealogical libraries; some of the National Archives Field Branches have donated collections of these privately compiled indexes.

head of household by name. In addition they display only very broad age categories for household members. The 1850 census was the first to show the names and ages of all household members.

The labor involved in census validation of centenarian deaths reported in DMF-PR may be reduced as digital scanning of population schedule microfilm by private enterprises accelerates. For example, Ancestry, Inc. makes the 1790-1870 census indexes created by Accelerated Indexing Systems (AIS) available online and plans to create online head of household indexes for later censuses not indexed in its AIS census index database. Ancestry, Inc.'s digital population schedule images (scheduled for completion for the full U.S. census 1790-1920 in 2001) are browseable, not searchable, so users must enter utilizing a reference found in the searchable index database or a geographical locality.[33] Whether the federal government will elect to adopt new technology for future census releases, remains to be seen. Advances in digital scanning may eventually make possible the automated creation of reliable name indexes directly from the census schedules themselves. This will be particularly important for the 1930 and later censuses since the soundex surname indexes compiled by the Works Progress Administration in the 1930s are primarily limited to the 1880, 1900, 1910 and 1920 censuses.[34]

The primarily state-based vital record system in the United States is different from the national-based systems which have long been in place in many western European countries. Demographers seeking to perform age-identity validations for centenarians utilizing vital records in the United States face a more complex assignment than their peers utilizing national record-keeping systems in western Europe. However, the demographer well-armed with the genealogical literature bearing upon age-identity research in U.S. vital records is likely to be rewarded with a significant degree of success in verifying (or disproving) age for individuals or sample populations drawn from DMF-PR.

The DMF-PR has become an indispensable tool for twentieth-century genealogical and probate work and has generated its own share of literature bearing upon those disciplines.[35] The DMF-PR can

---

[33] Subscription prices for the Ancestry, Inc. census indexes and images may be found on the company's Internet site at http://www.ancestry.com.

[34] Refer to footnote 30 for discussion of the very limited soundex indexes for the 1930 and 1940 censuses.

[35] Jake Gehring's article "The Social Security Death Master File: A Much Misunderstood Index" may be found on the Ancestry.com website on the Internet at: http://www.ancestry.com/search/rectype/vital/ssdi/article.htm. Ancestry.com also provides a document entitled "Social Security Death Index - FAQs [Frequently Asked Questions]" on the Internet at http://ancestry.com/search/rectype/vital/ssdi/faq.htm. Rootsweb maintains a helpful tutorial "A Unique Finding Aid: Social Security Death Index" on the Internet at: http:/www.rootsweb.com/~rwguide/lesson10.htm.

provide a low-cost entry vehicle for research to be followed up in vital records maintained by NCHS or state bureaus of vital statistics or in the population schedules of the U.S. census.

The Importance of Centenarian Research

Centenarians are no longer a curiosity in our society. The "centenarian cult" as described by Peter Laslett[36] is gradually giving way to a more objective view of the oldest members of our society.

[36] P. Laslett, "The Bewildering History of the History of Longevity" in Jeune and Vaupel (1999).

The Census Bureau report "Centenarians in the United States" [Krach and Velkoff (1999)] tells why the subject should be of interest to physicians, scientists and public policy makers. Table 14 (taken from Table 2 of this report) shows the projected number of centenarians in decennial years 2000 through 2050 under low, middle and high series projection assumptions.[37]

A somewhat different middle series population projection posted by the Census Bureau on the Internet[38] enables one to compute projected prevalence rates for centenarians in the general population. This data for decennial years 2000-2050 is shown in Table 15.

These projected centenarian populations speak to the needs which the increasing prevalence of centenarians is likely to create in the United States and other developed nations as we progress into the twenty-first century. According to Krach and Velkoff (1999), nearly fifty percent of U.S. centenarians in 1990 were living in nursing homes. Fewer than twenty percent of U.S. centenarians in 1990 had no self-care or mobility limitations. Excluding those living in institutions, 25.9% of centenarian females and 21.1% of centenarian males were living in poverty in the United States in 1990. Fewer than ten percent had the support of a living spouse; about fifteen percent were living alone.

---

[37] The actual enumeration of the population by age in the 2000 census is currently scheduled for release during the summer of 2001 (http://www.census.gov/population/www/cen2000/briefs/html).

[38] These projection data sets may be found at http://www.census.gov/population/projections/nation/summary/np-t3-a.txt through np-t3-g.txt. The methodology used for these projections may be found in Frederick W. Hollmann, Tammany J. Mulder, and Jeffrey E. Kallan, "Methodology and Assumptions for the Population Projections of the United States: 1999 to 2100" (U.S. Census Bureau, Population Division Working Paper No. 38), dated January 2000, which may be found at http://www.census.gov/population/www/documentation/twps0038.html.

Many causes affect the quality of data which can be gathered for centenarians. Bowerman (1939) noted the strong correlation of centenarian prevalence with illiteracy rates by state in the United States. Cognitive disability limits the ability of some centenarians to report their ages accurately. Several writers[39] have commented on the cultural factors impacting an individual's knowledge of his or her true age. There is undoubtedly some tendency of extremely aged persons (and their relatives and friends) to overstate their ages out of a sense of pride in long-living.[40] In the 1990 census[41], 83% of centenarians reported ages 100-104, 9% ages 105-109, and 7% ages 110 and over. Of 1505 whites who reported ages 110 and over in the 1990 census, 68% reported no mobility or personal care limitations; 30% lived alone; and 41% were married--all of which reported characteristics cast severe doubt of the validity of most of the supercentenarian age reports in the 1990 U.S. census.

[39] Clark (1909), p. 154; P. Laslett, *op. cit.*, p. 31; and Z. Weng, Y. Zeng, B. Jeune and J. W. Vaupel, "The Validation of Han Chinese Centenarians" in Jeune and Vaupel (1999), p. 209.

[40] Alter (1990) provides a statistical analysis of age misreporting in the U.S. censuses, 1900-1940. Myers (1966) analyzes the validity of centenarian data reported in the 1960 census. Krach and Velkoff (1999) provide a number of citations for studies of age misstatement in the 1970, 1980 and 1990 censuses.

[41] Refer to Krach and Velkoff (1999).

Some actuaries[42] have noted that the mortality improvements which have led to the increasing centenarian population have not been observed consistently at ages 85 and over. While Perls[43] reports a surprising heterogeneity in health status among the centenarians studied in the New England Centenarian Project, the observed death rate for the program's verified centenarian subjects has run about 50% per annum. While noting occasional cases of exceptional vigor in extreme old age, Clark (1909)[44] noted a marked increase in frailty between the ninth and tenth decades of life. Forette[45] cites survey studies of Finnish and French centenarians by Louhija and Allard which found high prevalences of severe physical and cognitive impairments among the populations studied.

The possibility of an emerging explosion in a predominantly frail centenarian population is a concern for all of those involved with related economic, health care and public policy issues. We must determine whether we can afford to provide the nursing home beds and the medical care which our

---

[42] Goss, Wade and Bell (1998) (p. 114) found that age-adjusted annual death rates for ages 85 and over in the United States actually worsened by 0.72% per year for males and by 0.52% for females during the observation period 1990-94.

[43] T. T. Perls, K. Bochen, M. Freemen, L. Alpert and M. H. Silver, "Age Validation in the New England Centenarian Study" in Jeune and Vaupel (1999). The website for the New England Centenarian Study [http://www.med.harvard.edu/programs/necs/] provides a vivid portrait of the heterogeneity of the centenarian population and many useful links. Perls, Silver and Laverman (1999) provide many diverse portraits of centenarian lifestyles.

Clark (1909) (p. 157) and Gould (1929) (p. 65) both comment on exceptionally vigorous centenarians. (The author has been unable to identify Clark's centenarian preacher Rev. Dr. Howe [certainly not Episcopal bishop Mark Antony DeWolfe Howe (1809-1895)] but Gould's Admiral Sir Provo William Parry Wallis (1791-1892) may be found in a number of standard reference works.)

The best-known centenarian at this writing is undoubtedly the English Queen Mother (Lady Bowes-Lyon, the widow of King George VI), who celebrated her one hundredth birthday on August 4, 2000. The Queen Mother is still alert and ambulatory in her 101st year of life. She is the first of European royalty to celebrate a one hundredth birthday and was preceded as a centenarian by only two English peers.

[44] Franklin C. Clark, M.D. (1847-1915) spent his lifetime in medical practice in Providence, RI. He had extensive personal and professional exposure to nonagenarians. Over various periods of his life, he lived with his maternal grandmother (1786-1878), his paternal grandmother (1789-1883) and his mother (1814-1906). By way of contrast only his maternal grandfather (ca. 1783-1863) was long-lived; his father (1816-1849) and his paternal grandfather (1790-1820) died prematurely even for their times. A biographical sketch of Clark and a bibliography of his writings may be found in his posthumously published novel *Susan's Obituary* (Glenview IL: Moshassuck Press, 1996).

[45] B. Forette, "Centenarians: Health and Frailty" in Robine, Vaupel, Jeune and Allard (1997).

future centenarian population is likely to require. That a significant portion of the future centenarian population of the United States will be unable to provide for such care using their own resources seems probable based on the demographics of U.S. centenarians in 1990.

We will surely need a broad scientific understanding of the phenomenon of centenarianism as we progress into the twenty-first century. New books addressed to the general public from experts in the field like Perls, Silver and Laverman (1999) and Olshansky and Carnes (2001) will help raise public consciousness of issues concerning the oldest old and one may anticipate more articles such as Stipp (1999) that address the quandaries which arise concerning future centenarian populations and their health status. We will need the best possible measurements of centenarian mortality and morbidity so that we can project populations and their probable needs. However, our work must also be informed by the broad perspective of the humanities.

The December 2000 DMF-PR is an important record of over 65 million completed human lives with the potential (in combination with state records and U.S. census population schedules) for facilitating strides in our understanding of the phenomenon of centenarianism in the United States. Demographic studies based on scientific sampling of the DMF-PR can help to increase knowledge relating to the oldest old, including the possible identification and validation of supercentenarians with lifespans clustering around those already proven for Jeanne Calment and Christian Mortensen.

The completion of the birth registration area in the United States will not celebrate its centennial until 2033. Even so, birth records remain among the most restricted of all state records. (By way of contrast, death records are public records in a significant number of states.) Thus, the methodologies available for the validation of centenarians in the United States described in this paper are likely to remain useful in the future.

Researchers from many fields depend upon unrestricted public access to the DMF-PR. The usefulness of the DMF-PR for demographic work will grow as the accuracy, timeliness and completeness of its death reporting continues to improve. The addition to the DMF-PR of deaths reported to SSA solely by state bureaus of vital statistics would be a welcome future enhancement as the NAPHSIS EDRS project progresses.[46]

Acknowledgements

The author thanks for their assistance in providing source material: Professor George Alter of Indiana University; Professor Robert L. Brown of the University of Waterloo; Ellen Bull of the Society of Actuaries Library; Bert Kestenbaum of the Office of the Actuary of the Social Security

[46] The information reported in the DMF-PR does not contain any information relating to cause of death.

Administration; Eric J. Margolis, M.D.; Professor Alan Shepard of Texas Christian University; and JoAnn Shepherd of the History Staff of the U.S. Census Bureau. The author also thanks Carol Faig for assistance in validating data and PolySystems, Inc. for providing research facilities.

The author is solely responsible for all of the assertions of fact and of opinion in this paper. The author owes his interest in centenarians to the pioneering work of Lieutenant Commander Rupert Thomas Gould, R.N.[47], in whose memory this paper is dedicated.

Appendix: Gould's Remarkable Outlier Supercentenarian Thomas Parr

Gould (1929) was impressed by the detail provided in John Taylor's contemporary (1635) verse biography of Thomas Parr; but his principal arguments in support of Parr's claims were two. First of all, Parr did not claim detailed memory of events witnessed in childhood, a claim made by many false supercentenarians. Secondly, none of his neighbors protested against deception when he was brought (with fatal results) to London to be introduced to King Charles I and his court and exhibited to the public. Weighing heavily against Parr's claims, however, are the results of his autopsy as performed by Sir William Harvey and his claim to have married for the first time at age 80 and to have performed penance at age 100 for begetting a bastard child by one Catherine Milton. The latter was his response when King Charles demanded of him: "You have lived longer than other men--what have you done more than other men?"

Regrettably, there appears to remain no primary evidence to support the age claim (152 years) made for Thomas Parr; and the credibility of his claim has not been aided by nineteenth-century claims of supercentenarian progeny. A note contributed to *The Topographer and Genealogist* by "H. P."

[47] Gould (born November 16, 1890; died October 5, 1948) is probably best known for his restoration of John Harrison's eighteenth-century marine chronometers and was recently portrayed by Jeremy Irons in a BBC presentation on this subject. He was also well known for his lively interchanges with Professor Cyril E. M. Joad on the popular wartime radio program "The Brains Trust" (1941-47). A sketch of Gould's life with a bibliography of his writings by Leslie Shepard may be found in *Oddities: A Book of Unexplained Facts* (New Hyde Park NY: University Books, 1964). Another sketch of Gould's life by Jonathan Betts may be found at http://www.rog.nmm.ac.uk/museum/harrison/gould.html. Betts is preparing a biography of Gould for publication.

and dated at Taunton, May 29, 1854, claimed that "Young Parr," his bastard child by Catherine Milton, lived to 113; a grandson, to 109, and a great-grandson Robert Parr (died 1757), to age 124. Gould (1929) cites the *Annual Register* for 1792 for a report of the death of an alleged great-grandson Robert Parr in July 1761 at age 127 and of an alleged great-granddaughter Catherine Parr in October 1792 at age 103. According to his "official" biography, both of Parr's children by his first wife Jane Taylor died in infancy and he had no children by his second wife Jane Lloyd (or Flood).

Parr's age claim was debunked by Thoms (1873) (pp. 85-94) and Young (1899). Despite the exaggerated claims of his contemporary biography, the author joins Gould in believing that Thomas Parr was probably at least a near-centenarian. A late first marriage would ordinarily have occurred 50 years prior to Parr's claimed first marriage at age 80; so the best round estimate of his actual age at death may be approximately 102 years rather than 152 years.

Little has been published about Parr and his claims since Gould (1929); note, however, the worthwhile discussion in Clair (1968) (pp. 163-165) and Alan Shepard's article " 'O seditious Citizen of the Physicall Common-Wealth!': Harvey's Royalism and His Autopsy of Old Parr" [*University of Toronto Quarterly* (vol. 65 no. 3) (1996), pp. 482-505], which interprets Harvey's autopsy of Parr in a political context. A modern discussion of Thomas Parr, including a facsimile reproduction of John Taylor's 1635 biography, with its famous portrait of the "old, old, very old man" (reproduced in Jeune and Vaupel (1999), p. 48 and in Gould (1929), frontispiece) is certainly long overdue.[48]

Gould (1929) reported that "Old Parr's Life Pills," invented by a Manchester druggist in the mid-nineteenth century, supposedly from a secret recipe obtained from one of Parr's descendants, still enjoyed some popularity among the poorer classes in the first decades of the twentieth century. Parr is unique among famous centenarians in being memorialized by a brand of premium scotch whiskey. One may hope some supercentenarians of the future will have the privilege of celebrating their birthdays with a glass of "Old Parr." Through the work of Perls and other researchers, the public's perception of

[48] Thoms (1873) reprinted John Taylor's verse biography of Parr (pp. 291-308) and Harvey's autopsy report on Parr (pp. 308-312). Taylor, a seventeenth-century poet and pamphleteer, was interested in the extremities not only of the human lifespan but also of human girth. He was also the author of *The Great Eater of Kent, or part of the admirable Teeth and Stomacks Exploits of Nicholas Wood, of Harrisom in the County of Kent* (London: Henry Gosson, 1630). Clair (1968) describes this work.

centenarians is gradually being freed from the popular press stereotypes which were described so well by Gould (1929) (p. 46):

"Centenarians, so far as one can judge by what one reads about them, are chiefly occupied in boring their attendants, or reading without glasses and intelligence, and otherwise demonstrating that their faculties are no worse than they always were. If they put anything on record except a malicious codicil or so, it is usually a testimonial to the virtues of their pet nostrum, or an affidavit certifying that throughout their lives they have always shunned (or, alternatively, eagerly consumed) tobacco and/or alcohol. Very few of them have left us any means of judging what they thought of old age, as a condition--probably they lacked the energy to do so."

References

Alter (1990). George Alter, "Old Age Mortality and Age Misreporting in the United States, 1900-1940," Working Paper #24, Population Institute for Research and Training, Indiana University.

Anderson (1999). R. N. Anderson, *Method for Constructing Complete Annual U.S. Life Tables*, National Center for Health Statistics (NCHS), *Vital and Health Statistics*, Series 2, no. 129.

Aziz and Buckler (1992). Faye Aziz and Warren Buckler, "The Status of Death Information in Social Security Administration Files," Social Statistics Section *Proceedings*, American Statistical Association, pp. 262-267.

Bowerman (1939). W. G. Bowerman, "Centenarians," *Transactions of the Actuarial Society of America*, vol. XL, pp. 360-378.

Brown (1990). Robert L. Brown, "Rectangularization of the Survivorship Curve - Facts Versus Appearances," University of Waterloo, Institute of Insurance and Pension Research, Research Report 90-04.

Clair (1968). Colin Clair, *Human Curiosities* (New York: Abelard-Schuman).

Clark (1909). Franklin C. Clark, M.D., "The Problem of Centenarianism," *The Providence Medical Journal*, vol. 10, pp. 143-158.

Eakle and Czerny (1984). Arlene Eakle and Johni Czerny (eds.), *The Source: A Guidebook of American Genealogy* (Salt Lake City UT: Ancestry Publishing Company). A revised edition of this compilaton (under the editorship of Loretto Dennis Szucs and Sandra Hargreaves Luebking) was published by Ancestry Publishing in 1997 and includes some coverage of DMF-PR.

Elo, Preston, Rosenwaike, Hill and Cheney (1995). Irma T. Elo, Samuel H. Preston, Ira Rosenwaike,

Mark Hill and Timothy P. Cheney, "Consistency of Age Reporting on Death Certificates and Social Security Administration Records among Elderly African-American Decedents," Working Paper Series No. 95-03, Population Aging Research Center, University of Pennsylvania.

Faig (1995). Kenneth W. Faig, Jr., "Life After Omega," *The Financial Reporter* (newsletter of the Financial Reporting Section of the Society of Actuaries), March 1995, pp. 7-8, 15.

Goss, Wade and Bell (1998). Stephen C. Goss, Alice Wade and Felicitie Bell, "Historical and Projected Mortality for Mexico, Canada, and the United States," *North American Actuarial Journal* (vol. 2 no. 4), October 1998, pp. 108-126.

Gould (1929). Rupert T. Gould, "Old Parr, and Others," *Engimas: Another Book of Unexplained Facts* (London: Philip Allan). Subsequent editions: (1) London: Geoffrey Bles, 1946 and (2) New Hyde Park NY: University Books, 1965. [Page citations in this article use the 1965 edition.]

Gould and Pyle (1896). George M. Gould, M.D. and Walter L. Pyle, M.D., *Anomalies and Curiosities of Medicine* (W. B. Saunders, 1896; reprinted, New York: The Julian Press, 1956).

Hetzel (1997), A. M. Hetzel, "History and Organization of the Vital Statistics System," in *U.S. Vital Statistics System: Major Activities and Developments 1950-1995*, National Center for Health Statistics, DHHS Publication no. (PHS) 97-1003, pp. 43-66. Reprinted from *Vital Statistics of the United States*, vol. 1, 1950, pp. 1-19. Electronic edition: http://www.cdc.gov/nchs/data/usvss.pdf.

Jeune and Vaupel (1995). Bernard Jeune and James W. Vaupel, eds., *Exceptional Longevity: From Prehistory to the Present*, Odense University Press, Monographs on Population Aging, 2. Electronic edition: http://www.demogr.mpg.de/Papers/Books/Monograph2.

Jeune and Vaupel (1999). Bernard Jeune and James W. Vaupel, eds., *Validation of Exceptional Longevity*, Odense University Press, Monographs on Population Aging, 6.

Kannisto (1994). Vaino Kannisto, *Development of the Oldest-Old Mortality: 1950-1990: Evidence From 28 Developed Countries*, Odense University Press, Monographs on Population Aging, 1. Electronic edition: http://www.demogr.mpg.de/Papers/Books/Monograph1.

Kannisto (1996). Vaino Kannisto, *The Advancing Frontier of Survival: Life Tables for Old Age*, Odense University Press, Monographs on Population Aging, 3. Electronic edition: http://www.demogr.mpg.de/Papers/Books/Monograph3.

Kestenbaum (1992). Bert Kestenbaum, "A Description of the Extreme Aged Population Based on Improved Medicare Enrollment Data," *Demography* (vol. 29 no. 4), November 1992, pp. 565-580.

Kestenbaum (1997). Bert Kestenbaum, "Recent Mortality of the Oldest Old, from Medicare Data," paper presented at the 1997 meetings of the Population Association of America, March 27-29, 1997.

Krach and Velkoff (1999). Constance A. Krach and Victoria A. Velkoff, *Centenarians in the United States: 1990*, U.S. Census Bureau, Current Population Reports: Special Studies, publication no. P23-199RV. Electronic edition: http://www.census.gov/prod/99pubs/p23-199.pdf.

Livi-Bacci (1992). Massimo Livi-Bacci, *A Concise History of World Population* (Cambridge MA and Oxford UK: Blackwell Publishers). Originally published in Italian in 1989; translated by Carl Ipsen.

McWhirter (1977). Norris and Ross McWhirter, *Guiness Book of World Records: 1977 Edition* (New York NY: Sterling Publishing Company). [Older editions of this popular record book contain much more extended discussions of human longevity, height and weight records than more recent editions.]

Myers (1966). Robert J. Myers, "Validity of Centenarian Data in 1960 Census," *Demography* (vol. 3), pp. 470 *et seq.*

Myers (1975). Robert J. Myers, *Social Security*, Homewood IL: Richard D. Irwin.

Myers and Shudde (1955). Robert J. Myers and Louis O. Shudde, "Mortality Experience of Union Civil War Veterans," *Society of Actuaries Transactions* (vol. 7 no. 3), pp. 63-68.

Olshansky and Carnes (2001). S. Jay Olshansky and Bruce A. Carnes, *The Quest for Immortality: Science at the Frontiers of Aging* (New York NY: W. W. Norton & Company).

Perls, Silver and Laverman (1999). Thomas T. Perls, Margery Hutter Silver and John F. Laverman, *Living to 100: Lessons in Living to Your Maximum Potential at Any Age* (New York NY: Basic Books).

Robine, Vaupel, Jeune and Allard (1997). J.-M. Robine, J. W. Vaupel, B. Jeune and M. Allard (eds.), *Longevity: To the Limits and Beyond* (Berlin, Heidelberg and New York: Springer Verlag).

Smith (1993). Smith, David W. E., M.D., *Human Longevity* (Oxford UK: Oxford University Press).

Spiegelman (1968). Mortimer Spiegelman, *Introduction to Demography* (Cambridge MA: Harvard University Press).

SSA-OIG (2000). Social Security Administration, Office of the Inspector General, *Improving the Usefulness of Social Security Administration's Death Master File*, Evaluation Report no. A-09-98-61011, July 2000.

Stipp (1999). David Stipp, "Hell No, We Won't Go!" *Fortune* (vol. 140 no. 2), July 19, 1999, pp. 102 *et seq.*

Szucs and Luebking (1988). Loretto Dennis Szucs and Sandra Hargreaves Luebking, *The Archives: A Guide to the National Archives Field Branches* (Salt Lake City UT: Ancestry Publishing).

Thatcher, Kannisto and Vaupel (1998). A. R. Thatcher, V. Kannisto and J. W. Vaupel, *The Force of Mortality at Ages 80 to 120*, Odense University Press, Monographs on Population Aging, 5. Electronic edition: http:/www.demogr.mpg.de/Papers/Books/Monograph 5.

Thoms (1873). William J. Thoms, *Human Longevity: Its Facts and Its Fictions* (London: John Murray). Revised edition, 1879 (London: Frederic Norgate). The 1873 edition is stamped "Longevity of Man" on its spine, which has resulted in many citations to this work using that title.
Wachter and Finch (1997). Kenneth W. Wachter and Caleb E. Finch, *Between Zeus and the Salmon: The Biodemography of Longevity* (Washington DC: National Academy Press).

Young (1899). T. E. Young, *On Centenarians, And the Duration of the Human Race: A Fresh and Authentic Enquiry; With Historical Notes, Criticisms and Speculations* (London: Charles and Edwin Layton). Revised edition, 1905.