Paper for the 2013 PAA Annual meeting, New Orleans, LA

Session 92: "*The Oldest Old: Determinants of Very Long Lifespans*"

Friday, April 12, 8:30 AM - 10:00 AM, Borgne Room, 3rd Floor

Determinants of Exceptional Longevity: Early-Life Conditions, Mid-Life Environment and Parental Characteristics

Leonid A. Gavrilov, Natalia S. Gavrilova

Center on Aging, NORC/University of Chicago
1155 East 60th Street, Chicago, IL 60637

Address for correspondence:
Dr. Leonid A. Gavrilov, Center on Aging
NORC at the University of Chicago
1155 East 60th Street, Chicago, IL 60637
E-mail: gavrilov@longevity-science.org

Determinants of Exceptional Longevity:
Early-Life Conditions, Mid-Life Environment and Parental Characteristics

Abstract

Earlier studies found that parental characteristics as well as early-life conditions and mid-life environment play a significant role in survival to advanced ages. However, little is known about simultaneous effects of all these three factors on longevity. This ongoing study attempts to fill this gap by comparing American centenarians born in 1890-1891 with their short-lived peers born in 1890 and died at age 65 years. The records were taken from computerized family histories, which were then linked to 1900 and 1930 U.S. censuses. The study found that parental longevity and some mid-life characteristics proved to be significant predictors of longevity while the role of childhood conditions was less important. More centenarians were born in the second half of the year compared to controls suggesting early origins of longevity. The results of this study suggest that familial background, early-life conditions and mid-life characteristics play an important role in longevity.

Keywords: exceptional longevity; maternal age; family histories; season-of-birth; centenarians

Introduction

Studies of centenarians (people living to 100 and older) could be useful in identifying factors leading to long life and avoidance of fatal diseases. Even if some individual characteristics have a moderate protective effect on risk of death, people with this trait/condition should be accumulated among long-lived individuals, because of cumulative survival advantage. Thus, study of centenarians may be a sensitive way to find genetic, familial, environmental, and life-course factors associated with lower mortality and better survival.

Most studies of centenarians in the United States are focused on either genetic (Hadley et al., 2000; Murabito et al., 2012; Perls and Terry, 2003; Sebastiani et al., 2012; Zeng et al., 2010) or psychological (Adkins et al., 1996; Hagberg et al., 2001; Margrett et al., 2010; Martin et al., 2010; Murabito et al., 2012) aspects of survival to advanced ages. On the other hand, several theoretical concepts suggest that early-life events and conditions may have significant long-lasting effect on survival to advanced ages. These concepts include (but are not limited to) the reliability theory of aging and the High Initial Damage Load (HIDL) hypothesis in particular (Gavrilov and Gavrilova, 2001b; Gavrilov and Gavrilova, 2003a; Gavrilov and Gavrilova, 2006); the theory of technophysio evolution (Fogel, 2004; Fogel and Costa, 1997); the idea of fetal origin of adult diseases (Barker, 1998; Kuh and Ben-Shlomo, 1997); and a related idea of early-life programming of aging and longevity (Gavrilov and Gavrilova, 2004). These ideas are supported by studies suggesting significant effects of early-life conditions on late-life mortality (Barker, 1998; Costa and Lahey, 2005; Elo and Preston, 1992; Finch and Crimmins, 2004; Fogel and Costa, 1997; Gavrilov and Gavrilova, 2003b; Hayward and Gorman, 2004; Kuh and Ben-Shlomo, 1997; Smith et al., 2009a). The role of early-life conditions in shaping late-life mortality is now well recognized and studies of centenarians can contribute to this area of research.

Our search for appropriate data resources for centenarian studies revealed an enormous amount of lifespan data that could be made readily available for subsequent full-scale studies (Gavrilov et al., 2002; Gavrilova and Gavrilov, 1999). Millions of genealogical records are already computerized and after their strict validation could be used for the study of familial and other predictors of human longevity. Computerized genealogies provide the most complete information on the lifespan of centenarians' relatives, when compared to other data sources, such as death certificates, census data and the U.S. Medicare database.

Studies of centenarians require serious work on age validation (Jeune and Vaupel, 1999; Poulain, 2010; Poulain, 2011) and careful design including the choice of an appropriate control group. In this article we consider several approaches to the choice of control population in centenarian studies: (1) use of population-based control, (2) the within-family analysis, as well as (3) the between-family analysis with selection of controls from the same population universe. These approaches are illustrated using data on American centenarians, their relatives and unrelated shorter-lived controls obtained from the same online genealogies.

1. Using population-based control in centenarian studies

Taking general population as a control group is one of the most popular approaches in centenarian studies. For example, the study of survival of siblings of centenarians used a U.S. cohort life table for comparison (Perls et al., 2002). A scientific group led by Preston suggested an original methodology to study longevity in the Unined States (Preston et al., 1998). The researchers collected individual death certificates for persons who died at ages 85+ during January 1-14, 1985. Death certificate data were then linked to the 1900 US census. Individual data from the 1900 U.S. census were used as a control group. Population-based census data are available as a part of the IPUMS project at the University of Minnesota (Ruggles et al., 2004).

We applied method suggested by Preston in our earlier study of centenarians found in computerized family histories (Gavrilova and Gavrilov, 2007). To this aim, the detailed family data for 991 alleged centenarians born between 1875 and 1899 in the United States were extracted from publicly available computerized family histories available at the Rootsweb website. In order to validate the age of the centenarians, these records were linked first to the U.S. Social Security Administration Death Master File records (for death date validation) and then to the records of the 1900, 1910 and 1920 U.S. censuses (for birth date validation). The results of this cross-validation study demonstrated that computerized genealogies may serve as a useful starting point for developing a reliable family-linked scientific database on exceptional human longevity. The resulting database on centenarians with validated ages was used in the study of the predictors of exceptional human longevity for a restricted sample of 358 centenarians born in 1890-1899 (Gavrilova and Gavrilov, 2007). This was done through the comparison of households where children (future centenarians) were raised (using data obtained through linkage of genealogies to early U.S. censuses) to control households drawn from the Integrated Public Use Microdata Series (IPUMS) for the 1900 U.S. census.

The IPUMS sample represents one percent of Caucasian households enumerated in 1900 (households where the head of household was Caucasian) (Ruggles et al., 2004). The linkage to early U.S. censuses in our study found that most centenarians in our sample were Caucasian (with the exception of two American Indian families), so we used a sample of Caucasian population from the IPUMS dataset as a control. At this initial stage of data analysis, we conducted a comparison of households that raised a future centenarian and linked to the 1900 census (358 cases) to the population-based sample of Caucasian households enumerated by the 1900 census, which had children below age 10 to make these households comparable to the set of centenarians who were born in 1890-1899 and hence were below age 10 in 1900 (31,322 control households).

We applied a method of multiple logistic regression in order to compare these two sets of households. The variables used for describing a household were similar to those applied by Preston et al. (1998). Our tested hypothesis was that if early childhood conditions are important for survival to age 100, then the households with children-future centenarians would be different from the general population.

The results of this study demonstrate that both the region of childhood residence and the household property status were the two most significant variables that affect the chances of a household to produce a future centenarian (for both sons and daughters). Spending a childhood in the Mountain Pacific and West Pacific regions in the U.S. were found to increase chances of long life (by a factor of three) compared to the Northeastern part of the country (Gavrilova and Gavrilov, 2007). Also a farm (particularly owned farm) residence in childhood was associated with better survival to advanced ages. These findings are consistent with the hypothesis that lower burden of infectious diseases during childhood expressed as lower child mortality in families of farm owners and families living in the West (Preston and Haines, 1991) may have far-reaching consequences for survival to extreme old ages. Some of these results are consistent with earlier studies of childhood conditions and survival to age 85+ (Hill et al., 2000; Preston et al., 1998). These studies, also based on linkage to early censuses, demonstrated a significant advantage in survival to age 85 for children living on farms for both African Americans (Preston et al., 1998) and native-born Caucasians (Hill et al., 2000). On the other hand, the Northeast and Midwest were found to be the best regions of childhood residence for subsequent survival to age 85+ (Hill et al., 2000).

Although the use of population-based controls facilitates the research and often helps to obtain reproducible results it has a serious limitation. If centenarians and controls are sampled differently then the results of the longevity study may be biased by factors unrelated to differential survival. For instance, genetic composition of centenarians and younger controls (often used in gene association studies) may be differently affected by migration. Centenarians in The New England Centenarian Study were found to be better educated compared to the general population (Perls et al., 2002). Nevertheless, the researchers compared survival of siblings of centenarians in this study to the general population. Although researchers adjusted their calculations for educational status there is still a possibility that the relative survival of siblings of centenarians could be overstated, because of residual confounding. In our earlier study we made an assumption that persons in computerized family histories do not differ from the general population (Gavrilova and Gavrilov, 2007) and found some agreement with other studies (Hill et al., 2000; Preston et al., 1998). However, there is still a possibility of biased results if this assumption is not true. For this reason, better research approaches to centenarian studies should be developed as it is considered in this article.

2. Comparative study of biological and non-biological relatives

Numerous studies found that biological relatives of long-lived individuals have substantial survival advantage compared to relatives of shorter-lived individuals (Gavrilov and Gavrilova, 2001a; Gavrilov et al., 2002; Kerber et al., 2001; Pearl and Pearl, 1934; Perls et al., 2007; Smith et al., 2009b; Willcox et al., 2006) while relatively few studies analyze lifespan of non-biological relatives (Mazan and Gagnon, 2007; Montesanto et al., 2011; Schoenmaker et al., 2006; Westendorp et al., 2009). At the same time, non-biological relatives may serve as a nonbiased environmental control group in contrast to the general population control. Information about biological relatives may be useful in another type of design called “the within-family-analysis” where siblings of centenarians serve as a control group. This type of design and its examples are described in the next section. Family histories (genealogies) proved to be a good source of information for different types of relatives and were successfully used in historical demography (Adams and Kasakoff, 1984; Anderton et al., 1987; Bean et al., 1992) and biodemography (Caselli et al., 2006; Gavrilov et al., 2002; Gavrilova et al., 1998; Smith et al., 2009a; Smith et al., 2009b).

Studies of biological and non-biological relatives usually deal with relatively small numbers of relatives. In our study, we intended to create a large sample of centenarians and their relatives, so we conducted a large-scale search in many hundreds of online family histories using a technique known as web automation (Sklar and Trachtenberg, 2002). This technique allowed us to search online databases on a large-scale basis for people with exceptional longevity (and some other traits). In particular, a technique is developed to scan more than 300,000 online databases in the Rootsweb WorldConnect project (http://wc.rootsweb.ancestry.com), a publicly available data source. Application of web-automation techniques to this online source identified more than 40,000 records of alleged centenarians born between 1880 and 1895 with known information about their parents.

After collecting data on centenarians, the next step was to collect detailed data on their parents from computerized genealogies using the same web-automation technique. After this procedure, we selected the most detailed genealogies where information on birth and death dates for both parents was available. Our prior experience working with computerized genealogies suggests that this procedure selects-out the majority of genealogies with poor quality. As a result of this procedure, the total number of centenarian records slightly decreased from 25,451 to 23,127 (see Table 1). In the next step, we collected data on siblings for those centenarians who had detailed data on parental birth and death dates. Using this procedure, we collected 172,091 records for siblings of centenarians. However, a significant proportion of these records did not contain information about the death dates of siblings, which created some difficulties for the within-family study of human longevity. So, the next step was to identify the most detailed data on families with complete information on birth and death dates for siblings. As a result of this identification procedure, we found 2,834 families where information on birth and death dates was known for more than 80 percent of siblings. Finally, the age at death of centenarians in our sample was verified using the U.S Social Security Administration Death Master File (DMF). This file contains information about birth and death dates as well as first/last names and residence for persons died in the United States. This procedure confirmed the age for 1,711 centenarians in our sample. Our previous study found that the ages of centenarians confirmed through the linkage to DMF are then usually confirmed through the linkage to early U.S. censuses (Gavrilova and Gavrilov, 2007), so there is no a strong need to make an additional verification with the early census records. Table 1 shows the number of records obtained in each stage of data collection.

[Table 1 About Here]

Finally, 1,711 validated centenarians born in 1880-1895 with their death dates verified using the Social Security Death Index are used for further analyses. The database of centenarians and their relatives contains information on lifespan for 398 male and 1,313 female centenarians, their 13,419 siblings, 1,307 spouses, and 7,924 siblings-in-law. This database is used for comparative analysis of longevity in different types of biological and non-biological relatives.

Each centenarian has about 7.8 siblings on average. The total sibship size (nine siblings on average) in the studied centenarian families is higher than the average number of children in American families reported by the 1900 U.S. Census: 5.12 ± 0.01; data obtained from the 5% sample of the US 1900 Census from the Integrated Public Use Microdata Series or IPUMS (Ruggles et al., 2010). Larger sibship size in the centenarian families compared to the general population can be explained by the fact that genealogies are more likely to be compiled for larger families and that longer-lived individuals in the United States were born more often in rural areas with higher fertility (Gavrilova and Gavrilov, 2007; Preston et al., 1998). This difference in sibship size with the general population is not critical for the within-family design of the study when an appropriate control group (shorter-lived siblings raised in the same family or spouses) is selected.

Spouses of male centenarians are 5.07 years younger, and spouses of female centenarians are 3.37 years older on average than their long-lived mates. This asymmetry in spousal age gap between male and female centenarians is statistically significant: the 95% confidence interval for mean spousal age gap is 4.40 - 5.74 years in male centenarian couples, and only 3.01 - 3.73 years in female centenarian couples. Such asymmetry in spousal age gap could happen if during their marriage ages the future centenarians were perceived at the marriage market as persons of younger age than they really were, thus obtaining a younger mate than typically expected. This phenomenon of delayed or postponed aging of future centenarians on marriage market is consistent with our other findings that centenarians tend to marry at later ages then usual (see section 4 of this article).

Comparison of mean lifespan for relatives survived to age 50 reveals a survival advantage of brothers and sisters of centenarians compared to the same-sex siblings-in-law: brothers lived 2.6 and sisters lived 2.9 years longer on average compared to siblings-in-law of corresponding sex (Table 2) with differences in lifespan being statistically significant ($p<0.001$).

[Table 2 About Here]

Table 1. Number of centenarians and their siblings at different stages of data collection and cleaning

Type of records	Number of Centenarians			Number of Siblings
	Males	Females	Total	
All initial non-duplicate records for centenarians born in 1880-1895 with names of parents available	7,174	18,277	25,451	
Centenarians having detailed information on birth and death dates of their parents	6,370	16,757	23,127	172,091
Centenarians having detailed information on birth and death dates of their parents and siblings	707	2,127	2,834	21,893
Centenarians after data cleaning with confirmed death dates through the linkage to U.S. Social Security Death Master File, DMF	398	1,313	1,711	13,419

Table 2. Mean life span conditional on survival to age 50 (LS50) with 95% confidence intervals (95% CI) for relatives of centenarians (compared to the 1900 US birth cohort).

Relatives of centenarians	Men		Women	
	Sample size, N	LS50 (95% CI), years	Sample size, N	LS50 (95% CI), years
Parents	1590	76.2 (75.7-76.8)	1557	77.2 (76.7-77.8)
All Siblings	5324	77.6 (77.3-77.9)	4877	82.4 (82.0-82.7)
Married Siblings	3221	77.7 (77.3-78.1)	3028	82.2 (81.8-82.6)
Spouses	876	75.4 (74.6-76.1)	283	81.4 (80.1-82.7)
Siblings-in-law	2349	75.0 (74.6-75.5)	2407	79.5 (79.0-79.9)
1900 US birth cohort		73.3		79.4

Wives of centenarians tend to live 0.8 year less on average than married sisters of centenarians, however this difference is not statistically significant. Husbands of centenarians live 2.3 years less on average than married brothers of centenarians (the difference is statistically significant, $p<0.001$). Although fathers of centenarians are born about 30 years earlier than brothers-in-law of centenarians they still have higher lifespan conditional on survival to age 50 than later-born non-biological relatives such as siblings-in-law ($p<0.001$) and husbands of centenarians ($p=0.04$). On the other hand, mothers of centenarians survived to age 50 have the shortest lifespan among all relatives – 77.2 years on average. Overall, siblings-in-law have the lowest lifespan compared to biological relatives and spouses born in a similar time period. At the same time, lifespan of siblings-in-law is still higher than mean lifespan of the general population (1900 U.S. birth cohort). This difference is particularly high for men – 1.7 years ($p<0.001$) while for women this difference is not statistically significant (Table 2). This finding is particularly important, because it indicates that using general population as a control group to compare survival of siblings of centenarians or other biological relatives may overstate survival advantage of biological relatives and hence overstate the genetic effect on lifespan. Thus, the use of general population to compare survival of relatives of long-lived individuals may be inappropriate, particularly for males. Therefore, siblings-in-law is a better control group for comparison than the general population.

Although the positive association of person's longevity with better survival of biological relatives is well documented, little is known about the effects of centenarian's gender on longevity of biological and non-biological relatives. In this article, we use the effects of centenarian's gender to explore genetic and environmental effects on longevity. We found that fathers of male centenarians lived significantly longer than fathers of female centenarians: mean life span conditional on survival to age 50 was 77.22 years vs 75.93 years ($p=0.023$, see Table 3). This effect is gender-specific and is observed for fathers of male centenarians but not for mothers of male centenarians (who have similar survival as mothers of female centenarians). Brothers of male centenarians also lived significantly longer compared to brothers of female centenarians: mean life span conditional on survival to age 50 is 79.25 and 77.09 years respectively ($p<0.001$, Table 3). Again, this effect is gender-specific: sisters of female centenarians have no statistically significant survival advantage over sisters of male centenarians (82.45 vs 82.06 years, $p=0.836$, Table 3). This finding is also supported by comparison of siblings' sex ratio in families of male and female centenarians. The sex ratio for siblings surviving to age 50 years is higher (more males) in families with male centenarian (sex ratio = 1.22), when compared to the sex ratio in families with female centenarian (sex ratio = 1.10).

[Table 3 About Here]

Figure 1 shows survival curves after age 30 for male siblings of centenarians depending on centenarian gender. Note that brothers of male centenarians have substantially better survival than brothers of female centenarians and this difference in survival is highly statistically significant ($p<0.001$) according to generalized Wilcoxon test (Breslow, 1970). This survival advantage is particularly strong after age 65 years while before the age 50 years the differences in survival are not observed. Thus, having a centenarian brother is associated with better late-life survival for males.

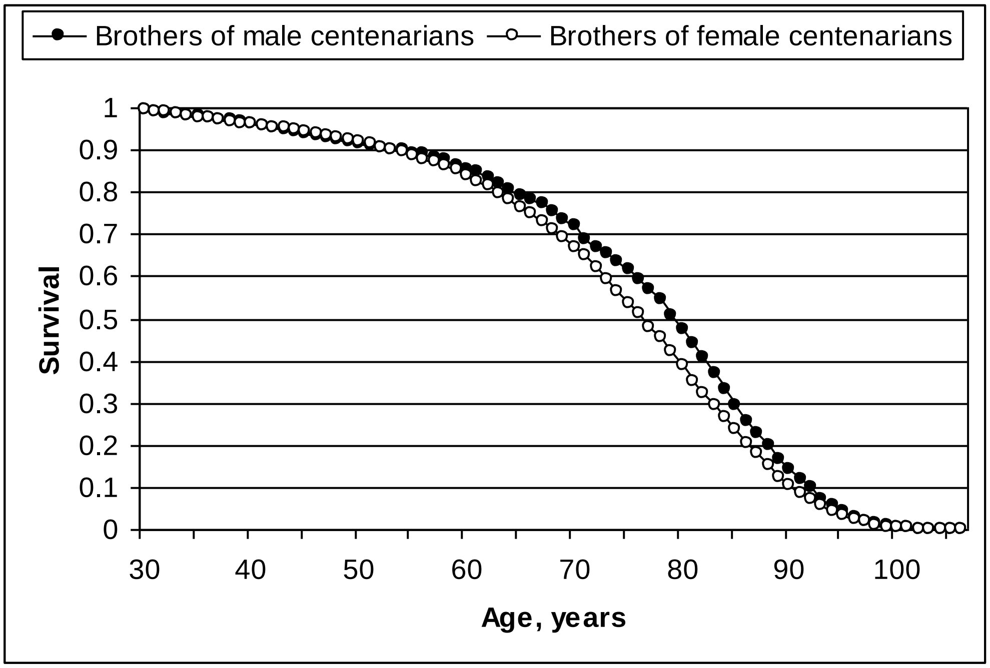
[Figure 1 About Here]

Table 3. Mean life span conditional on survival to age 50 (LS50) for biological and non-biological relatives of centenarians, by gender of centenarian.

Relatives of centenarians	Male centenarians		Female centenarians		p-value†
	Sample size N	LS50, years	Sample size N	LS50, years	
Parents					
Fathers	374	77.22	1216	75.93	0.023
Mothers	362	77.96	1195	77.03	0.087
Siblings					
Brothers	1268	79.25	4056	77.09	<0.001
Sisters	1071	82.06	3806	82.45	0.836
Siblings-in-law					
Brothers-in-law	492	74.00	1857	74.55	0.577
Sisters-in-law	611	79.22	1796	79.55	0.730

†p-values are related to the difference between life spans of corresponding relatives of male vs female centenarians

Figure 1. Survival of male siblings (brothers) of centenarians, by centenarian gender.



Note: Difference between survival curves is statistically significant ($p<0.001$ according to generalized Wilcoxon test).

Taking into account that female gender of centenarian has a much weaker effect on survival of sisters compared to the effect of male gender of centenarian on the survival of brothers, we may hypothesize that male centenarians and their brothers share living conditions and lifestyle favorable for men. To test this hypothesis, we used data on mean life span for siblings-in-law of centenarians as a control group. Siblings-in-law do not share specific genes and family environment with centenarians but they usually come from similar socio-economic background (because of assortative mating). Table 3 shows mean life span conditional on survival to age 50 years for siblings and siblings-in-law of centenarians depending on centenarian gender. If survival advantage of brothers of male centenarians is related to better socio-economic status of their families compared to families of brothers of female centenarians then it may be expected that wives of brothers of male centenarians (siblings-in-law) would also have better survival than siblings-in law of female centenarians. Note that centenarian gender has no effect on life expectancy of siblings-in-law (Table 3). This result suggests that survival advantage of brothers of centenarians is not related to better socio-economic status of families of male centenarians and their brothers compared to brothers of female centenarians.

One hypothesis explaining the survival advantage of brothers of male centenarians may suggest stronger genetic influence on male longevity. This hypothesis of stronger heritability of longevity among males was recently suggested by Italian researchers who studied 202 nonagenarians born around 1910 and their relatives in the province of Calabria in Italy (Montesanto et al., 2011). The researchers found that brothers and sisters of male nonagenarians have better survival compared to siblings-in-law of male nonagenarians. On the other hand, only brothers of female nonagenarians have reduced mortality compared to same-sex siblings-in-law. The authors made a conclusion that genetic factors in males have a higher impact than in females on attaining longevity (Montesanto et al., 2011). In our study we found that both brothers and sisters have higher longevity compared to siblings-in law regardless of centenarian gender. Only male biological relatives of male centenarians demonstrated survival advantage compared to biological relatives of female centenarians. These differences between the earlier study and our results may be explained by relatively small sample in the Italian study (202 nonagenarians with only 76 males) and largely rural and underdeveloped Calabrian society with strong social differences (Montesanto et al., 2011). Although our results may point to genetic factors as an explanation of the observed effects of male centenarian gender, we need to consider alternative non-genetic explanations as well.

One possible nongenetic explanation of the observed phenomenon comes from the past family traditions. In the past, men often continued to live in the place of their childhood while women more often left parental household after marriage. Favorable living conditions and/or lifestyle of male centenarians could be more likely shared by their brothers (rather than sisters) as well as by their spouses. If this hypothesis is correct, then spouses of male centenarians (wives) should have higher life expectancy compared to wives of brothers of centenarians. Results presented in Table 4 confirm this hypothesis. Thus, we may suggest that living in household with male centenarian has positive effect on survival of wives (presumably because of better environment and healthier lifestyle determined by the head of household in the past). To illustrate this point, consider, for example, that smoking and binge drinking in households in the past were driven more often by husbands' habits rather than habits of their wives.

[Table 4 About Here]

Table 4. Mean life span conditional on survival to age 50 (LS50) for spouses of centenarians and spouses of siblings of centenarians, by gender.

Relatives	Spouses of centenarians		Spouses of siblings of centenarians		p-value†
	Sample size N	LS50, years	Sample size N	LS50, years	
Husbands	876	75.38	2349	75.04	0.221
Wives	283	81.40	2407	79.46	0.004

†p-value is related to difference between mean life span of spouses of centenarians vs spouses of siblings of centenarians.

These findings are also consistent with our previous results as well as results of other studies, which found positive effects of farming and farm background on late-life survival (Gavrilova and Gavrilov, 2007; Preston et al., 1998). Farm childhood background turned out to be particularly favorable for men who usually continue to work on a farm. This study suggests that a significant portion of lifespan advantage for siblings of centenarians may be related to better lifestyle and living conditions rather than pure genetic effects only (otherwise, wives of centenarians would not benefit much from husbands' longevity).

3. The Within-Family Approach:

How centenarians are different from their shorter-lived siblings?

One promising approach to the choice of control group in centenarian studies is to use shorter-lived siblings of centenarians as controls. This within-family approach allows researchers to eliminate between-family variation including the differences in genetic background and childhood living conditions. We used a sample of 1711 centenarian larger families described in the previous section to explore the effects of early-life factors (birth order, paternal age, maternal age, month of birth) on the likelihood of survival to advanced ages. Centenarians (cases) were compared to their "normal" shorter-lived siblings (controls) using a within-family approach. During the process of data inspection, we found that some siblings were born after 1910 and their death dates were not indicated (hence, they potentially could become centenarians). To avoid this kind of data truncation, we used data for our index centenarians who were born earlier, between 1880 and 1889 rather than in 1880-1895. For this subgroup, there were no siblings born after 1910 with unknown death dates.

The study applied a case-sibling design, a variant of a matched case-control design in which siblings of cases (long-lived individuals) are used as controls (Woodward, 2005). This approach allows investigators to study within-family differences, not being confounded by between-family variation and unobserved between-family differences. Long-lived people born mostly between 1880 and 1889 are used as cases. Siblings were born between 1850 and 1910. The main approach used in this study is based on comparison of children within rather than across families. Within a family, children are born to parents at different ages and this variation may be used to estimate the net effect of parental age more conclusively (Kalmijn and Kraaykamp, 2005). Similarly, we can estimate the net effects of birth months.

The statistical analyses of within-family effects are performed using a conditional multiple logistic regression model (fixed-effect model) to investigate the relationship between an outcome of being a case (long-lived person) and a set of prognostic factors (Breslow and Day, 1993; Hosmer and Lemeshow, 2001). Only within-family variation is taken to influence the uncertainty of results (as reflected in the confidence interval) of a within-family study using a fixed-effect model. Variation between the estimates of effect from each family (heterogeneity) does not affect the confidence interval in a fixed-effect model. The likelihood to survive to advanced ages (to be in the long-lived group) is used as a dependent variable. Analyses were conducted using Stata Statistical Software, Release 11 (StataCorp, 2009). The following variables were included in the model: birth order, paternal age, maternal age, month of birth and sex (male or female). We found no statistically significant effects of birth order on the chances to survive to advanced ages in this particular data sample when parental age was also included in the analysis (data not shown).

We explored the role of the father's age as a potential predictor for survival to age 100. When the first child is born, the father is younger and can provide resources for a longer period than for his later-born children. We found that siblings born to fathers younger than 40 years had higher chances to survive to 100 than siblings born to older fathers (50 and older). However, control for maternal age decreased this dependence and made it statistically insignificant. This suggests that the apparent effects of a young father's age on exceptional longevity are driven by the correlated effects of a young mother. For this reason, we focused our study on the effects of maternal age.

In the next step, we included into analysis maternal age at birth, and it turned out that a young maternal age at childbirth was the most important predictor of exceptional survival, while the effects of paternal age at birth have become statistically insignificant (Table 5, Model 2). We found that the odds to live to 100 are 1.6 times higher for children born to younger mothers compared to siblings (brothers and sisters) born to mothers older than age 40 in the same families and even after controlling for paternal age (see Table 5). For separately studied daughters, maximum chances of survival to 100 shifted from the youngest maternal ages (<20 years) to age group 25-29 years (Table 5, Model 3). For separately studied sons, effects of maternal age have become even stronger, particularly at maternal age 20-24 years (odds ratio=1.77, Table 5, Model 4).

[Table 5 About Here]

Next question explored was about the role of child mortality, which was very high a century ago, when the studied centenarians were born. To avoid the effects of child mortality on survival, we re-analyzed our data including only those siblings who survived to older ages. We found that for people who survived to age 50 younger maternal age still remains a significant predictor of longevity: the odds to live to 100 are 1.57 times higher for those born to mothers younger than 25 years compared to siblings born to 40-year-old mothers (Table 6). Moreover, even at age 70, it still helps to be born to a young mother—the odds to celebrate the 100th birthday are 1.6 times higher for siblings born to mothers younger than 20 years compared to those born to 40-year-old mothers (Table 6).

[Table 6 About Here]

Table 5. Effects of maternal age at person's birth on human longevity. Odds ratios (with p-values) to live to 100 as predicted by conditional logistic regression (fixed effects)†.

Variable	Model 1	Model 2	Model 3 Daughters only	Model 4 Sons only
Maternal age:				
<20	1.60 (0.022)	1.66 (0.029)	1.43 (0.121)	1.72 (0.162)
20-24	1.49 (0.007)	1.51 (0.013)	1.37 (0.067)	1.77 (0.042)
25-29	1.46 (0.008)	1.44 (0.018)	1.57 (0.006)	1.24 (0.435)
30-34	1.13 (0.404)	1.12 (0.492)	1.07 (0.708)	1.29 (0.360)
35-39	1.05 (0.747)	1.04 (0.814)	1.10 (0.552)	0.92 (0.769)
40+	Reference	Reference	Reference	Reference
Paternal age:				
<25		0.90 (0.501)		
50+		0.94 (0.798)		
Female sex	3.21 (<0.001)	3.21 (<0.001)		
Pseudo R^2	0.0691	0.0693	0.0062	0.0109
Number of observations	6,413	6,413	4,732	1,681

†Calculated using Stata 11 statistical package (procedure clogit).

Table 6. Effects of maternal age on human longevity. Odds ratios (p-values) to live to 100 as predicted by conditional logistic regression (fixed effects) for different subgroups[†].

Variable	Siblings survived to age 50	Siblings survived to age 70	Siblings survived to age 20; smaller family size (<9)	Siblings survived to age 20; larger family size (9+)
Maternal age:				
<20	1.57 (0.029)	1.61 (0.026)	2.33 (0.012)	1.29 (0.337)
20-24	1.52 (0.006)	1.53 (0.006)	1.94 (0.010)	1.29 (0.164)
25-29	1.46 (0.008)	1.48 (0.009)	1.63 (0.049)	1.41 (0.049)
30-34	1.15 (0.368)	1.13 (0.446)	1.13 (0.631)	1.17 (0.380)
35-39	1.05 (0.757)	1.05 (0.769)	1.26 (0.369)	0.94 (0.766)
40+	Reference	Reference	Reference	Reference
Female sex	3.25 (<0.001)	2.95 (<0.001)	3.34 (<0.001)	3.11 (<0.001)
Pseudo R^2	0.0731	0.0645	0.0832	0.0625
Number of observations	5,778	4,813	2,352	4,061

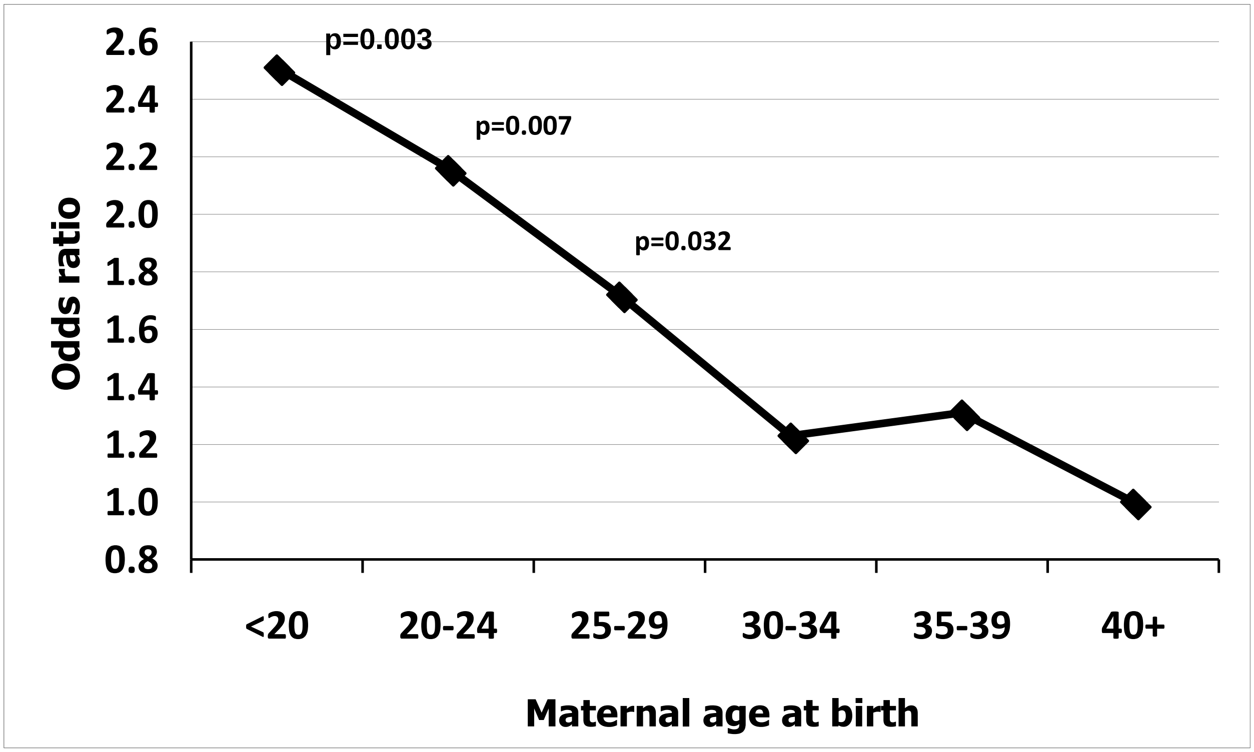
[†]Calculated using Stata 11 statistical package (procedure clogit).

Another studied question is related to family size. Families in our dataset are rather large with median size of nine children and with some families having up to 18 children. So we divided families where centenarians were born into smaller families (with less than nine children) and larger families (with nine children and more). It should be noted that in many larger families, some siblings died in infancy or early childhood. We found that the effect of maternal age on survival is significantly higher in smaller families compared to larger families. In smaller families, siblings born to mothers younger than 20 years had more than twice the chances to live to 100 compared to their brothers and sisters born to 40-year-old mothers (Figure 2 and Table 6). Thus, the within-family analysis of the paternal- and maternal-age effects on human longevity demonstrated that a young age of the mother increases the chances of children to reach longevity.

[Figure 2 About Here]

The finding of a beneficial effect of young maternal age on offspring survival to age 100 in humans may have biological explanation. There is empirical evidence that the quality of female eggs in human beings rapidly declines with age (Bickel, 2005; Pellestor et al., 2005) and this deterioration starts rather early—before age 30 (Heffner, 2004). Maternal age influences the biology of the mother-fetus relationship, with a negative effect on fetal development and predisposition to severe diseases such as type I diabetes (Gloria-Bottini et al., 2005).

Figure 2. Effects of maternal age at person's birth on odds to live to 100 conditional on survival to age 50.



The within-family study of 2,153 centenarians and their siblings <u>who survived to age 50</u>. Data on smaller families with less than nine children. Calculated using Stata 11 statistical package (procedure clogit).

Experiments on laboratory mice found that offspring born to younger mothers live longer (Tarin et al., 2005). This study also demonstrated that the largest survival advantage of offspring born to younger mothers is observed at later life. Beneficial effects of younger maternal age on offspring survival were observed in another study of mice, but only for females (Carnes et al., 2012). In the case of birds (*Parus major*), offspring hatched from older mothers suffer from an earlier onset, and stronger rate, of reproductive senescence later in life (Bouwhuis et al., 2010). Animal studies found that hormonal profiles in pregnant mice are different depending on maternal age (Wang and vom Saal, 2000). This may explain why adult mouse offspring of adolescent and middle-aged mothers have lower body weight and delayed puberty; and male offspring have smaller reproductive organs than those born to younger mothers (Wang and vom Saal, 2000). Female offspring produce progeny whose birth weight depended on the age at pregnancy of their grandmothers, demonstrating a transgenerational effect of maternal age (Wang and vom Saal, 2000). Delayed motherhood in mice has also been demonstrated to have negative effects on behavioral traits of young adult offspring (Tarin et al., 2005). Data on the long-term effects of maternal age in human beings are scarce. One study found that the lifespan of children decreased with increasing maternal age (Kemkes-Grottenthaler, 2004). Our earlier studies did not detecte an association of maternal age with offspring mortality in historical populations of European aristocracy (Gavrilov and Gavrilova, 1997; Gavrilov and Gavrilova, 2000), but we believe this might be due to limitations in the data or the tools to analyze them. Specifically, our earlier studies were focused on mean life span rather than exceptional longevity and did not apply the within-family analysis.

The fact that lifespan of offspring depends on the mother's age at their birth even in laboratory animals indicates that some fundamental biological mechanisms may be involved. Such possible epigenetic mechanisms as changes in genomic imprinting in oocytes of aging females may be a plausible hypothesis (Comings and Macmurray, 2001; Comings and MacMurray, 2006). Another plausible biological hypothesis is the telomere theory of reproductive senescence in females (Keefe et al. 2005), which posits that eggs ovulating from older females have shorter telomeres because of late exit from the oogonial “production line” (Polani and Crolla, 1991) during fetal life, with incomplete restoration by telomerase (Keefe et al., 2005). Telomeres are DNA repeats that cap and protect chromosome ends, so that longer telomeres in eggs of younger females may be beneficial for offspring lifespan. However, in human beings, some additional sociobehavioral mechanisms may be also involved, on top of more general biological mechanisms.

In this study, we used a large sample of centenarians and their siblings to study early-life effects on human longevity. This study shows good agreement with our previous results obtained using another dataset with significantly smaller sample size (Gavrilov and Gavrilova, 2012; Gavrilova and Gavrilov, 2007). We found significant positive effects of a young maternal age on survival to age 100 with maximum effect observed predominantly at age 20 to 24 years. Paternal age effects were also observed, but they were mainly driven by correlated young maternal age. Effect of a young mother is particularly prominent in smaller families, which is important taking into account smaller family sizes in contemporary population.

In addition to maternal age, the within-family approach can be applied to the season of birth study of exceptional longevity. Month of birth is a useful proxy characteristic for environmental effects acting during in-utero and early infancy development. To analyze net effects of birth month on exceptional longevity, not confounded by possible changes in birth and infant death seasonality, childhood conditions and genetic background, we conducted a matched study using a multivariate conditional logistic regression method (Gavrilov and Gavrilova, 2011). In this study, we used the whole sample of centenarians born between 1880 and 1895. To discriminate between effects due to differential survival early in life from effects of birth month acting later in life, we analyzed survival to age 100 among siblings conditional on their survival to different adult ages. The results demonstrate that people born in September to November have significantly higher chances of exceptional longevity than people born in March. This month-of-birth effect is observed even for siblings who survived to age 70, suggesting a very long-lasting influence of season of birth on longevity (Gavrilov and Gavrilova, 2011).

These results are in agreement with previous publications on the effects of birth month on lifespan in the Northern hemisphere (Abel and Kruger, 2010; Doblhammer and Vaupel, 2001; Doblhammer et al., 2005; Gavrilov and Gavrilova, 1999; Lerchl, 2004; Vaiserman et al., 2009) and in the United States in particular (Doblhammer, 2004; Gavrilov and Gavrilova, 2008). These studies found better survival for people born in September through December compared to people born in the middle of the year. At the same time, our study does not demonstrate significant differences in survival for siblings born during other seasons. This does not agree with some other studies, which found decline in mean age at death for people born in the summer months and relatively high mean age at death for people born in winter months (Doblhammer, 2004; Lerchl, 2004; Vaiserman et al., 2009). These differences in the month-of-birth pattern between our study and other publications can be partially explained by changes in seasonality of births and seasonality of infant mortality over time. Studies based on the analysis of cross-sectional death records do not have information about population at risk (Doblhammer, 2004) and hence may be affected by secular changes in seasonality of births and infant deaths. Although these secular effects probably do not distort completely the entire month-of-birth pattern in life expectancy, they can modulate amplitudes observed for specific months. It would be reasonable to suggest that a decreasing trend of summer infant deaths led to over-representation of summer-born individuals in later-born age groups, leading to an apparent drop in the mean age at death for these months. Our study, based on cohort survival analysis, is not affected by secular changes in seasonality of births and infant deaths (Gavrilov and Gavrilova, 2011).

There are several explanations of season-of birth effects on longevity pointing to the effects of early-life events and conditions: seasonal exposure to infections, seasonal nutritional deficiencies, environmental temperature and sun exposure. All these factors were shown to play role in later-life health and longevity. More details on the within-family study of seasonal effects on longevity can be found elsewhere (Gavrilov and Gavrilova, 2011).

The results obtained in the within-family studies demonstrate that factors acting early in life may have significant long-lasting effects on survival to advanced ages. These results are consistent with the reliability theory of aging and the High Initial Damage Load (HIDL) hypothesis in particular (Gavrilov and Gavrilova, 1991; Gavrilov and Gavrilova, 2004), which emphasizes the importance of the initial level of damage in determining future human longevity. More specific explanation of the observed effects of early-life conditions on longevity can be provided by the inflammation hypothesis suggested by Finch and Crimmins (2004). According to this hypothesis, a strong acute-phase inflammatory response required for survival early in life initiates chronic inflammation, which promotes chronic diseases of aging. The results obtained in our study suggest that optimizing the process of early-development can potentially result in avoiding many diseases in later life and significantly extending healthy life span.

The within-family approach has great advantages over other methods because it is free of confounding caused by between-family differences. At the same time, this approach allows researchers to analyze a limited number of predictor variables. Only variables that vary across siblings in the family can be analyzed. The within-family approach described in this section can be extended further to include adulthood variables in addition to the early-life conditions.

4. The Between-Family Approach:

Sampling Centenarians and Genetically Unrelated Controls from the Same Population Universe

In this study we compare centenarians born in the United States to their peers belonging to the same birth cohort who were also born in the United States and died at age 65 years. Both cases and controls were randomly sampled from the same population universe (computerized family histories) and had the same birth year window (1890-91). These records were then linked to historical U.S. censuses (1900, 1910, 1930). The main focus of the study is on the 1900 and 1930 censuses that correspond to the childhood and adulthood periods of their individual lives. The age at death for controls is selected assuming that the majority of deaths at age 65 occur due to chronic age-related diseases rather than injuries or infectious diseases.

Sample sizes of male centenarians are small in the majority of longevity studies, and in order to resolve this problem and to have a sample balanced in regard to gender, males are oversampled in this study. This oversampling does not affect the analyses because male and female data are studied separately taking into account that men and women may respond differently to the same set of risk factors. In order to obtain a more homogeneous birth cohort regarding the secular changes in mortality and life course events, a narrow birth date window was used: 1890-91.

Prevalence of centenarians in modern populations is very low: about 1 per 10,000 population (Hadley et al., 2000), and therefore traditional methods of population sampling are difficult and not feasible for obtaining large samples of centenarians. Case-control design proved to be the most appropriate and cost-effective approach for studies of rare conditions (Breslow and Day, 1993; Woodward, 2005) and hence is extremely useful for centenarian studies. Breslow and Day (1993) suggested that the classic case-control design can be expanded in a variety of ways. One such expansion is a design suggested by Samuel H. Preston (Preston et al., 1998). According to this design, a survival to advanced ages (rather than disease or death) is considered to be a case and relative survival probabilities are used instead of odds ratios. In this study we draw centenarians and controls randomly from the same universe of online family histories in order to ensure comparability and avoid possible selection bias when centenarians and controls are drawn from different populations. Also we used data from historical sources collected when centenarians and controls were children or young adults thereby avoiding a limitation related to self-report or recall bias. Only records from genealogies of presumably good quality with available information on exact (day, month, year) birth dates and death dates (for centenarians) as well as information on birth and death dates of both parents are used in the sampling procedure for both cases and controls.

Persons born in 1890-91 represent an interesting birth cohort to study. These individuals experienced high exposure to infections during childhood and decreasing infectious disease load later in life. It is important to note that nonagenarians and centenarians living now in the United States have very similar experiences as persons born in the end of the 19^{th} century. Therefore, more detailed analysis of past history and life course of this birth cohort may be important for understanding the underlying factors and causes of mortality among the currently living old age cohorts.

Centenarians represent a group with really rare condition of successful survival (only two men and 14 women out of 1000 from 1900 US birth cohort survived to age 100) but common enough for obtaining samples of sufficient size. In this study we analyzed early-life and adulthood effects that operate throughout the life by comparing centenarians of each gender to the respective control groups.

Data quality control procedure in this study was similar to the procedure described in section 2 and included: (1) preliminary quality control of computerized family histories (data consistency checks); (2) verification of the centenarian's death date; (3) verification of the birth date (for centenarians and controls); (4) verification of family information (parents, spouses and siblings). These methods of age validation were based on the approaches proposed by the experts in this area (Jeune and Vaupel, 1999; Poulain, 2010) and our own research experience. All records (for centenarians and controls) were subjected to verification and quality control using several independent data sources. Our primary concern was the possibility of incorrect dates reported in family histories. Previous studies demonstrated that age misreporting and age exaggeration in particular are more common among long-lived individuals (Elo et al., 1996; Hill et al., 2000; Rosenwaike and Stone, 2003; Shrestha and Rosenwaike, 1996). Therefore, the primary focus in this study was on the age verification for long-lived individuals, which involved death date verification using the U.S. Social Security Administration Death Master File (DMF), and birth date verification using early U.S. censuses.

According to our experience, the linkage to DMF selects out the majority of incorrect records for alleged centenarians (Gavrilova and Gavrilov, 2007). Definite match was established when information on first and last names (spouse last name for women), day, month and year of birth matches in DMF and family history (Sesso et al., 2000). In the case of disagreement in day, month or year of birth, the validity of match is verified on the basis of additional agreement between place of the last residence and place of death.

The procedure of death date verification using DMF is not feasible for validating death dates of controls, because data completeness of DMF is not very high for deaths occurred before the 1970s. We found that approximately 30 percent of deaths in the control group could be confirmed through the US state death indexes, cemetery records and obituaries, which cover longer periods of time. Taking into account that exact ages of death for controls are not particularly important for the study design, it is possible to rely on death date information recorded in family histories for controls not found in external sources (as it was done in the Utah Population Database for individuals died before 1932 (Kerber et al., 2001)).

Verification of birth dates was accomplished through a linkage to the 1900 U.S. census data recorded when the person was a child (when age exaggeration is less common compared to claims of exceptional longevity made at old age). The preference is given to the 1900 census because it is more complete and detailed in regard to birth date verification (contains information on month and year of birth) compared to the 1910 and 1920 censuses. If person cannot be found in 1900 census, then he/she was searched in 1910 census. We obtained a good linkage success rate (92-95%) in our study, because of availability of powerful online indexes provided by the Ancestry.com service and supplemental information in family histories (Gavrilova and Gavrilov, 2007). These indexes allowed us to conduct search on the following variables: first, last names (including Soundex), state, county, township, birthplace, birth year (estimated from census), immigration year, relation to head-of-household. Data on birth dates, birth places and names of siblings produced unambiguous matches in overwhelming majority of cases.

Ancestry.com has a powerful search engine, which helps researchers to find a person in multiple historical sources simultaneously (including all historical U.S. censuses, up to 1940 census) based on all information available in computerized genealogies. Use of this service greatly facilitates the linkage procedure and helps to obtain unambiguous link in practically all studies cases. After the linkage to early censuses, the final database on centenarians and controls combined information on family characteristics (taken from family histories), data on the early-life conditions taken from the 1900-1910 U.S. censuses and adult socio-economic status taken from the 1930 Census. Eearly U.S. censuses contain a rich set of variables, which can be used to study the effects of both childhood and adulthood living conditions on human longevity (see Table 7).

[Table 7 About Here]

Below we summarize the core topical domains of the variables analyzed in this study.

Childhood living conditions at household level. This information was obtained from 1900 and 1910 censuses. Selection of variables was guided by the results obtained in the previous studies on child mortality at the turn of the 20th century (Preston and Haines, 1991). These studies demonstrated that child mortality is affected by household structure (including presence of a boarder in household), paternal occupation, mother's work, the occupation of household head, maternal and paternal literacy, family structure (whether the proband lived with both parents; his/her father and stepmother; a stepfather and mother; his/her father only; mother only; on his/her own – for example, in orphanage) (Preston and Haines, 1991). An important factor of survival to advanced age is childhood farm residence – a result found in our earlier study (described in section 1) as well as in other studies (Hill et al., 2000; Preston et al., 1998).

Table 7.

Information available in early U.S. censuses for the search of longevity predictors.

Variables	**Early U.S. census**						
	1860	**1870**	**1880**	**1900**	**1910**	**1920**	**1930**
Age, sex, color/race	+	+	+	+	+	+	+
Month and year of birth				+			
Marital status			+	+	+	+	+
Marriage duration (for married)				+	+		+
Literacy	+	+	+	+	+	+	+
School attendance (for children)	+	+	+	+	+	+	+
Place of birth	+	+	+	+	+	+	+
Places of birth for parents			+	+	+	+	+
Parental nativity		+	+	+	+	+	+
Mother tongue						+	+
Home ownership				+	+	+	+
Farm status				+	+		+
Value of real and personal estate	+	+					+
Number of children born and surviving (for women)				+	+		
Whether deaf and/or dumb					+		
Radio in household							+
Occupation	+	+	+	+	+	+	+
Employment			+	+	+	+	+
Citizenship		+		+	+	+	+
Year of immigration				+	+	+	+
Veteran status					+		+

Infectious burden. The main hypothesis we studied here is that the early exposure to infections decreases chances of survival to advanced ages affecting mortality later in life. Infectious burden is estimated as the within-family infectious burden. Information on children ever born and children surviving allowed us to estimate proportion of surviving children for each family, where biological mother is present. Child mortality served as a proxy of infectious disease burden in the family characterizing the living environment, as suggested by other researchers (Bengtsson and Lindstrom, 2000; Bengtsson and Lindstrom, 2003; Finch and Crimmins, 2004; Preston and Haines, 1991). We based our estimates of child mortality on information available in 1910 census whenever possible, because by this time the majority of studied mothers finished their reproductive period.

Seasonal early-life conditions. Effects of seasonal conditions on survival to extreme ages are studied using month of birth as an integral proxy for environmental seasonal conditions (e.g., seasonal infections) before and shortly after the birth. Existing literature on the U.S. mortality and our own results based on the within-family approach (see section 3) shows that month of birth may be a significant predictor of mortality not only during childhood but also in later life (Costa and Lahey, 2005; Doblhammer, 2004; Doblhammer and Vaupel, 2001; Gavrilov and Gavrilova, 1999; Gavrilov and Gavrilova, 2011).

Adulthood social conditions. Socioeconomic achievement at adult ages for men was estimated using occupation status and dwelling ownership status (measured as in 1900 Census). In particular, we tested a hypothesis that farm background is particularly favorable for male survival because sons of farmers also become farmers (Preston et al., 1998). In this case the farm status in both 1900 and 1930 should bring a significant advantage for survival to 100. In the case of females, estimation of socioeconomic achievements through their occupation is not feasible, because in 1930 proportion of women in the labor force was relatively small in the United States. A reasonable proxy variable describing social status of non-working adult women is an occupation of husband (for married women) or occupation of the head of household for single, widowed or divorced women. Urban/rural residence in 1930 is another variable used in the study. Preston and Haines found that child mortality in 1900 was significantly higher in urban than in rural areas (Preston and Haines, 1991). Urban adults in the contemporary United States also have higher mortality despite better infrastructure and access to health services (Hayward et al., 1997).

Familial longevity and other family characteristics. Family histories allow us to obtain information on lifespan of biological and non-biological relatives. For this particular study, the most important variables are lifespans of mother and father. As yet, no studies have simultaneously examined the net effects of parental longevity and early-life conditions. Studies suggest that effects of parental longevity on longevity of the offspring may be substantial (Gavrilov et al., 2002; Kerber et al., 2001; Pearl and Pearl, 1934) and heritability of lifespan estimates increase dramatically when parents live longer than 80 years (Gavrilova et al., 1998). Therefore, we believe that parental longevity (measured as paternal and maternal lifespan 80 years and over) may have significant moderating influence on the effects of childhood conditions and it can be used as a proxy for genetic influences on lifespan. Other family variables of interest are paternal and maternal ages at person's birth, sibship size and birth order.

In this ongoing study we have identified 838 centenarians born in 1890-91 in the United States and 439 shorter-lived controls born in the United States and died at age 65 years. Further linkage to 1900 census resulted in 98.2% success rate for centenarians and 98.6% success rate for controls. 94.9% of centenarian records and 96.4% of control records were successfully linked to the 1930 census. Linkage to the 1900 census revealed that 95.6% centenarians and 96.0% controls lived with one or both biological parents. 67% fathers of studied individuals were farmers according to 1900 census. Centenarians and controls had approximately equal sibship sizes (7.6 and 7.8 respectively), which are higher compared to the general population in 1900 census (5.6) suggesting larger sizes of families presented in computerized genealogies. In further analyses we restricted our sample with records where information was available for both 1900 and 1930 census. To study effects of marriage history on survival to age 100 years, only records for individuals being married in

1930 were taken into account. Finally, data for 765 centenarians and 423 shorter-lived controls were used in our analyses.

Multivariate logistic regression model was used to study survival to age 100. Our main focus was on the following three types of variables: (1) Early-life conditions drawn from the 1900 census (type of parental household - farm or non-farm, own or rented, parental literacy, parental immigration status, paternal occupation, number of children born/survived by mother, size of parental household in 1900, places of birth for household members), (2) midlife conditions drawn from the 1930 census (type of person's household, availability of radio in household, person's age at first marriage, person's occupation or husband's occupation in the case of women, industry of occupation, number of children in household, veteran status), (3) family characteristics drawn from computerized genealogies (paternal and maternal lifespan, paternal and maternal age at person's birth, number of siblings).

In the first step we studied familial, childhood and adulthood variables separately using univariate analyses. Study of familial characteristics taken from genealogies revealed that paternal and maternal longevity was significantly associated with survival to age 100 for both men and women. Being born in the second half of year was significantly associated with male longevity. However loss of parents early in life (before 1910) had no effect on the chances to become a centenarian. Childhood conditions recorded in the 1900 census included: paternal and maternal literacy and immigration status, paternal occupation, status of dwelling (owned or rented farm, owned or rented house), household size, grandparent or boarder in household, proportion of surviving children reported by mother, region of birth. Larger household size, birth in the North East or Midwest regions and having farmer father were found to be significant predictors of male (but not female) longevity in univariate analyses. Birth in the North-East region is also predictive for survival to advanced ages in men. This result agrees with findings by Hill and colleagues for persons survived to age 85 (Hill et al., 2000), but does not agree with the results of our earlier study, which compared centenarians drawn from computerized family histories with population-based controls (Gavrilova and Gavrilov, 2007). This contradictory finding may indicate that the earlier use of population-based control could produce biased results if the studied sample of genealogical records does not represent the general population. Female longevity revealed no significant associations with any of 1900 census variables. Adulthood conditions in the 1930 census included: dwelling status, occupation of self (husband or head of household for females), radio in household, veteran status of self (or husband), marital status, age at first marriage, availability of children (composite variable based on information taken from 1930 census and genealogies). Univariate analyses showed that farmer occupation in 1930 was a very strong predictor of longevity for men. In the case of women, having husband-farmer had no effect on the chances of survival to age 100. For women, availability of radio in household and higher age at first marriage were the strongest predictors of longevity among the studied midlife variables.

In multivariate analyses, when familial, early-life and midlife chanracteristics are combined the region of birth and having farmer father are no longer associated with longevity of men. Parental longevity turned out to be one of the strongest predictors of survival to age 100. Table 8 presents the results of multivariate analyses for men. Note that the farmer occupation in 1930 is one of the strongest predictors of survival to age 100, which agrees with results of other studies including our own study of centenarians based on population-based sample of survivors to age 100 from the 1887 birth cohort (Gavrilov and Gavrilova, 2012).

[Table 8 About Here]

Table 9 presents results of multivariate analyses for women. For women, having a husband-farmer has no effect on survival to age 100. Long-lived women also had higher age at first marriage (Table 9). Interestingly, having radio in household in 1930 has positive effect on longevity for women but not for men. This finding can be explained by the fact that women in 1930 spent most of their time at home and were much more exposed to radio (as an educational and entertainment source) compared to men. Listening to radio improves people's feelings of happiness and energy, and electro-encephalographic (EEG) study found that listening to radio creates high levels of positive engagement in the brain according to the recent findings of the “Media and the mood of the nation” research project conducted by Sparkler Research in Spring 2011 (Redican and Barber, 2012).

[Table 9 About Here]

Finally, we tested our previous results that season of birth may be predictive for survival to long life and compared season-of-birth among centenarians and shorter-lived controls in this database. Figure 3 shows proportion of persons born in the first and the second halfs of the calendar year for centenarians and controls. Note that more centenarians than controls were born in the second half of the year and this difference is statistically significant (p=0.032, chi-square test). This result confirms our results obtained using the within-family analysis (Gavrilov and Gavrilova, 2011), which showed that centenarians were born more often in September-November (see section 3).

Table 8 . Predictors of male survival to age 100: effects of parental longevity, early-life and midlife conditions. Results of multivariate logistic regression.

Variable	**Odds ratio**	**95% CI**	**p-value**
Father lived 80+	1.69	1.18-2.44	0.005
Mother lived 80+	1.99	1.37-2.87	<0.001
Farmer in 1930	2.03	1.38-2.97	<0.001
Age at first marriage	1.01	0.99-1.03	0.265
Radio in household	1.12	0.77-1.63	0.541
Born in the second half of year	1.47	1.02-2.10	0.037

N=550. Farm childhood in 1900 were found to be non-significant predictors for males. Calculated using Stata 11 statistical package (procedure logistic).

Table 9. Predictors of female survival to age 100: effects of parental longevity, early-life and midlife conditions. Results of multivariate logistic regression.

Variable	**Odds ratio**	**95% CI**	**p-value**
Father lived 80+	1.97	1.34-2.90	0.001
Mother lived 80+	2.37	1.63-3.44	<0.001
Husband (or head of household) farmer in 1930	1.05	0.72-1.54	0.789
Age at first marriage	1.03	1.01-1.05	0.004
Radio in household in 1930	1.61	1.09-2.38	0.017
Born in the second half of year	0.99	0.69-1.43	0.966

N=594. Calculated using Stata 11 statistical package (procedure logistic).

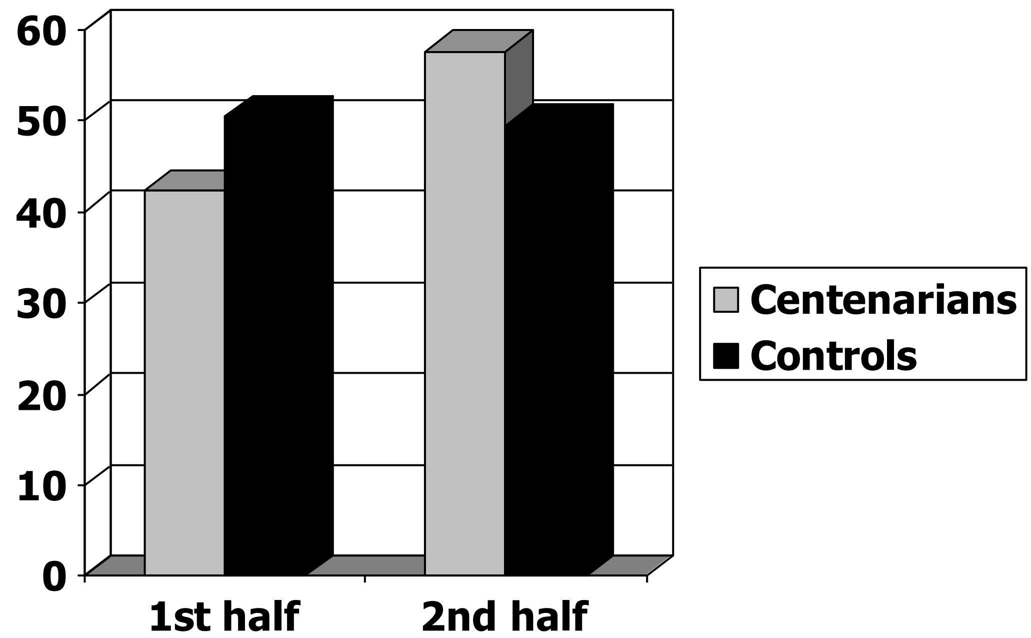
[Figure 3 About Here]

This study demonstrated that both midlife and early-life conditions affect survival to age 100 with some gender specificity. It is also important to note that parental longevity turned out to be one of the strongest predictors of survival to age 100. Thus, we may conclude that information about such an important predictor as parental longevity cannot be ignored and should be collected in contemporary longitudinal studies. At the same time, we found no effects of higher child mortality in household (a proxy of infectious burden) on longevity as suggested by inflammatory hypothesis of aging (Finch and Crimmins, 2004).

Conclusion

We considered in this article several approaches to centenarian studies. Some cases of exceptional longevity may represent particularly interesting outcomes of successful natural experiments on delaying human aging, and preventing age-related diseases. Therefore, studies on centenarians could become a gold-mine for unraveling the secrets of human longevity through careful epidemiological analysis of fortunate unintended natural experiments on life-extension and disease prevention. A comparison with the gold-mine is appropriate here not only in terms of high expected gains for possible dramatic extension of healthy human life, but also in terms of required effort: we had to scan over 300,000 online family histories and then to do a tedious work on data validation and cleaning, in order to obtain just a few hundred reliable records and some meaningful findings. Therefore we consider the results of this study as the beginning of a subsequent large-scale research effort with the promise of potentially breathtaking findings in the future.

Figure 3 . Season of birth and survival to 100. Proportion (percent) of persons born in the first half and the second half of the calendar year among centenarians and controls (who died at age 65).



Acknowledgments

This study was supported by the U.S. National Institute on Aging (grant R01 AG028620).

References

Abel, E.L., and Kruger, M.L., 2010. Birth Month Affects Longevity. Death Stud. 34, 757-763.

Adams, J.W., and Kasakoff, A.B., 1984. Migration and the Family in Colonial New-England - the View from Genealogies. J Fam Hist 9, 24-43.

Adkins, G., Martin, P., and Poon, L.W., 1996. Personality traits and states as predictors of subjective well-being in centenarians, octogenarians, and sexagenarians. Psychol Aging 11, 408-416.

Anderton, D.L., Tsuya, N.O., Bean, L.L., and Mineau, G.P., 1987. Intergenerational Transmission of Relative Fertility and Life Course Patterns. Demography 24, 467-480.

Barker, D.J.P., 1998. Mothers, Babies, and Disease in Later Life. Churchill Livingstone, London.

Bean, L.L., Mineau, G.P., and Anderton, D.L., 1992. High-Risk Childbearing - Fertility and Infant-Mortality on the American Frontier. Soc Sci Hist 16, 337-363.

Bengtsson, T., and Lindstrom, M., 2000. Childhood misery and disease in later life: The effects on mortality in old age of hazards experienced in early life, southern Sweden, 1760-1894. Pop Stud-J Demog 54, 263-277.

Bengtsson, T., and Lindstrom, M., 2003. Airborne infectious diseases during infancy and mortality in later life in southern Sweden, 1766-1894. Int J Epidemiol 32, 286-94.

Bickel, S.E., 2005. Aging (not so) gracefully.[comment]. Nature Genetics 37, 1303-4.

Bouwhuis, S., Charmantier, A., Verhulst, S., and Sheldon, B.C., 2010. Trans-generational effects on ageing in a wild bird population. J. Evol. Biol. 23, 636-642.

Breslow, N., 1970. Generalized Kruskal-Wallis Test for Comparing K Samples Subject to Unequal Patterns of Censorship. Biometrika 57, 579-&.

Breslow, N.E., and Day, N.E., 1993. Statistical Methods in Cancer Research. Vol.1. The Analysis of Case-Control Studies. Oxford University Press, Oxford.

Carnes, B.A., Riesch, R., and Schlupp, I., 2012. The Delayed Impact of Parental Age on Offspring Mortality in Mice. J Gerontol a-Biol 67, 351-357.

Caselli, G., Pozzi, L., Vaupel, J.W., Deiana, L., Pes, G., Carru, C., Franceschi, C., and Baggio, G., 2006. Family clustering in Sardinian longevity: A genealogical approach. Exp Gerontol 41, 727-736.

Comings, D.E., and Macmurray, J., 2001. Maternal age as a confounding variable in association studies. American Journal of Medical Genetics 105, 564-564.

Comings, D.E., and MacMurray, J.P., 2006. Maternal age at the birth of the first child as an epistatic factor in polygenic disorders. Am J Med Genet B 141B, 1-6.

Costa, D.L., and Lahey, J., 2005. Becoming oldest old: evidence from historical U.S. data. Genus 61, 125-161.

Doblhammer, G., 2004. The Late Life Legacy of Very Early Life. Demographic Research Monographs. Springer, Heidelberg.

Doblhammer, G., and Vaupel, J.W., 2001. Lifespan depends on month of birth. P Natl Acad Sci USA 98, 2934-2939.

Doblhammer, G., Scholz, R., and Maier, H., 2005. Month of birth and survival to age 105+: Evidence from the age validation study of German semi-supercentenarians. Exp Gerontol 40, 829-835.

Elo, I.T., and Preston, S.H., 1992. Effects of early-life condition on adult mortality: A review. Popul Index 58, 186-222.

Elo, I.T., Preston, S.H., Rosenwaike, I., Hill, M., and Cheney, T.P., 1996. Consistency of age reporting on death certificates and social security records among elderly African Americans. Soc Sci Res 25, 292-307.

Finch, C.E., and Crimmins, E.M., 2004. Inflammatory exposure and historical changes in human life-spans. Science 305, 1736-1739.

Fogel, R.W., 2004. Technophysio evolution and the measurement of economic growth. J Evol Econ 14, 217-221.

Fogel, R.W., and Costa, D.L., 1997. A theory of technophysio evolution, with some implications for forecasting population, health care costs, and pension costs. Demography 34, 49-66.

Gavrilov, L.A., and Gavrilova, N.S., 1991. The Biology of Life Span: A Quantitative Approach. Harwood Academic Publisher, New York.

Gavrilov, L.A., and Gavrilova, N.S., 1997. Parental age at conception and offspring longevity. Reviews in Clinical Gerontology 7, 5-12.

Gavrilov, L.A., and Gavrilova, N.S., 1999. Season of birth and human longevity. Journal of Anti-Aging Medicine 2, 365-366.

Gavrilov, L.A., and Gavrilova, N.S., Human longevity and parental age at conception, in: Robine, J.-M., et al., Eds.), Sex and Longevity: Sexuality, Gender, Reproduction, Parenthood, Springer-Verlag, Berlin, Heidelberg 2000, pp. 7-31.

Gavrilov, L.A., and Gavrilova, N.S., 2001a. Biodemographic study of familial determinants of human longevity. Population: An English Selection 13, 197-222.

Gavrilov, L.A., and Gavrilova, N.S., 2001b. The reliability theory of aging and longevity. J Theor Biol 213, 527-545.

Gavrilov, L.A., and Gavrilova, N.S., 2003a. The quest for a general theory of aging and longevity. Science of Aging Knowledge Environment 28, 16.

Gavrilov, L.A., and Gavrilova, N.S., Early-life factors modulating lifespan, in: Rattan, S. I. S., (Ed.), Modulating Aging and Longevity, Kluwer Academic Publishers, Dordrecht, The Netherlands 2003b, pp. 27-50.

Gavrilov, L.A., and Gavrilova, N.S., 2004. Early-life programming of aging and longevity - The idea of high initial damage load (the HIDL hypothesis). Ann Ny Acad Sci 1019, 496-501.

Gavrilov, L.A., and Gavrilova, N.S., Reliability Theory of Aging and Longevity, in: Masoro, E. J. and Austad, S. N., Eds.), Handbook of the Biology of Aging, Academic Press, San Diego 2006, pp. 3-42.

Gavrilov, L.A., and Gavrilova, N.S., Mortality measurement at advanced ages: a study of the Social Security Administration Death Master File, Living to 100 and Beyond: Survival at Advanced Ages [online monograph]. SOA Monograph M–LI08–1. 32p., The Society of Actuaries, Shaumburg, IL 2008.

Gavrilov, L.A., and Gavrilova, N.S., 2011. Season of birth and exceptional longevity: comparative study of american centenarians, their siblings, and spouses. Journal of aging research 2011, 104616.

Gavrilov, L.A., and Gavrilova, N.S., 2012. Biodemography of Exceptional Longevity: Early-Life and Mid-Life Predictors of Human Longevity. Biodemography and Social Biology 58, 14-39.

Gavrilov, L.A., Gavrilova, N.S., Olshansky, S.J., and Carnes, B.A., 2002. Genealogical data and the biodemography of human longevity. Soc Biol 49, 160-173.

Gavrilova, N.S., and Gavrilov, L.A., 1999. Data resources for biodemographic studies on familial clustering of human longevity. Demographic Research 1, 1-48.

Gavrilova, N.S., and Gavrilov, L.A., 2007. Search for Predictors of Exceptional Human Longevity: Using Computerized Genealogies and Internet Resources for Human Longevity Studies. North American Actuarial Journal 11, 49-67.

Gavrilova, N.S., Gavrilov, L.A., Evdokushkina, G.N., Semyonova, V.G., Gavrilova, A.L., Evdokushkina, N.N., Kushnareva, Y.E., Kroutko, V.N., and Andreyev, A.Y., 1998. Evolution, mutations, and human longevity: European royal and noble families. Hum Biol 70, 799-804.

Gloria-Bottini, F., Cosmi, E., Nicotra, M., Cosmi, E.V., and Bottini, E., 2005. Is delayed childbearing changing gene frequencies in western populations? Hum Biol 77, 433-441.

Hadley, E.C., Rossi, W.K., Albert, S., Bailey-Wilson, J., Baron, J., Cawthon, R., Christian, J.C., Corder, E.H., Franceschi, C., Kestenbaum, B., Kruglyak, L., Lauderdale, D.S., Lubitz, J., Martin, G.M., McClearn, G.E., McGue, M., Miles, T., Mineau, G., Ouellett, G., Pedersen, N.L., Preston, S.H., Page, W.F., Province, M., Schachter, F., Schork, N.J., Vaupel, J.W., Vijg, J., Wallace, R., Wang, E., Wijsman, E.M., and Wor, N.A.G.E., 2000. Genetic epidemiologic studies on age-specified traits. Am J Epidemiol 152, 1003-1008.

Hagberg, B., Alfredson, B.B., Poon, L.W., and Homma, A., 2001. Cognitive functioning in centenarians: A coordinated analysis of results from three countries. J Gerontol B-Psychol 56, P141-P151.

Hayward, M.D., and Gorman, B.K., 2004. The long arm of childhood: The influence of early-life social conditions on men's mortality. Demography 41, 87-107.

Hayward, M.D., Pienta, A.M., and McLaughlin, D.K., 1997. Inequality in men's mortality: The socioeconomic status gradient and geographic context. J Health Soc Behav 38, 313-330.

Heffner, L.J., 2004. Advanced maternal age--how old is too old? New Engl J Med 351, 1927-9.

Hill, M.E., Preston, S.H., Rosenwaike, I., and Dunagan, J.F., Childhood conditions predicting survival to advanced age among white Americans., 2000 Annual meeting of the Population Association of America, Los Angeles 2000.

Hosmer, D.W., and Lemeshow, S., 2001. Applied Logistic Regression. Wiley & Sons, New York.

Jeune, B., and Vaupel, J., 1999. Validation of Exceptional Longevity. Odense University Publisher, Odense.

Kalmijn, M., and Kraaykamp, G., 2005. Late or later? A sibling analysis of the effect of maternal age on children's schooling. Soc Sci Res 34, 634-650.

Keefe, D.L., Franco, S., Liu, L., Trimarchi, J., Cao, B., Weitzen, S., Agarwal, S., and Blasco, M.A., 2005. Telomere length predicts embryo fragmentation after in vitro fertilization in women - Toward a telomere theory of reproductive aging in women. American Journal of Obstetrics and Gynecology 192, 1256-1260.

Kemkes-Grottenthaler, A., 2004. Parental effects on offspring longevity - evidence from 17th to 19th century reproductive histories. Ann Hum Biol 31, 139-158.

Kerber, R.A., O'Brien, E., Smith, K.R., and Cawthon, R.M., 2001. Familial excess longevity in Utah genealogies. Journals of Gerontology Series A-Biological Sciences & Medical Sciences 56, B130-9.

Kuh, D., and Ben-Shlomo, B., 1997. A Life Course Approach to Chronic Disease Epidemiology. Oxford University Press, Oxford.

Lerchl, A., 2004. Month of birth and life expectancy: role of gender and age in a comparative approach. Naturwissenschaften 91, 422-425.

Margrett, J., Martin, P., Woodard, J.L., Miller, L.S., MacDonald, M., Baenziger, J., Siegler, I.C., Davey, A., Poon, L., and Study, G.C., 2010. Depression among Centenarians and the Oldest Old: Contributions of Cognition and Personalityu. Gerontology 56, 93-99.

Martin, P., Cho, J., MacDonald, M., and Poon, L., 2010. Personality, Functional Capacity, and Well-Being among Centenarians. Gerontologist 50, 50-50.

Mazan, R., and Gagnon, A., 2007. Familial and environmental influences on longevity in historical Quebec. Population 62, 315-338.

Montesanto, A., Latorre, V., Giordano, M., Martino, C., Domma, F., and Passarino, G., 2011. The genetic component of human longevity: analysis of the survival advantage of parents and siblings of Italian nonagenarians. European Journal of Human Genetics 19, 882-886.

Murabito, J.M., Yuan, R., and Lunetta, K.L., 2012. The Search for Longevity and Healthy Aging Genes: Insights From Epidemiological Studies and Samples of Long-Lived Individuals. J Gerontol a-Biol 67, 470-479.

Pearl, R., and Pearl, R.D.W., 1934. The Ancestry of the Long-Lived. The John Hopkins Press, Baltimore.

Pellestor, F., Anahory, T., and Hamamah, S., 2005. Effect of maternal age on the frequency of cytogenetic abnormalities in human oocytes. Cytogenetic & Genome Research 111, 206-12.

Perls, T., and Terry, D., 2003. Genetics of exceptional longevity. Exp Gerontol 38, 725-730.

Perls, T., Kohler, I.V., Andersen, S., Schoenhofen, E., Pennington, J., Young, R., Terry, D., and Elo, I.T., 2007. Survival of parents and siblings of supercentenarians. J Gerontol a-Biol 62, 1028-1034.

Perls, T.T., Wilmoth, J., Levenson, R., Drinkwater, M., Cohen, M., Bogan, H., Joyce, E., Brewster, S., Kunkel, L., and Puca, A., 2002. Life-long sustained mortality advantage of siblings of centenarians. P Natl Acad Sci USA 99, 8442-8447.

Polani, P.E., and Crolla, J.A., 1991. A Test of the Production Line Hypothesis of Mammalian Oogenesis. Hum Genet 88, 64-70.

Poulain, M., 2010. On the age validation of supercentenarians. Supercentenarians, 3-30.

Poulain, M., 2011. Exceptional longevity in Okinawa: A plea for in-depth validation. Demographic Research 25, 245-284.

Preston, S.H., and Haines, M.R., 1991. Fatal years. Child mortality in late nineteenth-century America. Princeton University Press, Princeton, NJ.

Preston, S.H., Hill, M.E., and Drevenstedt, G.L., 1998. Childhood conditions that predict survival to advanced ages among African-Americans. Social Science & Medicine 47, 1231-1246.

Redican, S., and Barber, M., Radio: The Emotional Multiplier., London 2012.

Rosenwaike, I., and Stone, L.F., 2003. Verification of the ages of supercentenarians in the United States: Results of a matching study. Demography 40, 727-739.

Ruggles, S., Alexander, J.T., Genadek, K., Goeken, R., Shroeder, M.B., and Sobek, M., 2010. Integrated Public Use Microdata Series (IPUMS): Version 5.0 [Machine-readable database]. University of Minnesota, Minneapolis, MN.

Ruggles, S., Sobek, M., Alexander, T., Fitch, C.A., Goeken, R., Hall, P.K., King, M., and Ronnander, C., Integrated Public Use Microdata Series (IPUMS): Version 3.0., Minnesota Population Center, Minneapolis, MN 2004.

Schoenmaker, M., de Craen, A.J.M., de Meijer, P., Beekman, M., Blauw, G.J., Slagboom, P.E., and Westendorp, R.G.J., 2006. Evidence of genetic enrichment for exceptional survival using a family approach: the Leiden Longevity Study. European Journal of Human Genetics 14, 79-84.

Sebastiani, P., Solovieff, N., DeWan, A.T., Walsh, K.M., Puca, A., Hartley, S.W., Melista, E., Andersen, S., Dworkis, D.A., Wilk, J.B., Myers, R.H., Steinberg, M.H., Montano, M., Baldwin, C.T., Hoh, J., and Perls, T.T., 2012. Genetic Signatures of Exceptional Longevity in Humans. Plos One 7.

Sesso, H.D., Paffenbarger, R.S., and Lee, I.M., 2000. Comparison of National Death Index and World Wide Web death searches. Am J Epidemiol 152, 107-111.

Shrestha, L.B., and Rosenwaike, I., 1996. Can data from the decennial census measure trends in mobility limitation among the aged? Gerontologist 36, 106-109.

Sklar, D., and Trachtenberg, A., 2002. PHP Cookbook. O'Reilly.

Smith, K.R., Mineau, G.R., Garibotti, G., and Kerber, R., 2009a. Effects of childhood and middle-adulthood family conditions on later-life mortality: Evidence from the Utah Population Database, 1850-2002. Social Science & Medicine 68, 1649-1658.

Smith, K.R., Gagnon, A., Cawthon, R.M., Mineau, G.P., Mazan, R., and Desjardins, B., 2009b. Familial Aggregation of Survival and Late Female Reproduction. J Gerontol a-Biol 64, 740-744.

StataCorp, 2009. Stata User's Guide. Stata Press.

Tarin, J.J., Gomez-Piquer, V., Rausell, F., Navarro, S., Hermenegildo, C., and Cano, A., 2005. Delayed motherhood decreases life expectancy of mouse offspring. Biology of Reproduction 72, 1336-43.

Vaiserman, A.M., Khalangot, M.D., Carstensen, B., Tronko, M.D., Kravchenko, V.I., Voitenko, V.P., Mechova, L.V., Koshel, N.M., and Grigoriev, P.E., 2009. Seasonality of birth in adult type 2 diabetic patients in three Ukrainian regions. Diabetologia 52, 2665-2667.

Wang, M.H., and vom Saal, F.S., 2000. Maternal age and traits in offspring - The timing of a mouse's first litter influences the development of her pups. Nature 407, 469-470.

Westendorp, R.G.J., van Heemst, D., Rozing, M.P., Frolich, M., Mooijaart, S.P., Blauw, G.J., Beekman, M., Heijmans, B.T., de Craen, A.J.M., Slagboom, P.E., and Grp, L.L.S., 2009. Nonagenarian Siblings and Their Offspring Display Lower Risk of Mortality and Morbidity than Sporadic Nonagenarians: The Leiden Longevity Study. J Am Geriatr Soc 57, 1634-1637.

Willcox, B.J., Willcox, D.C., He, Q.M., Curb, J.D., and Suzuki, M., 2006. Siblings of Okinawan centenarians share lifelong mortality advantages. J Gerontol a-Biol 61, 345-354.

Woodward, M., 2005. Epidemiology. Study Design and Data Analysis. Chapman & Hall/CRC, Boca Raton, FL.

Zeng, Y., Cheng, L.G., Chen, H.S.A., Cao, H.Q., Hauser, E.R., Liu, Y.Z., Xiao, Z.Y., Tan, Q.H., Tian, X.L., and Vaupel, J.W., 2010. Effects of FOXO Genotypes on Longevity: A Biodemographic Analysis. J Gerontol a-Biol 65, 1285-1299.