ANNALES DE DÉMOGRAPHIE HISTORIQUE 2003 n° 2 p. 177 à 198

EARLY-LIFE PREDICTORS OF HUMAN LONGEVITY: ANALYSIS OF THE XIXth CENTURY BIRTH COHORTS

by Natalia S. GAVRILOVA, Leonid A. GAVRILOV,

Galina N. EVDOKUSHKINA, Victoria G. SEMYONOVA

The idea of fetal origins of adult degenerative diseases and early-life programming of late-life health and survival is being actively discussed in the scientific literature (Lucas, 1991; Barker, 1992; 1998; Kuh & Ben-Shlomo, 1997; Leon *et al.*, 1998; Lucas *et al.*, 1999; Blackwell *et al.*, 2001). The historical improvement in early-life conditions may be responsible for the observed significant increase in human longevity through the process called 'technophysio evolution' (Fogel & Costa, 1997; Fogel, 1997; 1999). Additional arguments suggesting the importance of early-life conditions in later-life health outcomes are coming from the reliability theory of aging and longevity (Gavrilova, Gavrilova, 1991; 2001a). According to this theory, biological species (including humans) are starting their lives with extremely high initial load of damage, and, therefore, they should be sensitive to early-life conditions affecting the level of initial damage (Gavrilova, Gavrilova, 1991; 2001a). All these ideas require further testing, more studies, and more data.

There are two major goals in this exploratory study: (1) To find out whether early-life conditions may have significant effects on adult lifespan. We also tried to determine whether our data set on European aristocratic families could be useful to explore the role of early-life conditions, and to be used in future more detailed studies. (2) To determine whether early-life conditions may have significant effect on sex disparities in adult health and longevity. These sex disparities are well documented (Van Poppel, 2000), but they still have to be explained and fully understood. For example, the following research question could be posed: are the long-lasting effects of early-life conditions identical for both sexes, or, on the contrary, they are sex-specific? This question stimulated us to conduct this study on the sex specificity of the effects of early-life conditions on adult lifespan.

Acknowledgments

The results of this study were presented at the 2001 Annual meeting of the Social Science and History Association and we would like to thank the participants of this meeting as well as three anonymous reviewers for valuable comments and suggestions.

We are most grateful to Mr. Brian Whiteley for useful editorial suggestions. We would also like to acknowledge partial support from the National Institute on Aging grants.

The study of sex differences has also important methodological implications for research work on early-life effects. This is because in many cases the available data sets are limited in their size and/or the studied outcomes are rare events, thus creating a temptation to pool the data for both sexes together in order to increase the statistical power (Blackwell *et al.*, 2001). It is important, therefore, to find out whether gender differences in response to early-life conditions are indeed similar (so that the data could be pooled together with simple adjustment for sex just by one indicator variable), or they are fundamentally and qualitatively different (so that each sex should be studied separately).

In this study we addressed these scientific problems (fundamental and methodological) by studying the effects of early-life conditions on adult lifespan of men and women separately, using the methodology of historical prospective study of extinct birth cohorts. We found significant sex differences in adult lifespan responses to early-life conditions that justify the need for further full-scale research project on related topic.

Data and Methods

Main Data Source

In this study we collected, computerized, cross-checked and analyzed the detailed genealogical records on lifespan of over 16,000 adult persons (8,052 men and 7,979 women survived by age 30) and their parents, using particularly reliable and complete data on European royal and noble families for extinct birth cohorts (born 1800-1880). The main advantage of these data is their high accuracy, reliability and completeness (to be discussed later). Another advantage of this kind of data is the relative homogeneity of this Caucasian population regarding social class and educational background. Since this privileged social group lived in favourable conditions for many centuries, one could expect less influence of adverse social factors (poverty, for example) on life span and hence lower bias caused by these factors. This kind of data allows us to minimize the social heterogeneity of the population under study. Thus, although the sample analyzed in this study does not represent the whole human population (as laboratory animals do not represent species in the wild), it is one of the best possible samples to test biodemographic hypotheses since the effects of population heterogeneity are minimized with regard to social status.

The database on European royal and noble families (a family-linked database) was developed as a result of 6 years of our continued efforts because of extensive data cross-checking and data quality control. The earlier intermediate versions of this database were already used in our previous studies (Gavrilov, Gavrilova, 1997a; 1997b; 1999a; 2000a; 2001b; Gavrilov *et al.*, 1997; Gavrilova, Gavrilov, 2001; Gavrilova *et al.*, 1998). To develop this database we have chosen one of the best professional sources of genealogical data available—the famous German edition of the "Genealogisches Handbuch des Adels" (Genealogical Yearbook of Nobility). This edition is known world wide as the "Gotha Almanac" - "Old Gotha" published in Gotha in 1763-1944, and "New Gotha" published in Marburg since 1951 (see Gavrilova, Gavrilov,

1999, for more details). Data from the Gotha Almanac were often used in early biodemographic studies of fertility (see Hollingsworth, 1969, 199-224, for references) and proved to be useful now in the studies of human longevity (Gavrilov, Gavrilova, 1997a; 1997b; 1999a; 2000a; 2001b; Gavrilov *et al.*, 1997; Gavrilova, Gavrilov, 2001; Gavrilova *et al.*, 1998).

Each volume of the New Gotha Almanac contains about 2,000 genealogical records dating back to the XIVth-XVIth centuries (to the founder of a particular noble genus). More than 100 volumes of this edition are already published, so more than 200,000 genealogical records with well-documented genealogical data are available from this data source. The high quality of information published in this edition is ensured by the fact that the primary information is drawn from the German Noble Archive (Deutsches Adelsarchiv). The Director of the German Noble Archive (Archivdirektor) is also the Editor of the New Gotha Almanach. Our own experience based on cross-checking the data, has demonstrated that the number of mistakes (mostly misprints) is very low in the "New Gotha Almanac" (less than 1 per 1000 records), so this source of data is very accurate compared to other published genealogies.

The information on noble families in the New Gotha Almanac is recorded in a regular manner. The description of each particular noble genus starts with information on 2-3 generations of founders of male sex only. Then three to four the most recent generations are described in more detail, including information on individuals (e.g., first and last names; event data: birth, death, marriage dates and places; descriptive data: noble degrees, occupation if available, information on death circumstances if available), information on parents (e.g., first and last names; event data: birth and death dates and places), information on spouses (e.g., first and last names; birth and death dates and places; first and last names of parents) and information on children (detailed as for each individual).

The process of data computerization is not yet completed—instead this is an ongoing project because of tremendous amounts of published data available for further computerizing. The present study represents, therefore, our intermediate findings.

Supplementary Data Sources

Some other supplementary sources of data were used in the development of database. These data sources included: (1) computerized database on European royalty named "Royal92" and distributed on the Internet by Brian C. Tompsett at University of Hull, UK; (2) computerized database on British Peerage distributed on CD by S&N Genealogy Supplies; (3) relevant computerized data for European aristocratic families available in World Family Tree Archive CDs (Gavrilova, Gavrilov, 1999); (4) over 100 genealogical publications on Russian nobility listed elsewhere (Gavrilov *et al.*, 1996).

These data were used as a supplement to the main data source since their quality was not as high compared to the Gotha Almanac. Although data on European royalty were recorded in computerized data sources ("Royal92", British Peerage CD, see above) with sufficient completeness, data on lower rank nobility (landed gentry) were less complete and accurate. The same was true for the data on Russian nobility. All

supplementary data were matched with the Gotha Almanac data, in order to cross-check the overlapping pieces of information. This cross-checking procedure allowed us to increase the completeness of the database by complementation of information taken from different sources.

The Structure of the Genealogical Database

The database approach used in this study is similar to the approach used for existing family-linked databases, such as the Utah Population Database (Skolnick *et al.*, 1979), Laredo Epidemiological Project (Buchanan *et al.*, 1984) or other historical databases (Gutmann *et al.*, 1989).

Each record in the database represents an individual's event data (birth and death dates and places) and individual's descriptive information, that is, identification number, sex, first and last names, nobility rank, occupation, birth order, cause of death (violent/nonviolent), ethnicity, marital status, data source code number, data source year of publication. Individual information is supplemented by data for parents (identification numbers, first and last names, birth, death and marriage dates, cause of death) and spouses. Thus, the database that is used in this project is organized in the form of triplets (referred to as the "ego" and two parents). This structure of records is widely used in human genetics and is adequate for studies of parent-child relationships. Similar database structure was used in the recent study of kinship networks (Post *et al.*, 1997).

Data Quality Control

Data quality control was an important part of our study designed to develop high quality family-linked database and to use it for scientific research.

The genealogical data sources were checked for the following: (1) *completeness*, in reporting birth and death dates, which is crucial for calculating individual life span, the variable of particular interest in our study; (2) *accuracy*, whether the percentage of mistakes and inconsistencies between reported dates (such as, for example, birth by the dead mother) is low enough to be acceptable; (3) *representativeness*, whether the characteristics of investigated data sets (distribution by age, sex, marital status, age at death, etc.) is close to demographic characteristics of populations in similar geographic areas, historical periods and social groups. In our study we referred to the well-known publication by Thomas Hollingsworth (1962) on British peerage as a standard for European aristocracy, to check for data representativeness.

The *completeness* in birth and death dates reporting in the New Gotha Almanac was very high: dates of all vital events were reported for nearly 95% of all persons. Such high completeness is not common for many other genealogical data sources. For example, for British Peerage data published by Burke almanac in most cases there are no birth dates for women, which makes calculation of their life span impossible. In fact, this problem (with British aristocratic women) was first noticed a century ago by Karl Pearson (Beeton and Pearson, 1899, 1901). He used the British Peerage data to study the longevity inheritance and had to exclude women from his consideration for the following

reason: "The limitation to the male line was enforced upon us partly by the practice of tracing pedigrees only through the male line, partly by the habitual reticence as to the age of women, even at death, observed by the compilers of peerages and family histories." (Beeton and Pearson, 1901, 50-51)

The *accuracy* of data published in the New Gotha Almanac is also very high: the frequency of inconsistent records is less than 1 per 1000 records while for many other genealogical data sources it falls within 1 per 300-400 records.

As for *representativeness*, the comparison of our data with Hollingsworth's analysis of British peerage (Hollingsworth, 1962) revealed good agreement between his findings and our data on mortality patterns, including male/female gap in life expectancy (up to 7-10 years of female advantage in lifespan).

The genealogies for the members of European aristocratic families presented in the "Gotha Almanac" are of descending type, tracing almost all the descendants of relatively few founders. This is an important advantage of this data source over other genealogies that are often of ascending type (pedigrees). It is known in historical demography that the ascending genealogies are biased, over-representing more fertile and longer-lived persons who succeed to become ancestors, and for this reason such genealogies should be treated with particular caution (Jetté and Charbonneau, 1984; Fogel, 1993).

Another important advantage of this data set is that the data are not spoiled by selective emigration (common problem for local registers), because every person is traced until his/her death in this data set.

Thus, the genealogical data published in the Gotha Almanac are characterized by high quality and accuracy. We have, however, encountered two problems regarding the data completeness, which are discussed below, along with proposed solutions.

Censored, truncated observations and missing death dates

Our study revealed that the percentage of cases with unreported death dates is rather small in our main data sources (Gotha Almanac), and is caused mainly by right censoring of long-lived persons who were still alive by the date of data collection and publication. The percentage of non-reported death dates varies from 0 to 7% in extinct birth cohorts (1800-1880), while it is higher in later birth cohorts (1880-1899)—23% for women and 8% for men, since some individuals were still alive by the date of data collection and volume publication. Note that women, who live longer, have a higher proportion of right-censored observations. The high proportion of censored observations in genealogies is not desirable, since the exact dates of censoring are often unknown. This uncertainty creates problems for data analysis, so the researchers working with genealogies prefer to use non-censored, extinct birth cohorts in their studies (Mayer, 1991; Pope, 1992; Kasakoff and Adams, 1995). We also used extinct (non-censored) birth cohorts in our study. For this purpose only those birth cohorts were used in the study that were born at least 100 years before the year of data publication (to be sure that the birth cohort under study is almost extinct). In addition to that, we focused our study on old cohorts born in 1800-1854 to exclude censored data from the analysis whenever possible.

Underreporting of women and children

In many genealogical books and databases non-married women as well as children died in infancy are often missed or reported with less completeness. Since genealogical records are focused on family names which are transmitted by males only, women could be lost in genealogies when they marry and change their family names (Hollingsworth, 1976). Also, in many cases data for women do not contain information on their birth and death dates resulting in biased sex ratio in the sample with complete dates. We have also encountered this problem in our studies although for somewhat different reason. Our analysis revealed that the main cause of the sex bias in the New Gotha Almanac is related to the manner of data representation: more recent generations are presented completely, while the earlier generations are limited mainly to the male ancestors (in order to avoid repetitive publication of individuals already presented in previous volumes). That is why, the sex ratio among early birth cohorts (1800-1860) is biased in favor of males while for more recent birth cohorts (1880-1899) it is within normal range. Since in our previous studies the most recent volumes of the New Gotha Almanac (published after 1980) were computerized and analyzed (in order to avoid censoring), the proportion of males in extinct birth cohorts (early generations) was substantially higher than expected (Gavrilov, Gavrilova, 1997a; b; 1999a; 2000a; 2001a; Gavrilov *et al.*, 1997; Gavrilova *et al.*, 1998)). Sex bias is an important issue, particularly when gender differences are studied (as it is done in the present study). Therefore, every effort is made to ensure that the data set used in this particular study is sex-balanced (Table 1).

Data presented in Table 1 demonstrate no significant sex bias in our data set.

The underreporting of children who died in infancy may be also a serious problem, especially for studies that include fertility analysis. Fortunately, in the Gotha Almanac the families that belong to the higher nobility rank (kings, princes, earls) are described with remarkable completeness. In particular, all ever born children are recorded, including those who died the same day when they were born. Another indicator of data completeness is the normal sex ratio at birth (101 to 108) observed among these families (based on analysis of our sample). In our database, over 90 aristocratic genuses belonged to the upper nobility were recorded completely, although data for lower rank nobility were not yet completed. Underreporting of children is not a problem for this particular study that is focused on adult life span for those who survived by age 30 years.

Analytical Methods

Since the data collected for this study are characterized by remarkable accuracy and completeness, it was possible to apply simple and straightforward methods of data analysis without making heavy assumptions. In particular, since the length of life is known for every person (there were no right censored observations) it was possible to analyze the duration of life directly as a dependent, outcome variable in multivariate regression model. There was no need to apply the Cox proportional hazard model and to make a heavy assumption about multiplicative effects of co-variates on

hazard rate. Instead, the persons' lifespan is studied directly as a dependent outcome variable and a function of other explanatory and potentially confounding predictor variables (see below).

In this study we applied a multivariate regression analysis with nominal variables, which is a very flexible tool to control for effects of both quantitative and qualitative (categorized) variables. This method also allows researchers to accommodate for complex non-linear and non-monotonic effects of predictor variables. The beauty of this method is that it does not require any assumptions about the analytical function describing the effects of predictor variables. Instead, the model allows researchers to calculate directly a conditional mean lifespan in a group of individuals with a particular set of predictor variables values. The regression coefficients obtained in this model (named as differential intercept coefficients) have a clear interpretation of additional years of life gained (or lost) due to change in a particular predictor variable.

We applied the methodology of prospective historical study to the data for extinct birth cohorts, free of censored observations. We focused our study on old cohorts born in 1800-1854, for which secular changes in life expectancy were minimal. We also tested a long list of explanatory and potentially confounding variables (described below) to avoid possible artifacts.

Life span of adult (30 +) progeny (sons and daughters separately) was considered as a dependent outcome variable in multivariate regression with dummy (0-1) variables using the SAS statistical package (procedure REG).

In this study we used the following multivariate model:

$$y_i = \beta_0 + \beta_1 x_i + r_i$$

The independent predictor variables included 14 types of binary variables:

(1) calendar year of birth (to control for historical increase in life expectancy as well as for complex secular fluctuations in lifespan). The whole birth year period of 1800-1854 was split into 5-year intervals (11 intervals) presented by 11 binary (0-1) variables with reference level set at 1850-1854 birth years. For data set of more recent birth cohorts (born in 1855-1880) five 5-year intervals were used with reference level at 1875-1880 birth cohort.

(2) maternal lifespan (to study maternal lifespan effects through combined genetic effects and shared environment). The maternal lifespan data were grouped into 5-year intervals (14 intervals) with the exception of the first (15-29 years) and the last (90 + years) longer intervals with small number of observations. The data were coded with 13 dummy variables with reference level set at 55-59 years for maternal lifespan.

(3) paternal lifespan (to study paternal lifespan effects through combined genetic effects and shared environment). The data were grouped and coded in a way similar to maternal lifespan (see above).

(4) maternal age when a person (proband) was born. This variable is used to control for possible confounding effects of maternal age on offspring lifespan. The data for mother's age were grouped in 5-year intervals (6 intervals to cover the age range of 15-60 years) with the exception of the last longer interval of 40 + years with small number of observations. Maternal age of 25-29 years is selected as a reference category.

(5) father's age when a person was born. This explanatory variable is used to study paternal age effects on offspring lifespan. The data were grouped and coded in 5-year intervals (9 intervals to cover the age range of 15-80 years) with the exception of the first (15-24 years) and the last (60-79 years) longer intervals with small number of observations.

(6) birth order. This variable is represented by 8 binary variables with the first birth order initially taken as the reference level.

(7) nationality. The nationality of individual is represented by a set of 6 categories—Germans, British, Italians, Poles, Russians and 'others'. Germans (the largest group in our sample) is selected as a reference group.

(8) cause of death ('extrinsic' versus 'natural'). The death is coded as extrinsic or premature in the following cases: (1) violent cause of death (war losses, accidents, etc.), (2) death in prison and other unfavorable conditions (concentration camp, etc.), (3) death from acute infections (cholera, etc.) and (4) maternal death (for women only). Deaths from all other causes combined were considered as a reference outcome. The proportion of reported 'extrinsic' deaths in our data set was about 5% for males and about 1% for females.

(9) loss of the father in the formative years of life (before age 10). This is a binary variable coded as 1 when father was lost before the age 10 and coded as zero otherwise.

(10) loss of the mother before age 10. This binary variable is coded as 1 in those cases when mother was lost before the age 10 and coded as zero otherwise.

(11) month of birth. This variable was included into analysis, because previous studies indicated that month of birth may be an important predictor of adult lifespan (Gavrilov, Gavrilova, 1999a; Doblhammer, Vaupel, 2001), particularly for daughters (Gavrilov, Gavrilova, 1999a). This variable was represented as a set of 11 dummy variables with those born in February considered as a reference group. The main focus of this particular study is on sex-differences in the month-of-birth effects that were not well studied before.

(12) death of sibling before age 18. This variable was included into analysis, because previous studies demonstrated that death of siblings during childhood may be an important predictor of late-life mortality (Alter *et al.*, 2001). This binary variable is coded as 1 in those cases when at least one sibling died before age 18 and coded as zero otherwise.

(13) nobility rank of the father. Families are categorized in groups of ruling royal families, non-ruling princes, dukes and counts (coded as 1) and other nobility (coded as 0).

(14) family size. Families are categorized in groups according to sib ship sizes: 1-2, 3-4, 5-6, 7-8, 9-10, and 11 + . Later some categories were collapsed together during an iterative fitting procedure.

Sensitivity analysis

In order to determine how robust are our findings, the sensitivity analysis was made. Specifically, the data were re-analyzed in several different ways, when either the initial data set was partially changed, or the set of predictor variables was modified. Changes in the data set included separate analyses for German and Russian nobility, or separate analyses

for old (born in 1800-1854) and more recent birth cohorts (born in 1860-1880). Changes in predictor variables included consideration of such additional variables as nobility rank, sib ship size, and reproductive lifespans (ages at last childbirth) both for mother and father.

Results and Discussion

The characteristics of analyzed data set are presented in Table 1. There are several notable features to mention here:

First, the numbers of males and females are rather similar in all studied birth cohorts (no apparent sex bias). The sex ratio in the entire data set is 1.01 (8,052 males/7,979 females). In the data set of cohorts born in 1800-1854 the sex ratio is 3,994/3,879 = 1,03. This is close to the normal sex ratio, in contrast to the sex ratio of 1.42, observed in the British peerage database with many missing records for women (Gavrilov, Gavrilova, 1999b). Thus, it seems to be possible to study sex differences in response to early-life conditions without particular concerns about selective sex bias in this data set.

Second, the values for mean lifespan are rather high—more than 63 years for males and 65 years for females. It indicates that lifespan in this socially elite population is to some extent comparable with modern lifespan values observed now in some countries of the world.

Third, there is a significant increase in lifespan over studied historical period, in particular for females (10 years gain). Therefore, the data should be adjusted for secular trends in lifespan (which has been done in this study) and/or analyzed for more narrow time periods (which was also done in this study). Finally, the temporal changes in lifespan are clearly not linear (no significant improvement in lifespan during the first 30 years of the XIXth century birth cohorts), and sometimes even not monotonic which justifies the method of analysis used in this study (multivariate regression with nominal variables and treating the year of birth as categorized predictor variable).

Season of birth and human longevity

Table 2 presents data suggesting that the month of birth is an important predictor for the life expectancy of adult women (30 years and above) born in 1800-1854. The dependence has M-shaped form with two peaks (bimodal distribution). In particular, women born in July and December tend to live 2-3 years longer on average compared to those born in February (significant at $p < 0.05$). The effects of months of birth are expressed in Table 2 as a difference from the reference level in February and are point estimates of the differential intercept coefficients adjusted for effects of other variables. It is interesting that life expectancy of women from more recent birth cohorts demonstrates similar M-shaped pattern with May and December as "good" months to be born (Table 3). It is important to emphasize that the month of birth continues to be an important predictor for women's lifespan, even after adjustment for the effects of all other explanatory variables mentioned earlier in the "Data and Methods" section.

It is interesting to note that the months of February and August are "bad" months to be born, and this finding corresponds well with some previous published data. For example, a similar bimodal month-of-

birth distribution was found for birth frequencies of cystic fibrosis disease with peak births in February and August (Brackenridge, 1980). Further studies are required to find out whether this is just a coincidence of findings or a general seasonal pattern.

Seasonal effects on human survival in historical population were demonstrated for infant mortality in Sardinia (Breschi, Livi Bacci, 1986). However, the fact that such an early circumstance of human life as the month of birth may have a significant effect 30 years later on the chances of human survival at adult ages is particularly intriguing. It indicates that there may be critical periods early in human life particularly sensitive to seasonal variation in living conditions in the past (e.g., vitamin supply, seasonal exposure to infectious diseases, etc.) with long-lasting effects in later life.

It is known that the deficiency of vitamins B_{12}, folic acid, B_6, niacin, C, or E, appears to mimic radiation in damaging DNA by causing single- and double-strand breaks, oxidative lesions, or both, and may contribute to premature aging (Ames, 1998). The seasonal lack of these vitamins in late winter/early spring, in coincidence with one of the two critical periods in fetus or child development (the third critical month of pregnancy and the first months after birth), may explain a dramatic life span shortening among those born in August and February. This finding is also consistent with the reliability theory of aging, which emphasizes the importance of the initial level of damage that determines the future length of human life (Gavrilov, Gavrilova, 1991; 2001b).

These general explanations, however, do not match well with data for males presented in Table 4. In contrast to females, the male lifespan does not depend on month of birth, at least in this particular data set. Males born in 1855-1880 also do not demonstrate any relation between their lifespan and month of birth (Table 5). This observation is the first example in our study when sex differences in response to early-life conditions are observed.

The sex specificity of the month-of-birth effects on adult lifespan is a puzzling observation, but it is also a reassuring one from the methodological point of view. Indeed, the data for men and women are taken from the same sources and are represented by the same set of parental variables (because they are brothers and sisters to each other). Therefore, any possible flaws in data collection and analysis (such as omission of important predictor variable, for example) should produce very similar biases both in males and females data. Instead we observe a clear-cut sex-specific effect, which is reassuring from the methodological perspective.

While discussing the greater response of female lifespan to the season of birth, it is interesting to see whether other traits such as female childlessness are also affected by the month of birth. Indeed, studies on Dutch women found that the birth distribution of childless women, as compared with fecunds, was best represented with bimodal curve with zeniths in January and July (Smits *et al.*, 1997). It is interesting to note that the two peaks for childlessness (January and July) seems to correspond well with the two observed minimums for female adult lifespan observed in our study (February and August—just only one month shift compared to childlessness findings).

Our finding that the month of February is "bad" month to be born for

female corresponds well with schizophrenia studies. The risk of schizophrenia is higher for persons, whose birth date is close to February, and this seasonal effect is more marked among females (Dassa *et al.*, 1995). It was also found that pre-natal exposure to influenza epidemic is associated with later development of schizophrenia in females but not in males (Takei *et al.,* 1993; 1994).

While discussing studies of month-of-birth effects, it is important to be aware of methodological problems and pitfalls. In some cases a simplistic approach is applied to study the effects of month of birth on human lifespan: mean ages at death are calculated for people born in different months using cross-sectional data, *i.e.*, death certificates collected during a relatively short period of time (Doblhammer, Vaupel, 2001). This methodology can produce both false positive and false negative findings. For example, if the seasonality of births and infant mortality were more expressed in the past, then the month-of-birth distribution of people would differ in different age groups of the population, thus producing a spurious month-of-birth effect on lifespan (if erroneously estimated through mean age at death). This mistake happens because the mean age at death depends on the age distribution of living people, which may differ depending on month-of-birth. Thus, even if the month of birth does not affect adult lifespan, nevertheless a false positive finding may occur, simply because the effects of population age structure are not taken into account. On the other hand, month-of-birth effects could be overlooked by this cross-sectional method if the seasonal effects on age-specific mortality rates are proportional. This false negative finding happens because proportional changes in death rates produce a proportional changes in the numbers of deaths in all age groups, and such proportional changes in numbers have no effect on the mean age at death. Thus, a false negative finding may occur, because cross-sectional analysis of death records is blind to proportional changes in age-specific death rates. In our study we avoided this simplistic cross-sectional analysis of death records as a flawed methodology. Instead we applied a cohort approach by following people born in the same calendar years until the last person died (method of extinct generations).

Finally, we would like to comment on the importance to control for socioeconomic status while studying the effects of month of birth. This is very important issue because there are significant differences in birth seasonality between different social classes (Smithers, Cooper, 1984; Bobak, Gjonca, 2001). Therefore, studies of aggregated data for whole countries (Doblhammer, Vaupel, 2001) may simply reflect the well-known differences in procreation habits of different socioeconomic classes. In our study we control for socioeconomic status by stratification (only aristocratic families are included into analysis).

Paternal age at person's birth and longevity

The dependence of female lifespan on paternal age at reproduction (when daughter was born) is presented in Table 6. Note that there is an optimal age to father a daughter, which is rather late — about 35-39 years (considered as reference level in this study). Daughters born

to older or younger fathers tend to live shorter lives on average (significant for older fathers at P<0.01, see Table 6). These are the net effects of paternal age, when all other co-variates (see "Data and Methods" section) are taken into account, including maternal age effects that surprisingly proved to be less important.

Shorter lifespan of daughters conceived to older fathers could be explained by age-related accumulation of mutations in DNA of paternal germ cells (Crow, 1997; Gavrilov, Gavrilova, 2000a; 2001a). Advanced paternal age at person's conception is an important risk factor for such disease of adult age as schizophrenia (Malaspina, 2001; Malaspina *et al*., 2001), and such disease of old age as sporadic (non-familial) Alzheimer disease (Bertram *et al*., 1998).

The practical importance of these findings is obvious: the age constrains for the donors of sperm cells in the case of IVF (in vitro fertilization) perhaps need to be revised to exclude not only the old donors, but also those donors who are too young. Of course, more detailed studies are required, before such an important practical recommendation could be made.

Again, all these interesting ideas and suggestions seems to fail when data on males are analyzed (Table 7).

In contrast to females, the male lifespan does not depend on paternal age at person's birth, at least in this particular data set. This observation is the second example in our study when sex differences in response to early-life conditions are observed.

Since only daughters inherit paternal X-chromosome, a hypothesis was suggested that mutation accumulation in X-chromosome of old sperm cells may be responsible for specific lifespan shortening of daughters conceived to older fathers (Gavrilov, Gavrilova, 1997a; 1997b; 2000a; 2001b).

Death of siblings in early life

Previous studies by other authors found that the disease load in early life (estimated through infant mortality rate), is an important early predictor for mortality in later life (Alter *et al*., 2001; Bengtsson, Lindstrom, 2000; 2001; Costa, 2000). In particular, study of Alter and colleagues found that death of siblings (a proxy for exposure to infectious diseases) has positive impact on late-life survival suggesting either selection of the most fit or early immunization against infectious diseases (Alter *et al*., 2001). Other studies found that exposure to infections early in life has negative effect on health many years later (Costa, 2000; Bengtsson, Lindstrom, 2000). The results of our analyses are presented in Table 8. They suggest that in early birth cohorts (1800-1854) the effect of sibling death on adult lifespan is not significant both for males and females. On the other hand, in more recent birth cohorts (1855-1880), when mortality declined, sibling death became significant negative predictor of survival later in life (Table 7).

Prospects for future research

There are several interesting directions for further development of these studies. The first research direction is related to the finding by George Alter and his colleagues that early-life conditions might have different effects in different

periods of later life (Alter *et al.*, 2001). We plan to elaborate on this issue by applying hazard modeling to our data set in future studies.

The second research direction is related to the finding made by Frans van Poppel and his colleagues that women's fecundability is associated with month of birth (Smits *et al.*, 1997). We plan to replicate this finding and to include fecundability variable in the future data analyses as the outcome variable, as well as the predictor/confounding variable for adult lifespan.

The third possible research direction is related to findings of George Alter and his colleagues on the importance of interfamily differences in adult mortality (Alter *et al.*, 2001). It would be interesting to take into consideration the "family effects" using the random effects model and to see how it affects our preliminary findings made in this study.

Finally, we acknowledge that the findings presented in this study should be interpreted with caution and need to be replicated on other data sets and with variety of analytical methods. However, the results of this study clearly indicate the need for separate analysis of data for males and females when late-life consequences of early-life conditions and events are explored. There is a definite need for subsequent full-scale studies of the effects of early-life conditions on sex-specific health outcomes in later life, and our pilot study presented here justifies the need of further work in this direction.

Natalia S. GAVRILOVA
Leonid A. GAVRILOV
Center on Aging, NORC
University of Chicago
1155 East 60th Street
Chicago, IL 60637
E-mail: nsgavril@midway.uchicago.edu
lagavril@midway.uchicago.edu.

Galina N. EVDOKUSHKINA
Victoria G. SEMYONOVA
Central Research
Institute of Public Health and Informatics,
Moscow
Russia.

BIBLIOGRAPHICAL REFERENCES

ALTER, G., ORIS, M., and BROSTRÖM, G. (2001), "The Family and Mortality: A Case Study from Rural Belgium", *Annales de Démographie Historique* 1, 11-31.

AMES, BN. (1998), "Micronutrients Prevent Cancer and Delay Aging", *Toxicology Letters,* 102-103, 5-18.

AUROUX, M., NAWAR, NN., NAGUIB, M., BAUD, M., LAPAQUELLERIE, N. (1998), "Post-Pubescent to Mature Fathers: Increase in Progeny Quality?", *Hum. Reprod.,* 13, 55-59.

BARKER, DJP. (1992), *Fetal and Infant Origins of Adult Disease*, London, BMJ Publishing Group.

BARKER, DJP. (1998), *Mothers, Babies, and Disease in Later Life*, 2nd edition, London, Churchill Livingstone.

BEETON, M., PEARSON, K. (1899), "Data for the Problem of Evolution in Man, II: A First Study of the Inheritance of Longevity and the Selective Death Rate in man", *Proceedings of the Royal Society of London*, 65, 290-305.

BEETON, M., PEARSON, K. (1901), "On the Inheritance of the Duration of Life and the Intensity of Natural Selection in Man", *Biometrika*, 1, 50-89.

BENGTSSON, T., 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", *Population Studies*, 54, 263-277.

BENGTSSON, T., LINDSTROM, M. (2001), "Early-life Conditions and Mortality in Later Life: Southern Sweden 1765-1894", Paper for the SSHA conference in Chicago, 15-18 November, 2001.

BERTRAM, L., BUSCH, R., SPIEGL, M., LAUTENSCHLAGER, NT., MULLER, U., KURZ, A. (1998), "Paternal Age is a Risk Factor for Alzheimer Disease in the Absence of a Major Gene", *Neurogenetics*, 1, 277-280.

BLACKWELL, DL., HAYWARD, MD. & CRIMMINS, EM. (2001), "Does Childhood Health Affect Chronic Morbidity in Later Life?", *Social Science & Medicine*, 52, 1269-1284.

BOBAK, M., GJONCA, A. (2001), "The Seasonality of Live Birth is Strongly Influenced by Socio-demographic Factors", *Human Reproduction,* 16, 1512-1517.

BRACKENRIDGE, CJ. (1980), "Bimodal Month of Birth Distribution in Cystic Fibrosis", *Am. J. Med. Genet.,* 5, 295-301.

BRESCHI, M., LIVI BACCI, M. (1986), "Season of Birth and Climate as Determinants of Infant Mortality in the Mainland Part of the Kingdom of Sardinia", *Genus*, 42, 87-101.

BUCHANAN, AV., WEISS, KM., SCHWARTZ, RJ., MACNAUGHTON, NL., MCCARTAN, MA., BATES, S.S. (1984), "Reconstruction of Genealogies from Vital Records: The Laredo Epidemiology Project", *Comput. Biomed. Res*., 17, 326-351.

COSTA, D. (2000), "Understanding the Twentieth-Century Decline in Chronic Conditions among Older Men", *Demography*, 37, 53-72.

CROW, JF. (1997), "The High Spontaneous Mutation Rate: Is it a Health Risk?", *Proc. Natl. Acad. USA,* 94, 8380-86.

DASSA, D., AZORIN, JM., LEDORAY, V., SAMBUC, R., GIUDICELLI, S. (1996), "Season of Birth and Schizophrenia: Sex Difference", *Prog. Neuro-Psychopharmacol. & Biol. Psychiat.*, 20, 243-251.

DOBLHAMMER, G., VAUPEL, JW. (2001), "Lifespan Depends on Month of Birth", *Proc. Natl. Acad. USA* 98, 2934-2939.

FOGEL, RW. (1993), "New Sources and New Techniques for the Study of Secular Trends in Nutritional Status, Health, Mortality, and the Process of Aging", *Historical Methods,* 26(1), 5-43.

FOGEL, RW. (1997), "Economic and Social Structure for an Ageing Population", *Phil. Trans. Royal Soc. London,* B 352, 1905-1917.

FOGEL, RW. (1999), "Catching up with the Economy", *Am. Economic Review,* 89, 1-21.

FOGEL, RW., COSTA, DL. (1997), "A Theory of Technophysio Evolution, with some Implications for Forecasting Population, Health Care Costs, and Pension Costs", *Demography*, 34, 49-66.

GAVRILOV, LA., GAVRILOVA, NS. (1991), *The Biology of Life Span: A Quantitative Approach*, NY, Chur, Harwood Academic Publisher.

GAVRILOV, LA., GAVRILOVA, NS. (1997a), "Parental Age at Conception and Offspring Longevity", *Reviews in Clinical Gerontology*, 7, 5-12.

GAVRILOV, LA., GAVRILOVA, NS. (1997b), "When Fatherhood Should Stop?", *Science,* 277, 17-18.

GAVRILOV, LA., GAVRILOVA, NS. (1999a), "Season of Birth and Human Longevity", *Journal of Anti-Aging Medicine,* 2, 365-366.

GAVRILOV, LA., GAVRILOVA, NS. (1999b), "Is There a Reproductive Cost for Human Longevity?", *Journal of Anti-Aging Medicine*, 2, 121-123.

GAVRILOV, LA., GAVRILOVA, NS. (2000a), "Human longevity and parental age at conception", 7-31, in *Sex and Longevity: Sexuality, Gender, Reproduction, Parenthood*, J.-M. Robine *et al.*, (eds), Berlin, Heidelberg, Springer-Verlag.

GAVRILOV, LA., GAVRILOVA, NS. (2000b), "Life Expectancy and the Month of Birth", *Healthy Life Expectancy*, REVES 12 Annual Meeting, March 20-22, Los Angeles, 2000, 34.

GAVRILOV, LA., GAVRILOVA, NS. (2001a), "The Reliability Theory of Aging and Longevity", *Journal of Theoretical Biology*, 213(4), (in press).

GAVRILOV, LA., GAVRILOVA, NS. (2001b), "Biodemographic Study of Familial Determinants of Human Longevity", *Population, English Selection,* 13(1), 197-222.

GAVRILOV, LA., GAVRILOVA, NS., EVDOKUSHKINA, GN., SEMYONOVA, VG., GAVRILOVA, AL., EVDOKUSHKINA, NN., LAPSHIN, EV. (1996), "Determinants of Human Longevity: Parental Age at Reproduction and Offspring Longevity", *Longevity Report* (ISSN 0964-5659), 10(54), 7-15.

GAVRILOV, A., GAVRILOVA, NS., KROUTKO, VN., EVDOKUSHKINA, GN., SEMYONOVA, VG., GAVRILOVA, AL., LAPSHIN, EV., EVDOKUSHKINA, NN., KUSHNAREVA, YuE. (1997), "Mutation Load and Human Longevity", *Mutation Research*, 377, 61-62.

GAVRILOVA, NS., GAVRILOV, LA.. (1999), "Data Resources for Biodemographic Studies on Familial Clustering of Human Longevity", *Demographic Research* [Online], vol.1(4), 1-48. Available: http://www.demographic–research.org/Volumes/Vol1/4/default.htm.

GAVRILOVA, NS., GAVRILOV, LA. (2001) "When Does Human Longevity Start?: Demarcation of the Boundaries for Human Longevity", *Journal of Anti-Aging Medicine*, 4(2), 115-124.

GAVRILOVA, NS., GAVRILOV, LA., EVDOKUSHKINA, GN., SEMYONOVA, VG., GAVRILOVA, AL., EVDOKUSHKINA, NN., KUSHNAREVA, YuE., KROUTKO, V.N., ANDREYEV, Ayu. (1998), "Evolution, Mutations and Human Longevity: European Royal and Noble Families", *Human Biology,* 70, 799-804.

GUTMANN, M., FLIESS, KH., HOLMES, AE., FAIRCHILD, AL., TEAS, WA. (1989), "Keeping Track of Our Treasures: Managing Historical Data with Relational Database Software", *Historical Methods,* 22(4), 128-143.

HOLLINGSWORTH, TH. (1962), "The Demography of the British Peerag", *Population Studies*, suppl., 18, 3-107.

HOLLINGSWORTH, TH. (1969), *Historical Demography,* Ithaca, N.Y., Cornell University Press.

JETTÉ, R., CHARBONNEAU, H. (1984), "Généalogies descendantes et analyse démographique", *Annales de Démographie Historique,* 45-54.

KASAKOFF, AB., ADAMS, JW. (1995), "The Effect of Migration on Ages at Vital Events: a Critique of Family Reconstitution in Historical Demography", *Eur. J. Pop.,* 11, 199-242.

KUH, D. & BEN-SHLOMO, B. (1997), *A Life Course Approach to Chronic Disease Epidemiology*, Oxford, Oxford University Press.

LEON, DA., LITHELL, HO., VÅGERÖ, D., KOUPILOVÁ, I., MOHSEN, R., BERGLUND, L., LITHELL, U-B. & MCKEIGUE, PM. (1998), "Reduced Fetal Growth Rate and Increased Risk of Death from Ischaemic Heart Disease: Cohort Study of 15000 Swedish Men and Women Born 1915-29", *Br. Med. J.*, 317, 241-245.

LUCAS, A (1991), "Programming by Early Nutrition in Man", 38-55I, in *The Childhood Environment and Adult Disease,* G.R. Bock, J. Whelan (eds), Chichester, Whiley.

LUCAS, A., FEWTRELL, MS. & COLE TJ. (1999), "Fetal Origins of Adult Disease—the Hypothesis Revisited", *Br. Med. J.,* 319, 245-249.

MALASPINA, D. (2001),"Paternal Factors and Schizophrenia Risk: de Novo Mutations and Imprinting", *Schizophr Bull.*, 27, 379-393.

MALASPINA, D., HARLAP, S., FENNIG, S., HEIMAN, D., NAHON, D., FELDMAN, D., SUSSER, ES. (2001), "Advancing Paternal Age and the Risk of Schizophrenia", *Arch. Gen. Psychiatry*, 58, 361-367.

MAYER, PJ. (1991), "Inheritance of Longevity Evinces no Secular Trend among Members of Six New England Families Born 1650-1874", *Am. J. Hum. Biol.,* 3, 49-58.

MCINTOSH, GC., OLSHAN, AF., BAIRD, PA. (1995), "Paternal Age and the Risk of Birth Defects in Offspring", *Epidemiology,* 6, 282-8.

OLSHAN, AF., SCHNITZER, PG., BAIRD, PA. (1994), "Paternal Age and the Risk of Congenital Heart Defects", *Teratology,* 50, 80-84.

POPE, CL. (1992), "Adult Mortality in America before 1900. A View from Family Histories", 267-296, in *Strategic Factors in Nineteenth Century American Economic History. A Volume to Honor Robert W. Fogel,* C. Goldin and H.Rockoff (eds.), Chicago and London, Univ. Chicago Press.

POST, W., VAN POPPEL, F., VAN IMHOFF, E., KRUSE, E. (1997), "Reconstructing the Extended Kin-network in the Netherlands with Genealogical Data: Methods, Problems, and results", *Pop. Studies,* 51, 263-278.

SKOLNICK, M., BEAN, LL., DINTELMAN, SM., MINEAU, G.(1979), "A Computerized Family History Data Base System", *Sociology and Social Research,* 63, 506-523.

SMITHERS, AG., COOPER, HJ. (1984), "Social-Class and Season of Birth", *Journal of Social Psychology,* 124, 79-84.

SMITS, LJ., VAN POPPEL, FW., VERDUIN, JA., JONGBLOET, PH., STRAATMAN, H., ZIELHUIS, GA. (1997),"Is Fecundability Associated with Month of Birth? An analysis of XIXth and Early XXth Century Family Reconstitution Data from the Netherlands", *Hum. Reprod.,* 12, 2572-2578.

TAKEI, N., O'CALLAGHAN, E., SHAM, PC., GLOVER, G., MURRAY, RM. (1993), "Does Prenatal Influenza Divert Susceptible Females from Later Affective Psychosis to Schizophrenia?", *Acta Psychiatr. Scand.,* 88, 328-336.

TAKEI, N., SHAM, P., O'CALLAGHAN, E., MURRAY, GK., GLOVER, G., MURRAY, RM; (1994), "Prenatal Exposure to Influenza and the Development of Schizophrenia: is the Effect Confined to Females?", *Am. J. Psychiatry,* 151, 117-119.

VAN POPPEL, F. (2000), "Long-term Trends in Relative Health Differences between Men and Women", *European Journal of Obstetrics & Gynecology*, 93, 119-122.

WASSER, SK., NORTON, G. (1993), "Baboons Adjust Secondary Sex-ratio in Response to Predictors of Sex-specific Offspring Survival", *Behavioral Ecology and Sociobiology*, 32, 273-281.

WYSHAK, G. (1978), "Fertility and Longevity of Twins, Sibs, and Parents of Twins", *Soc. Biol.,* 25, 315-30.

Tab. 1 *Characteristics of the data set*

Birth cohort (year of birth)	Mean Age at Death* ± Standard Error (years) Daughters (sample size)	Sons (sample size)
1800-1809	65.8 ± 0.7 (539)	63.8 ± 0.6 (543)
1810-1819	66.4 ± 0.7 (563)	63.5 ± 0.6 (629)
1820-1829	66.5 ± 0.6 (684)	64.0 ± 0.6 (668)
1830-1839	67.8 ± 0.6 (684)	63.6 ± 0.5 (587)
1840-1849	69.5 ± 0.6 (834)	63.7 ± 0.5 (890)
1850-1859	71.3 ± 0.5 (1,015)	64.1 ± 0.5 (1,081)
1860-1869	73.9 ± 0.4 (1,414)	66.3 ± 0.4 (1,472)
1870-1880	76.0 ± 0.3 (1,968)	65.8 ± 0.4 (1,840)

*Mean age at death is calculated for those persons who survived by age 30. This variable refers to "adult lifespan" in this study. The study data set consists 8,052 of males and 7,979 females.

Tab. 2 *Effects of month-of-birth on female lifespan. Birth cohorts: 1800-1854*

Month-of-birth	Net effect on adult lifespan, years (point estimate)	Standard Error	P value
February	0.00	Reference level	
March	0.88	1.29	0.4956
April	1.51	1.31	0.2494
May	2.21	1.31	0.0914
June	1.99	1.29	0.1212
July	*3.08*	*1.29*	*0.0173*
August	1.24	1.27	0.3297
September	1.29	1.33	0.3301
October	2.26	1.32	0.0861
November	2.07	1.29	0.1092
December	*2.86*	*1.30*	*0.0276*
January	1.01	1.30	0.4348
February	0.00	Reference level	

*Net effect corresponds to additional years of life gained (or lost) compared to the reference category. The data are point estimates of differential intercept coefficient adjusted for other predictor variables using multivariate regression with nominal variables.

Tab. 3 *Effects of month-of-birth on female lifespan. Birth cohorts: 1855-1880*

Month-of-birth	Net effect on adult lifespan, years (point estimate)	Standard Error	P value
February	0.00	Reference level	
March	0.51	1.14	0.6577
April	*2.39*	*1.13*	*0.0348*
May	*2.51*	*1.10*	*0.0222*
June	1.17	1.11	0.2913
July	0.94	1.11	0.3913
August	1.39	1.12	0.2142
September	1.12	1.10	0.3120
October	1.20	1.10	0.2765
November	2.00	1.18	0.0909
December	*2.32*	*1.14*	*0.0408*
January	0.57	1.14	0.6198
February	0.00	Reference level	

*Net effect corresponds to additional years of life gained (or lost) compared to the reference category. The data are point estimates of differential intercept coefficient adjusted for other predictor variables using multivariate regression with nominal variables.

Tab. 4 *Effects of month-of-birth on male lifespan. Birth cohorts: 1800-1854*

Month-of-birth	Net effect on adult lifespan, years (point estimate)	Standard Error	P value
February	0.00	Reference level	
March	0.21	1.19	0.8576
April	-0.23	1.20	0.8473
May	1.21	1.15	0.2945
June	0.90	1.18	0.4437
July	-0.84	1.16	0.4716
August	-0.53	1.16	0.6459
September	-0.12	1.16	0.9168
October	0.21	1.21	0.8647
November	-0.41	1.20	0.7297
December	1.50	1.20	0.2106
January	-0.31	1.17	0.7936
February	0.00	Reference level	

*Net effect corresponds to additional years of life gained (or lost) compared to the reference category. The data are point estimates of differential intercept coefficient adjusted for other predictor variables using multivariate regression with nominal variables.

Tab. 5 *Effects of month-of-birth on male lifespan. Birth cohorts: 1855-1880*

Month-of-birth	Net effect on adult lifespan, years (point estimate)	Standard Error	P value
February	0.00	Reference level	
March	0.71	1.12	0.5244
April	-0.94	1.13	0.4059
May	1.32	1.13	0.2439
June	0.83	1.12	0.4610
July	-0.33	1.09	0.7608
August	-0.37	1.12	0.7414
September	0.44	1.09	0.6888
October	-0.09	1.12	0.9330
November	-0.16	1.13	0.8851
December	-0.90	1.16	0.4391
January	1.29	1.14	0.2557
February	0.00	Reference level	

*Net effect corresponds to additional years of life gained (or lost) compared to the reference category. The data are point estimates of differential intercept coefficient adjusted for other predictor variables using multivariate regression with nominal variables.

Tab. 6 *Effect of paternal age at person's birth on female lifespan. Birth cohorts: 1800-1854*

Paternal Age	Net effect* on adult lifespan, years (point estimate)	Standard Error	P value
15-24	-1.27	1.38	0.3588
25-29	-0.98	0.86	0.2566
30-34	-0.18	0.76	0.8105
35-39	0.0	Reference level	
40-44	-0.31	0.85	0.7184
45-49	-0.10	1.08	0.9285
50-54	*-4.08*	*1.44*	*0.0046*

*Net effect corresponds to additional years of life gained (or lost) compared to the reference category. The data are point estimates of differential intercept coefficient adjusted for other predictor variables using multivariate regression with nominal variables.

Tab. 7 *Male lifespan as a function of paternal age at reproduction. Birth cohorts: 1800-1854*

Paternal Age	Net effect* on adult lifespan, years (point estimate)	Standard Error	P value
15-24	-2.06	1.28	0.1106
25-29	-0.62	0.81	0.4436
30-34	-0.58	0.69	0.4056
35-39	0.0	Reference level	
40-44	-1.22	0.79	0.1207
45-49	0.27	0.95	0.7789
50-54	-0.07	1.31	0.9600

*Net effect corresponds to additional years of life gained (or lost) compared to the reference category. The data are point estimates of differential intercept coefficient adjusted for other predictor variables using multivariate regression with nominal variables

Tab. 8 *Effect of sibling death (before age 18) on adult lifespan*

Birth cohort	Net effect on adult lifespan, years (point estimate)	Standard Error	P value
Females:			
1800-1854	-0.13	0.61	0.8371
1855-1880	*-1.87*	*0.61*	*0.0022*
Males:			
1800-1854	-0.21	0.55	0.7090
1855-1880	*-1.69*	*0.62*	*0.0062*

*Net effect corresponds to the years of life lost in a group with deceased sibling compared to the reference category when no siblings died before age 18. The data are point estimates of differential intercept coefficient adjusted for other predictor variables using multivariate regression with nominal variables.

SUMMARY

The idea of fetal origins of adult degenerative diseases and early-life programming of late-life health and survival is being actively discussed in the scientific literature. This idea is also important for understanding the historical changes in human lifespan through the mechanism of technophysio evolution as suggested by Robert Fogel and Dora Costa. Can this new concept also be useful to understand (at least partially) the observed sex disparities in adult health and longevity? Are the long-lasting effects of early-life conditions identical for both sexes, or, on the contrary, are they sex-specific? These questions stimulated us to conduct the present exploratory study on the sex specificity of late-life health outcomes for early-life effects.

In this study we explored the effects of early-life conditions on adult lifespan of 16,000 individuals (members of European aristocratic families born in 1800-1880) using methodology of historical follow-up study of extinct birth cohorts. Applying method of multivariate regression analysis with nominal predictor variables for individual lifespan as outcome variable, we found that sex differences in adult life span are indeed modulated by early-life events and conditions. Specifically, we found that such variables as (1) month of birth and (2) father's age at person's conception have statistically significant effects on adult lifespan (life expectancy at age 30) in females, but not in males. Female lifespan has bimodal distribution according to the month of birth (M-shaped curve), while male lifespan is less affected by the season of birth in our historical dataset. Similar M-shaped pattern of month-of-birth effects on adult lifespan was observed for females born in 1855-1880. Daughters born to old fathers (above 45 years) live significantly shorter lives, while sons are less affected by paternal age at conception. Death of siblings during childhood (often used as a proxy for childhood infections in family) had significant negative impact on adult lifespan of males and females for more recent birth cohorts (1855-1880), indicating possible increased selectivity of early mortality. The findings of this study provide scientific justification for the need of further more detailed studies on early-life programming of adult lifespan.

RÉSUMÉ

Les idées d'une origine fétale des maladies dégénératives adultes et d'une programmation dès l'aube de l'existence de la santé du grand âge ou de la survie sont au cœur d'un vaste débat scientifique. Ces idées sont importantes pour appréhender les variations historiques de la longévité humaine, selon les perspectives suggérées par Robert Fogel et Dora Costa. Peut-on y déceler une manière de rendre compte (au moins partiellement) des disparités existant entre sexes ? Cet article a pour objectif de mener une étude exploratoire sur les spécificités sexuelles de la santé adulte, conçues comme des effets de la situation prévalant aux jeunes âges.

Dans cet article, nous travaillons sur un échantillon de 16 000 individus, issus de familles aristocratiques européennes et nés dans les années 1800-1880, pour lesquels nous utilisons une méthode de suivi historique des cohortes de naissances. Grâce à une analyse régressive multivariée où l'espérance de vie individuelle est conçue comme une variable expliquée par un certain nombre de variables prédictives nominales, nous obtenons confirmation du fait que l'espérance de vie adulte est modulée par les évènements et les conditions du début de vie. Nous voyons notamment que des variables telles que le mois de naissance ou l'âge du père lors de la conception de l'enfant ont des effets statistiquement significatifs sur l'espérance de vie à 30 ans des femmes, mais non sur celle des hommes. La distribution des longévités féminines selon le mois de naissance obéit à une courbe bimodale (ou

en M), alors que le facteur saisonnier affecte peu les hommes. En outre, les filles conçues par un père âgé de plus de 45 ans ont des espérances de vie adulte significativement raccourcies, ce qui n'est pas le cas de leurs homologues masculins. L'existence de décès d'enfants dans la fratrie (souvent conçue comme un indicateur approximatif de la présence d'infections enfantines dans la famille) affecte négativement et significativement l'espérance de vie adulte des filles et des garçons, du moins pour les cohortes les plus récentes (1855-1880), suggérant un accroissement possible de la sélectivité de la mortalité précoce. De fait, les résultats de cette étude plaident en faveur d'un approfondissement des recherches scientifiques sur l'idée d'une programmation au début de l'existence de la longévité adulte.