Family matters: a genealogical inquiry into the familial component of longevity

Berg, N.M.A. van den

CHAPTER 2

HISTORICAL DEMOGRAPHY AND LONGEVITY GENETICS:

BACK TO THE FUTURE

Niels van den Berg[1]
Marian Beekman[1]
Ken Robert Smith[2]
Angelique A.P.O. Janssens[3]
P. Eline Slagboom[1]

[1] Department of Biomedical Data Sciences, section of Molecular Epidemiology,
Leiden University Medical Center, Albinusdreef 2, 2333 ZA Leiden, the Netherlands
[2] Department of Family and Consumer Studies; Population Sciences, Huntsman Cancer Institute,
University of Utah, 225 S. 1400 E. Rm 228 Salt Lake City, United States of America
[3] Radboud Group for Historical Demography and Family History,
Radboud University, Erasmusplein 1, 6525 HT Nijmegen, the Netherlands

Aging Research Reviews (September 2017)
https://doi.org/10.1016/j.arr.2017.06.005

Abstract

Research into the genetic component of human longevity can provide important insights in mechanisms that may protect against age-related diseases and multi-morbidity. Thus far only a limited number of robust longevity loci have been detected in either candidate or genome wide association studies. One of the issues in these genetic studies is the definition of the trait being either lifespan, including any age at death or longevity, i.e. survival above a diverse series of thresholds. Likewise heritability and segregation research have conflated lifespan with longevity. The heritability of lifespan estimated across most studies has been rather low. Environmental factors have not been sufficiently investigated and the total amount of genetic variance contributing to longevity has not been estimated in sufficiently well-defined and powered studies. Up to now, genetic longevity studies lack the required insights into the nature and size of the genetic component and the optimal strategies for meta-analysis and subject selection for Next Generation Sequencing efforts. Historical demographic data containing deep genealogical information may help in estimating the best definition and heritability for longevity, its transmission patterns in multi-generational datasets and may allow relevant additive and modifying environmental factors such as socio-economic status, geographical background, exposure to environmental effects, birth order, and number of children to be included. In this light historical demographic data may be very useful for identifying lineages in human populations that are worth investigating further by geneticists.

Introduction

During the past 200 years human life expectancy at birth significantly increased in western societies, with record female life expectancy increasing from 45 years in 1840 to 85 years in 2015[1]. Around 1950, even the oldest old (age 85 or older) started to show a pattern of extended life expectancy and today they are the fastest growing segment of older people[1]. This means that populations not only survive to higher ages than in the past, they also have a lower mortality rate, during their young and middle years[2]. Remarkably, the survival of a select few persons stands out of an otherwise aging population[3]. These persons were extremely long-lived and, most of all, showed little to no signs of age-related disease, allowing them to have extremely long and healthy lives[4–7]. Research into first-degree relatives of these long-lived individuals showed that they also had extremely long and healthy lives compared to relatives of individuals with more normative ages at death[8,9]. Hence, the familial component, including both genetic and environmental contributions, seemed to play a key role in gaining more knowledge about factors involved in healthy aging and in the capability to survive into extreme old ages (often called longevity)..

In the literature, the familial component of human longevity has been investigated using survival to extreme age and age at death as phenotypes of survival (see *Table 1*). The former actually refers to longevity whereas the latter refers to individual or population based lifespan. Both definitions are often used in the context of longevity research which is confusing and incorrect. Another complication is that most studies exclude infant and child mortality by applying a lower limit age threshold when considering the lifespan of a population or group of individuals. Unfortunately, there is no consensus on the age threshold for longevity studies. As a result of both the inconsistent use of terminology and different lower and upper limit age thresholds, the comparison of longevity studies is generally problematic[10]. We will refer to longevity as survival into extreme old ages whereas lifespan refers to age at death related measures (see *Table 1* and *Figure 1*).

Progress in longevity research is also hampered by the fact that longevity is likely dependent on an interplay between combinations of multiple genes and environmental factors[11–15] which makes it difficult to separate environmental from genetic influences. In fact, environmental influences likely moderate genetic effects on longevity[16–18]. Hence, in this review we describe how historical genealogical data can be used to study familial longevity by including family history information to identify longevous families with a high potential for genetic analysis, such as Next Generation Sequencing (NGS). We start by discussing the state of the art of genealogical heritability and segregation studies in the context of lifespan and longevity. Next we discuss the influence of environmental factors in longevity research, and finally we propose how historical genealogical and demographic data, and the results of genealogical studies can be included in genetic longevity research.

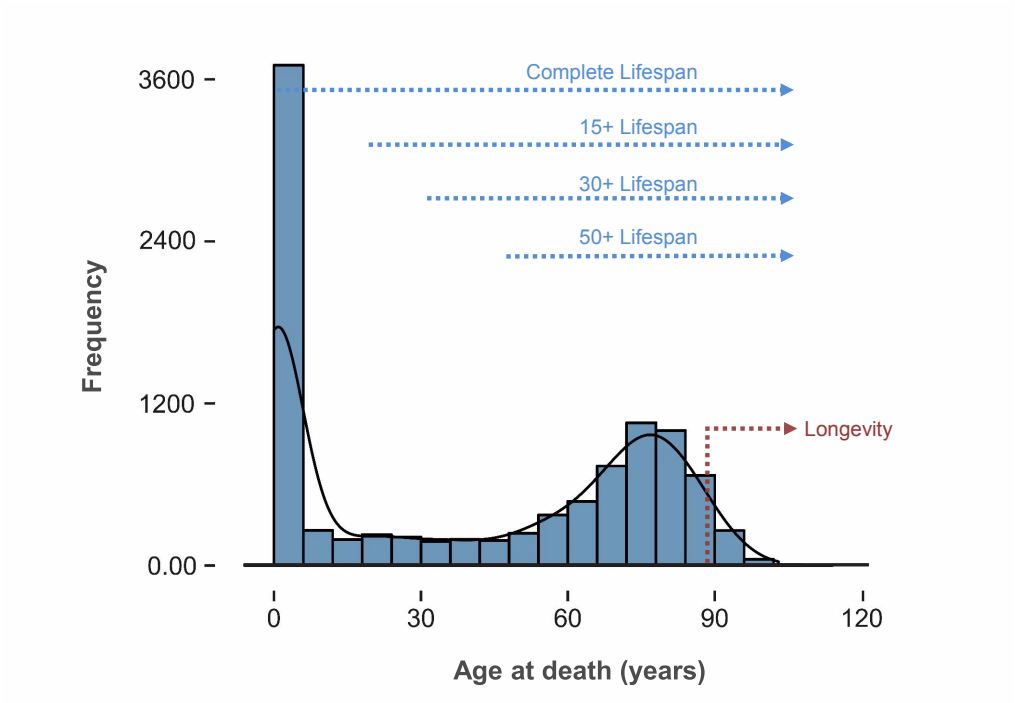
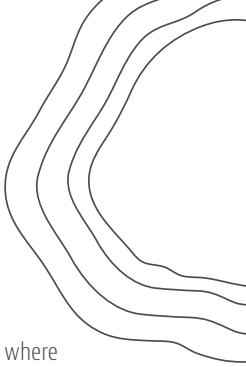


Figure 1: Difference between lifespan and longevity.

Figure is based on data from the Historical Sample of the Netherlands (1860 – 1875). This figure illustrates the distribution of "age at death" in the form of a histogram combined with a density plot. The bars in the histogram represent the number of individuals who died at the age depicted at the x-axis. The line is a density line representing the same concept as the bars. The x-axis represents age at death groups for HSN research persons born between 1860 and 1875. The y-axis represents the number of individuals who died in the different age at death groups. The distribution depicted in this figure is used to illustrate the difference between lifespan and longevity on an individual level in terms of the place of an individual within the distribution.



Heritability of longevity has not been established yet

The broad sense heritability (H^2) of a trait can be considered as the upper limit for genetic studies, where heritability coefficients can be seen as a progress indicator, indicating whether after identification of a first gene set for a trait, additional genes may still be determined. Heritability coefficients are differentially interpreted, depending on the type of data used for analysis. When estimated in genealogical data, heritability coefficients provide an estimation of the familial influence on a trait in which the combined effects of genes and shared environment within families are difficult to separate. As a consequence, heritability estimates depend on the environmental context[16,17]. Twin studies are more suitable than other genealogical studies to provide a first estimate of the influence of genetic, shared, and non-shared environmental influences on a trait. In practice, studies often report the narrow sense heritability (h^2), which is solely based on additive effects (see *Table 3* for a summary of key quantitative genetics concepts).

Research into siblings of centenarians showed that persons with a centenarian sibling have a four to eight times higher chance of becoming a centenarian as compared to persons with a sibling who died at a normative age[8]. A study into parent – offspring relations focused on parents belonging to the top 1 percent of their birth cohort and shows that these parents have a recurrence risk of 2.31 to have children who also belong to the oldest 1 percent of their birth cohorts[19]. Similarly, long-lived parents (>95th percentile) have a greater chance of having offspring who also live up to the 95th percentile or above[20]. Consistent with these findings, it has been shown that siblings of long-lived sib-pairs (men 89+ and women 91+), their parents, and their offspring live significantly longer than members of their own birth cohorts[21].

Table 1: Phenotypes of survival

Longevity and lifespan can be measured in multiple ways, mainly depending on specific research questions, time frame, and available information. Below the most common measures are set forth within the dichotomous framework of lifespan and longevity.

	Lifespan	**Longevity**
1. Age at death	X	-
2. Age at death > threshold	X	X
3. Cohort specific top X percentile	-	X

"X" means present and "–" means not present

age (at death): The most basic definition of lifespan refers to an individual's age at last observation or if known, the age at death. The age at death refers to the complete lifespan of an individual (Lutz et al., 2013). An advantage of using this definition is that it is easy to construct this measure and there is only little data loss.

age (at death) > certain threshold: This definition refers to the age at last observation or at death after surviving passed a specific age threshold. The advantage of this definition is that certain age specific biases can be controlled for by excluding individuals below the age threshold. Early life effects are often accounted for by using a lower limit age threshold of >15 or >30 years, whereas later life effects are often accounted for by using >90, or >100 years as an upper age threshold. Whether this measure represents lifespan or longevity depends on the height of the upper age threshold and its operationalization.

cohort specific top x percentile: This definition refers to the x percent most long-lived individuals depending on the cohort specific age at death distribution. The main advantage of this measure is that it can be used to eliminate the effects of secular trends.

Spouses of nonagenarian siblings did not show a survival advantage in the study of Schoenmaker et.al. (2006). Pedersen et al., however, did observe a survival advantage for spouses of long-lived siblings when comparing them to a birth cohort and sex matched control group. The authors attribute this survival advantage to assortative mating in their population[9]. An earlier Quebec study also reported a survival advantage of spouses[23] and a study of Southern Italy found male nonagenarians to outlive their spouses, whereas this was not the case for female nonagenarians[24]. Clearly, biological, environmental, and cultural factors influence survival to advanced ages in longevity families. These genealogical studies did not provide a quantification of the effects in terms of heritability estimates.

Several genealogical studies have attempted to estimate the heritability of lifespan and longevity (see supplemental data for a description of genealogical data). These studies can be divided into two categories based on the type of data they used; (1) twin data and (2) pedigree data. Unlike animal studies in a lab setting, the effects of the environment on longevity in human studies cannot be controlled. In twins at least the variation in early environment is minimized as compared to other family based studies. In all cases, heritability estimates and the effect of specific gene variants on lifespan and longevity depends on the populations studied and their past and present environmental conditions.

Twin studies

Twin studies have shown that genetic influences account for 1-27% of lifespan variation in the population (the overall heritability (h2 and H2) is between 0.01-0.27)[25–29]. In these studies minimum age thresholds were used, ranging from 15 to 37 years. Overall, twin studies rigorously differ, besides the variability in age thresholds, in their methodology, sample selection, and design. For example, a number of studies are unable to correctly establish twin zygosity[30]. Other studies result in inaccurate and overestimated heritability coefficients because they suffer from small sample sizes, censoring and truncation problems[26,31]. Taking these issues into account, we consider the twin studies of McGue et al., Herskind et al., and Ljungquist et al.[27–29], as the most robust (see *Table 2*).

McGue and colleagues estimated a heritability of 0.22 in a Nordic European twin sample of cohorts born between 1800 and 1950. They have found a minor and non-significant difference between men (H^2 = 0.23) and women (H^2 = 0.21) for lifespan, using an age threshold of 15 years. They have used structural equation modelling techniques to compare the fit of different models and concluded that there was significant evidence for non-additive effects and in particular for intra-locus interactions (dominance). Based on this dominance model a broad sense heritability coefficient of 0.22 was estimated which was larger than the heritability component for the additive model (h^2 = 0.13)[28]. These results have been replicated in the more recent study of Herskind et al. who came to the same conclusion, although the differences between the additive and the dominance model were more modest[29]. In addition, only one study distinguishes between twins reared together and twins reared apart, acknowledging the relevant environmental effects[32], which may limit the findings resulting from twin research[27]. The study has shown that the narrow sense heritability of lifespan beyond the age of 37 is 0.01 for men and 0.15 for women. However, these estimates are limited owing to low sample sizes for twins reared together (n_men = 82 and n_women = 97 pairs). Overall, the heritability of lifespan seems to be low and likely below 0.23.

Table 2: Overview of twin and genealogical heritability studies for lifespan and "longevity"

Twin studies

Study/population	Age	Method	Total n	h^2	Men n	h^2	Women n	h^2	Time span born	Phenotype	Ref.
Swedish twin registry | Sweden	37+	ICCs and SEM	358		164	0.01	194	0.15	1868 - 1925	Age at death (lifespan) and IMRs	27
Danish twin registry | Denmark	15+	ICCs based on ANOVA and SEM	1200	0.22| 0.33	652	0.23| 0.36	548	0.21| 0.32	1870 - 1880	Age at death in years and percentiles (lifespan)	28
Danish twin | Denmark	15+	ICCs based on ANOVA and SEM	5744		2816	0.26	2928	0.23	1870 - 1900	Age at death (lifespan)	29
GenomEUtwin | multiple countries	15+	Corrected ICCs	9334		4598	0.22	4736	0.32	1870 - 1910	Age at death (lifespan)	25
NAS-NRC twin registry | U.S.	19+	Corrected ICCs	31848		31848	0.54			1946 - 1978	Age at death (lifespan)	26

Pedigree studies

Study/population	Age	Method	Total n	h^2	Men n	h^2	Women n	h^2	Time span born	Phenotype	Ref.
MICROS study | Italy	50+	Variance components analysis	8277	0.15	4299	0.16	3978	0.18	1658 - 1907	Age at death (lifespan)	33
Genealogica sursilliana | Finland	15+	REML mixed-model	2614		1226	0.18	1388	0.17	1745 - 1903	Age at death (lifespan)	34
UPDB | US	30+	ANOVA with ML	14618	0.18	7601	0.14	7017	0.22	1850 – 1913	Excess longevity (lifespan)	35
UPDB | US	65+	Correlations (Rao model)	78994	0.15					1870 - 1907	Excess longevity (lifespan)	19
Royal & noble families | Europe	30+	Multiple linear regression	12150		8409	0.18	3741	0.20	?	Age at death (lifespan)	36
Villiage genealogies | Germany	0+	Correlation analysis for	9979		5315	0.20	4664	0.19	1650 - 1925	Age at death (lifespan)	37
OOA | US	30+	Variance components analysis	1655	0.25					1727 - 1890	Age at death (lifespan)	38
Valserine Valley | France	55+	ANOVA	1102	0.27					1745 - 1849	Age at death (longevity)	39
Village of Arthez d'Asson | Canada	20+	Correlations (Tau model)	2446	0.17					1686 - 1899	Age at death (lifespan)	40
OOA | US	30+	Variance components analysis	1655	0.25					1749 - 1890	Age at death (lifespan)	41

Heritability in all twin studies is based on differences between mono and dizygotic twins and heritability in all genealogical studies are based on parent offspring correlations. In twin studies we reported the broad sense heritabilities whereas narrow sense heritabilities are reported for the pedigree studies. For twin studies additional heritabilities are provided in the text. The Danish twin registry study from Denmark shows different heritability estimates. The top value in each column refers to "age at death in years" and the bottom value in each column refers to "age at death in percentiles". methods: 1.ICCs, 2.SEM, 3.ICCs based on ANOVA, 4.Variance components analysis, 5.REML mixed-model, 6.ANOVA with ML, 7.Correlations (Rao model), 8.Multiple linear regression, 9.Correlation analysis, 10.ANOVA, 11.Correlations (TAO model). List of abbreviations: intraclass correlation (ICC), analysis of variance (ANOVA), structural equation modelling (SEM), restricted maximum likelihood estimator (REML), Maximum likelihood (ML), Integrated Mortality Risk (IMR), age at death in years (AAD yrs), age at death in percentiles (AAD pct), excess longevity = EL. For an overview of the RAO and TAO model see the papers of Rao, et. al. and Cloninger et. al.[42,43]. All phenotypes are lifespan based, except Cournil et al. (2000) which is based on longevity. studies are prioritized, with the top rows showing the highest quality studies. twin studies: based on sample size, censoring, truncation, rearing, zygosity, and study design. pedigree studies: based on sample size, generalizability, and study design.

Table 3: Short introduction into quantitative genetics

A phenotypic trait can follow a Mendelian and a non-Mendelian inheritance pattern. If a Mendelian inheritance pattern is followed, the trait originates from the effect of only one gene and it is considered to have a discrete variance. However, most traits do not originate from only one gene and thus follow a non-Mendelian inheritance pattern. Such traits have a continuous variance and examples are: Intelligence and longevity.

Quantitative genetics focuses on mapping this continuous variance, distinguishing between additive and non-additive variance. Non-additive variance may be the result of gene interactions among gene effects either within (dominance) or between (epistasis) gene loci. Non-additive effects can be determined by establishing (dis) concordance between twins. Because of this, additive genetic variance always refers to genes directly transmitted from parents to their progeny.

A phenotypic trait is not only influenced by genetic effects but also by environmental factors. For example: monozygotic twins share 100 percent of their genes but as they age, they will phenotypically differentiate because of an accumulation of personal experiences and exposure to environmental factors.

In quantitative genetics the sum of variance for a phenotypic trait is designated as follows:

$$\sigma_p^2 = \sigma_g^2 + \sigma_e^2$$

Where σ_p^2 is the total phenotypic variance, σ_g^2 is the genetic variance and σ_e^2 is the environmental variance. The phenotypic variance can also be further broken down:

$$\sigma_p^2 = \sigma_a^2 + \sigma_d^2 + \sigma_i^2 + \sigma_e^2$$

Where σ_p^2 again is the total phenotypic variance, σ_a^2 is the additive variance, σ_d^2 is the variance due to dominant effects, σ_i^2 is epistatic variance, and σ_e^2 is environmental variance. The heritability of a phenotypic trait in a population (H^2) represents the amount of phenotypic variance explained by genetic differences. In its broadest meaning the heritability is given by:

$$H^2 = \frac{\sigma_g^2}{\sigma_p^2}$$

Where H^2 is the broad sense heritability, σ_g^2 is the is the genetic variance and σ_p^2 is the total variance. The heritability coefficient varies between 0 and 1 because the numerator in the fraction is smaller than the denominator and both are positive values. Selection can only affect additive genetic variance because dominant and epistatic components are broken by processes of recombination and independent segregation. Hence, a more strict definition of heritability is often used:

$$h^2 = \frac{\sigma_a^2}{\sigma_p^2}$$

Where h^2 is the narrow sense heritability, σ_a^2 represent the additive genetic effects, and σ_p^2 is the total variance.

In an attempt to investigate the heritability of surviving to advanced ages, Ljungquist et al. have estimated the heritability at different age cut-off values[27]. In this analysis the narrow sense heritability increased with age up to 0.28 in 80+ men and 0.23 in 85+ women which may be considered extreme ages (authors denote this as 'longevity') for the investigated birth cohorts (1886 – 1925). However, sample sizes at these extreme ages were small, and negative statistically insignificant heritability coefficients were estimated in the analysis for men at the age of 85 and for women at the age of 90, indicating statistical power problems. Moreover, it remains elusive whether the increase in heritability with age is statistically significant as this is not illustrated in the study[27]. Hence, compelling results have been obtained with regard to the heritability of lifespan, though the extreme heterogeneity in heritability estimates between studies may indicate that heritability estimates are strongly influenced by study size and environmental factors. The heritability of longevity has however not been robustly estimated as yet. Consequently, future heritability studies should make a more robust assessment of the role of the environment and of longevity as a trait.

Pedigree studies

Pedigree studies overall suffer from comparable issues as twin studies regarding small sample sizes, methodology, sample selection, and design issues. In pedigree studies the heritability of lifespan generally does not exceed 0.27[19,33–41] with the larger studies estimating the heritability to be below 0.20. Pedigree studies base their estimates on parent - offspring correlations (also indicating that they tend to report narrow sense heritabilities), which impedes the estimation of heritability coefficients that are less influenced by environmental effects. Hence, pedigree research is often conducted in extremely homogenous populations, such as the village of Arthez d'Asson, as environmental influences are relatively constant in such populations[36,40,41]. However, studies conducted in such homogenous populations have limited generalizability. An important benefit of pedigree studies over twin studies is the possibility of having access to a much larger sample size, especially for the older members of a population. Taking all these aspects into account, we consider four studies as the most accurate and robust[19,33–35]. These will be discussed in more detail.

Two studies from Utah have shown an estimated heritability for lifespan of 0.15 and 0.18 for persons above 65 and 30 respectively[19,35]. Another inquiry evinces that the heritability of lifespan for persons above 15 years is 0.18 and if stratified by sex it is 0.19 for men and 0.17 for women, although this difference is not statistically significant[34]. Another, elaborate study, conducted in three semi-isolated populations in Italy shows that the heritability of lifespan is 0.16 for men and 0.18 for women at the age of 50 and beyond. The joint heritability is estimated to be 0.15 and all estimates are corrected for confounding environmental effects. Moreover, this research illustrates that the heritability of lifespan above 50 years is constant during the 17th and 18th centuries and across different populations. This same study imposes different age thresholds for lifespan and concludes that the heritability increases with age at death to a maximum of 0.35. However, the heritability drops below 0.35 at the highest age thresholds and this is likely a function of small sample sizes. Furthermore, the study does not provide statistics of the increase in heritability estimates and besides that, ages at death were transformed into standardized scores which are difficult to interpret in relation to actual ages at death[33]. The heritability of lifespan seems comparable in pedigree and twin studies; it does not exceed 0.27. In pedigree studies, the heritability of longevity has been under-investigated and consequently, comparable to twin studies, the heritability of longevity has not been robustly estimated in pedigree studies. Moreover, pedigree studies also show a large variation in

heritabilities (0.15 – 0.27), which may be attributed to study size, selection criteria, and variation in environmental factors.

Longevity

The heritability of lifespan has been well documented by means of twin studies and pedigree studies, and it can be concluded that the heritability of lifespan is between 0.01 and 0.27 in the population at large. The large variation in the heritability estimates indicates a prominent role for differential environmental influences on the estimates. Studies showing that siblings of centenarians and longevous sib-pairs have a high probability to also become a centenarian or longevous, respectively, and studies, which show that longevous parents have a high probability to bear longevous offspring, provide indications that the heritability of longevity may be higher than that of lifespan[8,19–21]. However, the heritability of longevity has only been investigated once in a twin study design, though of limited sample size[27]. In addition, the heritability of longevity has been investigated more often in pedigree studies but the studies raise several questions about their design, sample size, and generalizability. Establishing the heritability of longevity is necessary for case definitions in genetic studies focused on gene mapping[21]. Hence, researchers' attention should shift from lifespan to longevity and the heritability of longevity should be estimated in an appropriate design with a sufficiently large sample. Both the heritability of lifespan and longevity should be investigated in different environments to investigate environmental influences.

Historical genealogical data in inheritance pattern research

Inheritance patterns of any complex trait generally provide insight into the contributions attributable to shared genes. Longevity is expected to be a complex trait with a complicated inheritance pattern, resulting from interactions between the environment and many genes[11–15]. Such effects may be additive or non-additive where one gene may be rate limiting over the action of another, or enhance or multiply the effect of another gene. A traits' genetic inheritance pattern can be investigated using historical genealogical data. In the context of survival to extreme ages the inheritance pattern has often been investigated by estimating correlations between the lifespan of parents and children[19,40,44–47] and, stratifying these correlations by sex[19,37,40,45–47]. The inheritance pattern of survival to extreme ages had also been investigated with survival analysis, logistic regression, and analysis of variance instead of basic correlations[20,39,48–55].

Table 4: Inheritance pattern studies using genealogical data

Study	Age	Method	Time-span	Sample size	Phenotype	Parent-offspring	Father-son	Father-daughter	Mother-son	Mother-daughter	Siblings / cousins
51	90+ versus 50-75	Linear and logistic regression	1850 -1884	1152	Lifespan and longevity	+	X	X	+	+	-
37	50+	Correlation	1650 - 1927	9979	Lifespan	+	X	X	+	+	-
48	Parents >90 Offspring >20	Survival analysis	1922 - 1930	8869	Lifespan	+	-	-	+	+	-
47	0+	Correlations	1860 - 1909	12317	Lifespan	-	+	X	+	+	-
44	0+	Correlation	1652 - 1850	32897	Lifespan	+	-	-	-	-	+
52	Parents >50 versus <50 Offspring 55+	ANOVA	1700 - 1976	2172	Lifespan	+	X	X	X	X	+
45	30+	Correlation	1800 - 1880	?	Lifespan	+	+	X	+	+	-
19	65+	Correlations	1870 - 1907	78994	Lifespan	+	-	-		-	+
56	30+	Regression	1800 - 1880	17136	Lifespan	-	+	+	+	+	-
53	Parents ? (>0) Offspring > 75	Logistic regression	700 - 1875	1132	Lifespan	-	X	+	X	+	-
39	Parents >50 versus <50 Offspring >55	ANOVA	1745 - 1849	3733	Lifespan	+	X	+	X	X	-
20	95th percentile	Logistic regression	1870 - 1900	1531	Longevity	+	-	-	-	-	+
40	>0, >20, and >50	Correlation	1686 - 1899	11102	Lifespan	-	+	+	+	+	+
49	50+	Cox regression	1850 - 1910	874	Lifespan	+	X	X	X	+	+
46	0+	Correlations	1820 - 1899	200	Lifespan	+	X	X	X	+	+
54	Parents >70 and > 80 Offspring <80 and >100	Logistic regression	1818 - 1898	8585	Longevity	-	+	X	X	+	-
55	Offspring 100+ Parents ?	Logistic regression And linear regression	?	1018	Longevity and lifespan	-	+	X	+	+	-
50	>0	Cox regression	<1919	759150	Lifespan	-	+	+	-	-	-
57	100+	Descriptives	1889 - 2003	207	Longevity	+	X	X	+	+	-
58	Parents 60+ Offspring <75 versus >90	T-tests	1887	193	Lifespan	-	+	+	X	X	-

-"+" ⟶ relation is found ; "-" ⟶ relation is not tested ; " X" ⟶ relation is not found. studies are prioritized with the top rows showing the highest quality studies. The prioritization is based on sample size, study design and population, control variables, and statistical analyses. phenotypes: 1=lifespan, 2=longevity. Methods: 1=linear regression, 2=logistic regression, 3=correlation , 4=survival analysis, 5=ANOVA, 6=descriptives, 7=T-tests.

Apart from the variety of analytical methods used in the literature, inheritance pattern research is very heterogeneous with regard to study designs. Most research has a cross-sectional nature using either a multiple cohort or a case – control design, in which a group of old persons is compared to a control group, over two generations[19,20,37,39,40,44–58]. Moreover, most studies focus on lifespan instead of longevity[20,53–55,57] and they often involve entire populations which are either extremely homogenous[44] or heterogeneous[19], depending on the research question. Homogenous populations suffer from generalizability problems whereas heterogeneous populations are difficult to analyze because of a larger amount of environmental variance and founder effects. Furthermore, a minimum of two generations should be available to conduct analyses (parents and their offspring). In practice a more than two generational approach has almost never been applied.

Patterns of inheritance

The main results of a range of pedigree studies are shown in *Table 4*. First, many studies have found evidence for a father – son inheritance pattern[19,40,45,47,50,54,55,58] although an equal number of studies has not found this evidence[37,39,49,51–53,57]. The same pattern can be observed for mother – son inheritance and the least evidence seems to point in the direction of a father – daughter inheritance pattern. Most evidence points in the direction of a mother – daughter pattern of inheritance with twelve confirming studies[37,40,45–51,53,55–57] and only three disconfirming studies[39,52,58].

Most of the evidence is not compelling because of persistent challenge of establishing a genetic inheritance pattern which is uninfluenced by environmental factors (e.g. Socio-economic status, mothers age at birth, and the physical environment)[33,44,46]. Furthermore, secular trends, caused by the increased average lifespan through improved nutrition, hygiene, and medical treatment, are important factors when comparing the lifespan of parents and their offspring[33,38,45,46]. Many attempts to control for secular trends and environmental factors by applying specific statistical techniques, including control variables, focusing on homogenous populations, and excluding infant and child mortality from the sample have not led to consistent results[19,33,37,44,46,52]. A few studies attempted a different approach by focusing on longevity instead of lifespan[20,54,55,57]. However, these studies did not examine more than two generations and focused on extremely homogenous populations. As a consequence the generalizability of their results may be limited. Thus, results of inheritance pattern studies have been largely inconsistent and strong differences exist between studies. A few studies stand out given their sample size and design, population, control variables, and statistical analyses and will be examined further here[37,47,51,55].

Kemkes-Grottenthaler (2004) found a significant correlation between maternal lifespan (lower limit age threshold 50 years) and the lifespan of sons and daughters by studying genealogies of two historical homogenous German villages during 1412 – 1912. Correlations between paternal lifespan and the lifespan of sons and daughters has also been estimated but no significant relationship was found[37]. Similarly, the study of Parman (2010) provided evidence for a correlation between maternal lifespan and the lifespan of sons and daughters. In contrast to Kemke-Grottenthaler (2004) the Parman (2010) study focused on the heterogenous setting of North Carolina during 1860 – 1909. The study included more than 12,000 individuals[47]. Deluty, Atzmon, Crandall, Barzilai, and Milman (2015) focused on a combination of lifespan and longevity instead of only lifespan. The authors defined longevity as being 100 years or above and focused

on 291 centenarians and their parents in the contemporary homogenous society of Ashkenazi Jews in the United States. They concluded that mothers of longevous men and women had significantly longer lifespans as compared to mothers of non longevous individuals. An attenuated but similar pattern could be seen for fathers although the lifespan of fathers did not differ between longevous and non longevous daughters. In addition, logistic regression models indicated that the odds of having longevous offspring increased for every 10 years of life achieved for mothers whereas this is not the case for fathers[55]. Lastly, a study focused on the homogenous population of Saguenay-Lac-St-Jean. The study investigated the familial transmission of longevity in a group of 576 individuals aged over 90 years as compared to an equally sized control group aged between 50 and 75 years in the time frame between 1950 and 1974. It was concluded that the probability of having longevous offspring (both boys and girls) was elevated with an increase in mothers' lifespan and that this is not the case for fathers[51]. Overall, there are several indications for maternal transmission of lifespan with some preference to daughters over transmission to sons. Furthermore we conclude that studies of the inheritance of longevity over multiple generation remain limited.

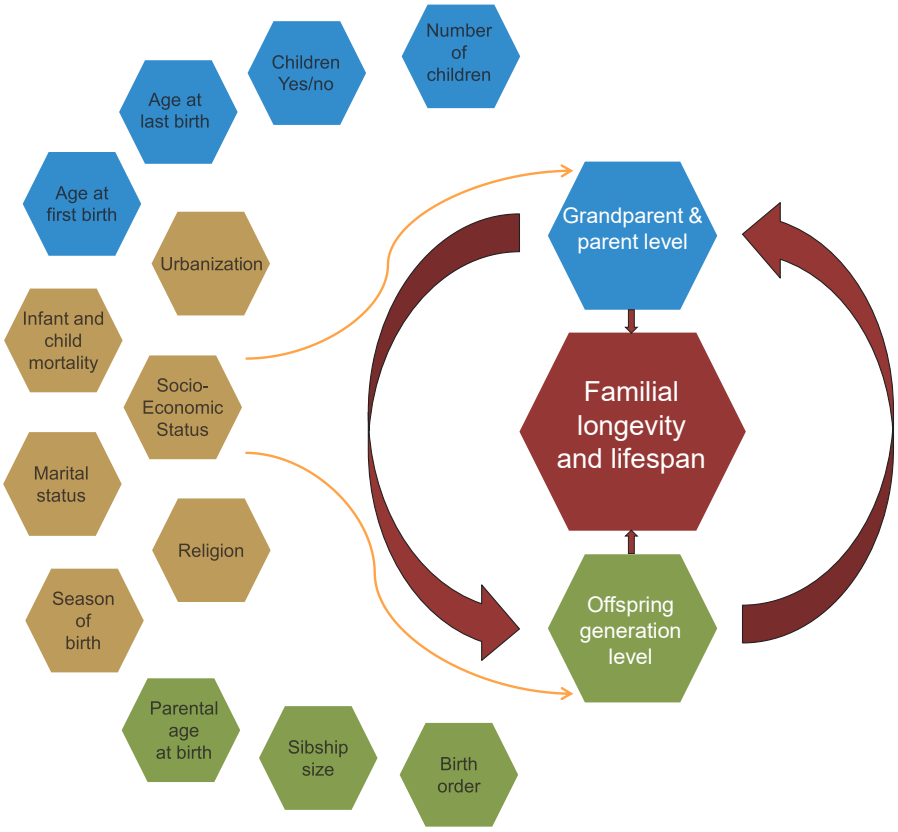


Figure 2: Covariates and confounders of lifespan and longevity.

The figure consist of two half circles. The outer rim mainly represents within family factors whereas the inner rim represents outside family factors. The blue color factors have a direct influence on the parental level (also depicted in blue). the green color factors have a direct influence on the offspring level (also depicted in green). The light-brown color predictors have an influence on both the parental and offspring levels which is illustrated by the arrows. The red arrows around the parental and offspring level indicate intergenerational effects of the factors illustrated in the left of the figure. The combining factor is the familial level of lifespan and longevity, which is depicted in the same red color as the arrows.

Up to now, the inheritance pattern of longevity has hardly been studied, with only two of the discussed studies estimating the (sex specific) effect of parental lifespan on the probability of having longevous offspring[51,55]. This is however not an optimal design to determine the inheritance pattern of longevity. Furthermore, longevity, is likely a polygenic trait, influenced by many environmental factors with small effects[11–15]. Hence, lifespan and longevity inheritance patterns may be influenced by environmental factors[52]. *Table 4* illustrates the large heterogeneity in inheritance pattern outcomes between studies. The study of Matthijs and colleagues (2002) is an example of how inheritance patterns may be influenced by the environment. The authors show that the inheritance pattern of lifespan in Moerzeke (Flanders) is different compared to a couple of Jura villages (France) using exactly the same methodological approach in a comparable time period (1700 - 1900)[39,52]. Ideally, multiple generations of families with a strong family history of extreme survival should be studied, which may reveal the interaction with environmental factors and may contribute to clarifiy the inheritance patterns by which longevity is transmitted.

Environmental influences in longevity research

Environmental factors such as socio-economic status, sibship size, parental age at birth, and geographical origin, reflecting exposure to epidemics, famines and war, are important variables within lifespan research[59–61]. This is because environmental factors can covary with and modify the lifespan of parents and children[18,39,52]. These factors can also confound the statistical relation between parents and their offspring, with respect to survival[19]. Longevity is derived from lifespan *(see Table 1)* and thus it can be expected that the same environmental factors which influence the results of genealogic research into lifespan also affect longevity research. In fact, some evidence for this exists but genealogical longevity research is scarce and sample sizes are generally small[54,62–67]. For example, one study found that environmental factors such as birth order and age at last birth slightly affected the relationship between parents and offspring longevity, defined as belonging to the oldest 5% of a person's birth cohort[54]. Hence, it is important to take environmental factors into account when inquiring into longevity, and because of this, we will outline the most important ones (see *Figure 2* for an overview).

Reproductive factors

Reproductive aging factors play a vital role in lifespan and longevity research[33,63–65,67–74]. The influence of reproductive factors on lifespan and longevity can take place on two levels; the level of the grandparents/parents and that of their offspring/subsequent generations. On the grandparents/parents level the following parameters will be described: Having children yes or no, parental age at first and last birth, and number of children. On the offspring/subsequent generations level, parental age at birth and birth order will be described.

Level of the grandparents/parents

The disposable soma theory suggest a biological trade-off between energy investment in reproduction and somatic maintenance[63,66,68,75–80]. This trade-off implies that an increase in the number of children causes a decrease in maternal lifespan. Such evolutionary trade-off has indeed been found in the study of laboratory animal models[81]. Just as the disposable soma theory, the maternal depletion theory suggests a trade-off between the number of children and maternal lifespan, although the theoretical mechanism is somewhat different[72,82]. In the maternal depletion theory the trade-off between number of children and maternal lifespan is explained by the emotional and physical investment of upbringing, and not necessarily a biological trade-off[72]. Hence, the maternal depletion theory also explains a paternal trade-off between the number of children and lifespan. In contrast to the maternal depletion and disposable soma theory, it is theorized that an increase in age at last birth is associated with an increase in maternal lifespan. One mechanism for this effect is that age at last birth may be a marker for general health and aging. Healthy aging persons may be predisposed to have slow aging tissues, which may subsequently cause the ability to reproduce late in life[83].

The theories that are described above have been extensively tested with genealogical data in natural fertility populations of various sample sizes, ranging from less than 100 to more than 10,000. One study found that on average women with children lived longer than women without children[84], but another study has not found evidence for this effect[85]. A few studies show that women without children reach older ages than women with children[71,86]. Similarly, many studies have

shown an increase in maternal lifespan if the number of children decreased[63,68,70,73,76,84,86–91], while only three studies found no effect at all[72,74,92] and two studies found the opposite effect[93,94]. For men, however, the relation between number of children and lifespan was inconsistent[63,66,68,74–76,87]. When it comes to mothers' age at last child, research has shown that an increase in age at last child is associated with an increase in maternal lifespan[75,80,83,89,94–96] and only two researchers found no link[68,76]. On the one hand, studies showed that if maternal age at birth increases, maternal lifespan equally increases[68,71,76,97]. Evidence for such an effect has also been provided for fathers[66,75]. On the other hand, studies have also shown negative effects for the relation between maternal age at birth and maternal lifespan6[3,98] or no relation at all[64,72,92]. All these reproductive effects on lifespan have typically been investigated for lifespan beyond 50 years in order to control for early deaths caused by childbearing.

For longevity, Tabatabaie (2011) showed that an increase in number of children correlates with a decrease in the odds of becoming 100+[63]. Tabatabaie et al. (2011) and sun et al. (2015) also showed that the odds of becoming longevous, defined as 100+, increase as the age at last child increases[63,65].

Level of the offspring generations

The human mutation rate of DNA base substitutions is high and increases with chronological age[99]. As a result, deleterious mutations in germ cells may cause a decrease in the lifespan of offspring as the parental age at conception increases[100]. The resource theory explains that being among the first children in the birth order may be beneficial for a persons' lifespan. Persons among the first in the birth order tend to receive more attention from their parents and do not have to share resources with multiple siblings[69].

A minority of studies focused on reproductive factors on the offspring level. The effects of birth order and paternal age at birth are relatively consistent. One small study showed that old fathers have daughters who die young as compared to young fathers. However, this study did not provide evidence for sons[100]. Furthermore, a large study of over 14,000 persons provided evidence that first born children live longer than those who are born later, regardless of their sex[69]. Lastly, all research on the offspring level focused on lifespan and no study looked into the effects on longevity.

Additional factors

Besides reproductive factors, other factors are also of significant importance for lifespan and longevity research. Persons with a high socio-economic status (SES) have a longer lifespan and a higher probability to become longevous than persons with a low SES[62,101–103]. This can mainly be attributed to the fact that high SES persons have better access to clean drinking water, high quality health care, and nutrition[13,95,104,105]. Furthermore, the season in which individuals are born has been shown to be an important measure. The effect of seasonality may be attributed to seasonal periods which encompass more danger for infections than others. Studies showed that the best months of birth are September until November[33,60,101]. Religion and marital status also influence lifespan. Religion is associated with a healthy lifestyle, causing religious persons to live a longer and healthier life than non-religious persons. Married persons also have a healthier lifestyle, explaining why they live longer and healthier than non-married persons[62,101]. Another important factor is the degree of urbanization, as urbanized areas have higher population densities than rural areas, making such areas

more susceptible to the distribution of diseases[101,106]. Infant mortality has also shown to associate with the survival of mothers, as infant mortality can be considered a proxy for maternal health[107,108]. Furthermore, both infant and child mortality can be considered proxies for children's early life circumstances[35]. Hence, not only reproductive factors are important for lifespan and longevity research, but also SES, marital status, religion, the degree of urbanization, and infant/child mortality.

Gene-environmental interactions in longevity research

Genetic effects in longevity studies are always influenced by different environmental factors[11–15] which can have additive and modifying influences. An example of additive influence is demonstrated if an individual is longevous because of a high SES and parental care, because the person was a first born. In this example, the modifying role of environmental factors is that the effect of some hypothetical gene set associated with longevity is more likely to become expressed in the phenotype when you are a firstborn child as compared to not being the firstborn child. The modifying nature of environmental factors can at best be addressed once genetic loci associating with longevity have been identified. In contrast, the additive nature of environmental factors can be used to screen potentially interesting persons for genetic analysis (this will be explained in more detail in chapter 5). In this regard genealogical data may provide new opportunities to record environmental effects which would otherwise have remained unknown in genetic studies.

Conclusions and future perspectives

In this review we focused on summarizing genealogical studies, and the beneficial role of genealogical studies and data for genetic longevity research. We conclude that lifespan and longevity research is very heterogeneous with regard to study designs, analytical methods, study populations, and results. This heterogeneity is problematic for comparative research. In addition, many studies have misused and misinterpreted the term longevity as it often refers to the complete lifespan of an individual or the lifespan beyond a certain lower limit threshold. As a result, many studies have investigated the heritability of lifespan instead of longevity. Irrespective of the twin or pedigree study design, the heritability estimates of lifespan are between 0.01 and 0.27 in the population at large, with an average of 0.25 (see *Table 2*). Inheritance pattern research has likely found evidence consistent with maternal transmission of both lifespan and longevity. This pattern has been identified on the basis of two generational analyses, which is a relatively weak design to identify a pattern of inheritance. When taking all inheritance pattern studies into account there is a large heterogeneity between the study results (see *Table 4*). As a next step, we suggest research into lifespan and longevity to take a standard minimum number of environmental factors into account (see *Figure 2*). Moreover, environmental factors in historical, demographic, and genealogical multi-generational data can be used to gain insight in the individual and family history of potentially interesting individuals for discovery genetics, such as by NGS. Selection of the most informative families and cases for these studies increases the probability to find longevity genes and one may gain insight in the differential role of the environment for specific gene variants. In conclusion, many studies have been conducted using different methodologies and focusing on different definitions of longevity. Hence, much knowledge has been gained from genealogical studies with regard to lifespan, though little is known about longevity and environmental influences (either additive or modifying) have been largely neglected in genetic lifespan and longevity research. Because of this, we propose a new approach and recommend integrating insights from genealogical studies into genetic studies to answer the still unsolved aspects of longevity.

Defining longevity in terms of the family over multiple generations

Lifespan and longevity are two distinct and incompatible concepts. In this paper, we defined longevity as survival into extreme old ages across an upper limit and lifespan as age at death behind a lower limit threshold such as 15, 30, or 50 years (see *Table 1*). However, in the literature, specific definitions aimed at quantifying the concept of "oldest old" are often used, where people need to have reached a certain age threshold (for example 75, 80, 85 years of age). Which persons actually are the oldest old in terms of absolute ages differs per time-period and population. By the use of absolute ages, comparisons across studies and populations become problematic and secular trends may cause extreme biases. Therefore, it has been suggested to define the age threshold for longevity at the oldest five percent of a population[10], allowing comparisons over time (including secular trends) and between populations. Limitations of taking a population percentage as age threshold for longevity come forward in certain specific study designs. When for example the oldest five percent is only 60 years of age such selection criteria will not refer to longevity. The percentiles should therefore always be weighted towards the life-tables of representative cohorts of an entire population. To sum up, lifespan and longevity are different concepts, which are preferably defined in terms of weighted percentiles instead of absolute ages.

Using the oldest five percent of a birth cohort seems appealing, but evidence that this oldest five percent of singletons indeed represents the best impact of genetic influence on longevity is not available. There is no evidence that using the oldest five percent creates new opportunities to distinguish genetic effects from environmental effects. Belonging to the oldest five percent may still harbor phenocopies caused by healthy lifestyles and excellent health care[109]. We propose to first investigate what the best definition of longevity is. This can be done by studying families over multiple generations in genealogic and demographic databases. Such data will allow the testing of different longevity definitions that reveal the most prominent familiarity in excess survival. One may compute whether an optimal familial clustering of longevity is observed for the 95th or 99th percentile surviving singletons, or for example, first degree relative-pairs (siblings, parents/offspring etc.).

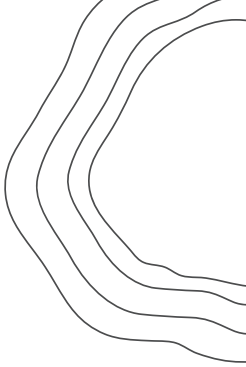
In addition to the definition problem of longevity hampering genetic research in detecting common genetic variants, research should focus on rare genetic variants in long-lived families; longevity is probably dependent on many genes with relatively small effects[11–15]. Some attempts have been made to identify rare variants contributing to human longevity by whole genome/exome sequencing of extremely long surviving individuals with as yet little robust findings[110,111]. In search for rare variants, we propose to select (in NGS efforts) families based on multiple generations of long-lived (top survival percentiles) descendants/first degree relatives, distant long-lived relatives, and to include environmental factors (such as birth order, SES, physical environment, the presence of a war or famine, and cause of death). Of course worldwide (joined) efforts will be needed for these analyses since the environment may modulate genetic effects, confining their detection.

Conflict of interest

The authors have nothing to disclose.

Funding

This work was supported by the Netherlands Organization for Scientific Research [360-53-180].



References

1 Oeppen, J. & Vaupel, J. W. Broken limits to life expectancy. *Science* **296**, 1029–1031 (2002).

2 Wachter, K. W. & Finch, C. E. *Between Zeus and the Salmon: The Biodemography of Longevity. Population (English Edition)* (Washington, DC: The National Academies Press, 1997).

3 Christensen, K. & Vaupel, J. W. Determinants of longevity: genetic, environmental and medical factors. *J. Intern. Med.* **240**, 333–341 (1996).

4 Andersen, S. L., Sebastiani, P., Dworkis, D. a., Feldman, L. & Perls, T. T. Health Span Approximates Life Span Among Many Supercentenarians: Compression of Morbidity at the Approximate Limit of Life Span. *Journals Gerontol. Ser. A Biol. Sci. Med. Sci.* **67A**, 395–405 (2012).

5 Ash, A. S. *et al.* Are Members of Long-Lived Families Healthier Than Their Equally Long-Lived Peers? Evidence From the Long Life Family Study. *Journals Gerontol. Ser. A Biol. Sci. Med.* Sci. **70**, 971–976 (2015).

6 Christensen, K., McGue, M., Petersen, I., Jeune, B. & Vaupel, J. W. Exceptional longevity does not result in excessive levels of disability. *Proc. Natl. Acad. Sci.* **105**, 13274–13279 (2008).

7 Evert, J., Lawler, E., Bogan, H. & Perls, T. Morbidity Profiles of Centenarians: Survivors, Delayers, and Escapers. *Journals Gerontol. - Med. Sci.* **58**, 232–237 (2003).

8 Perls, T. T. *et al.* Life-long sustained mortality advantage of siblings of centenarians. *Proc. Natl. Acad. Sci.* **99**, 8442–8447 (2002).

9 Pedersen, J. K. *et al.* The Survival of Spouses Marrying Into Longevity-Enriched Families. *Journals Gerontol. Ser. A Biol. Sci. Med. Sci.* **72**, 109–114 (2017).

10 Sebastiani, P., Nussbaum, L., Andersen, S. L., Black, M. J. & Perls, T. T. Increasing Sibling Relative Risk of Survival to Older and Older Ages and the Importance of Precise Definitions of "Aging," "Life Span," and "Longevity". *Journals Gerontol. Ser. A Biol. Sci. Med. Sci.* **71**, 340–346 (2016).

11 Deelen, J., Beekman, M., Capri, M., Franceschi, C. & Slagboom, P. E. Identifying the genomic determinants of aging and longevity in human population studies: progress and challenges. *Bioessays* **35**, 386–96 (2013).

12 Shadyab, A. H. & LaCroix, A. Z. Genetic factors associated with longevity: A review of recent findings. *Ageing Res. Rev.* **19**, 1–7 (2015).

13 Christensen, K., Johnson, T. E. & Vaupel, J. W. The quest for genetic determinants of human longevity: challenges and insights. *Nat. Rev. Genet.* **7**, 436–48 (2006).

14 Finch, C. E. & Tanzi, R. E. Genetics of Aging. *Science (80-.).* **278**, 407–411 (1997).

15 Kirkwood, T. B. L., Cordell, H. J. & Finch, C. E. Speed-bumps ahead for the genetics of later-life diseases. *Trends Genet.* **27**, 387–388 (2011).

16 Lewontin, R. C. Annotation: the analysis of variance and the analysis of causes. *Am. J. Hum. Genet.* **26**, 400–11 (1974).

17 Rose, S. P. R. Commentary: Heritability estimates - Long past their sell-by date. *Int. J. Epidemiol.* **35**, 525–527 (2006).

18 Montesanto, A. et al. Demographic, genetic and phenotypic characteristics of centenarians in Italy: Focus on gender differences. *Mech. Ageing Dev.* (2017). doi:10.1016/j.mad.2017.04.008

19 Kerber, R. A., Brien, E. O., Smith, K. R. & Cawthon, R. M. Familial Excess Longevity in Utah Genealogies. *Journals Gerontol. Ser. A Biol. Sci. Med. Sci.* **56**, 130–139 (2001).

20 Gudmundsson, H., Gudbjartsson, D. F. & Kong, A. Inheritance of human longevity in Iceland. *Eur. J. Hum. Genet. EJHG* **8**, 743–749 (2000).

21 Schoenmaker, M. et al. Evidence of genetic enrichment for exceptional survival using a family approach: the Leiden Longevity Study. *Eur. J. Hum. Genet.* **14**, 79–84 (2006).

22 *Vienna Yearbook of Population Research 2013.* (Verlag der ¨ Osterreichischen Akademie derWissenschaften, 2013).

23 Jarry, V., Gagnon, A. & Bourbeau, R. Survival advantage of siblings and spouses of centenarians in 20th-century Quebec. *Can. Stud. Popul.* **39**, 67 (2013).

24 Montesanto, A. *et al.* The genetic component of human longevity: analysis of the survival advantage of parents and siblings of Italian nonagenarians. *Eur. J. Hum. Genet.* **19**, 882–886 (2011).

25 Skytthe, A. *et al.* Longevity Studies in GenomEUtwin. *Twin Res.* **6**, 448–454 (2003).

26 Hrubec, Z. & Neel, J. V. Familial factors in early deaths: Twins followed 30 years to ages 51?61 in 1978. *Hum. Genet.* **59**, 39–46 (1981).

27 Ljungquist, B., Berg, S., Lanke, J., McClearn, G. E. & Pedersen, N. L. The Effect of Genetic Factors for Longevity: A Comparison of Identical and Fraternal Twins in the Swedish Twin Registry. *Journals Gerontol. Ser. A Biol. Sci. Med. Sci.* **53A**, 441–446 (1998).

28 McGue, M., Vaupel, J. W., Holm, N. & Harvald, B. Longevity Is Moderately Heritable in a Sample of Danish Twins Born 1870-1880. *J. Gerontol.* **48**, B237–B244 (1993).

29 Herskind, A. M. et al. The heritability of human longevity: A population-based study of 2872 Danish twin pairs born 1870–1900. *Hum. Genet.* **97**, 319–323 (1996).

30 Wyshak, G. Biodemography and Social Biology Fertility and longevity in twins , sibs , and parents of twins Parents of Twins. *Biodemography Soc. Biol.* **25**, 315–330 (2010).

31 Jarvik, L. F., Falek, A. & Kallmann, F. J. Survival Trends in a Senescent Twin Population. *Am. J. Hum. Genet.* **12**, 170–179 (1960).

32 Falconer, D. S. & Mackay, T. F. *Introduction to quantitative genetics.* (Pearson Education Limited, 1995).

33 Gögele, M. *et al.* Heritability analysis of life span in a semi-isolated population followed across four centuries reveals the presence of pleiotropy between life span and reproduction. *J. Gerontol. A. Biol. Sci. Med. Sci.* **66**, 26–37 (2011).

34 Pettay, J. E., Kruuk, L. E. B., Jokela, J. & Lummaa, V. Heritability and genetic constraints of life-history trait evolution in preindustrial humans. *PNAS* **102**, 2838–2843 (2005).

35 Hanson, H. A., Smith, K. R. & Hasstedt, *S. Heritability of longevity and the role of early and mid-life environments.* **022095**, (2012).

36 Gavrilova, N. *et al.* Evolution, mutations, and human longevity European royal and noble families. *Hum. Biol.* **70**, 799–804 (1998).

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

38 Mitchell, B. D. et al. Heritability of Life Span in the Old Order Amish. *Am. J. Med. Genet.* **352**, 346–352 (2001).

39 Cournil, A., Legay, J. M. & Schächter, F. Evidence of sex-linked effects on the inheritance of human longevity: a population-based study in the Valserine valley (French Jura), 18-20th centuries. *Proc. Biol. Sci.* **267**, 1021–5 (2000).

40 Bocquet-Appel, J. P. & Jakobi, L. Familial transmission of longevity. *Ann. Hum. Biol.* **17**, 81–95 (1990).

41 Sorkin, J. e*t al.* Exploring the genetics of longevity in the Old Order Amish. *Mech. Ageing Dev.* **126**, 347–50 (2005).

42 Cloninger, C. R., Rice, J. & Reich, T. Multifactorial Inheritance with Cultural Transmission and Assortative Mating II . A General Model of Combined Polygenic and Cultural Inheritance. *Hum. G* **31**, 176–198 (1979).

43 Rao, D. C., Morton, E. N. & Yee, S. Analysis of Family Resemblance. II. A Linear Model for FAmilial Correlation. *Hum. Genet.* **26**, 331–359 (1974).

44 Piraino, P., Muller, S., Cilliers, J. & Fourie, J. The transmission of longevity across generations: The case of the settler Cape Colony. *Res. Soc. Stratif. Mobil.* **35**, 105–119 (2014).

45 Gavrilov, L. a., Gavrilova, N. S., Olshansky, S. J. & Carnes, B. a. Genealogical data and the biodemography of human longevity. *Biodemography Soc. Biol.* **49**, 160–173 (2002).

46 Philippe, P. & Opitz, J. M. Familial correlations of longevity: An isolate-based study. *Am. J. Med. Genet.* **2**, 121–129 (1978).

47 Parman, J. Gender *and Intergenerational Mobility: Using Health Outcomes to Compare Intergenerational Mobility Across Gender and Over Time*. (2010).

48 Abbott, M. H., Abbey, H., Bolting, D. R. & Murphy, E. A. The Familial Component in Longevity-A Study of Offspring of Nonagenarians: 111 . Intrafamilial Studies. *Am. J. Med. Genet.* **2**, 105–120 (1978).

49 Salaris, L., Tedesco, N. & Poulain, M. Familial transmission of human longevity : a population-based study in an inland village of Sardinia (Italy), 1850 – 2010. *Vienna Yearb. Popul. Res.* **11**, 325–349 (2013).

50 Lach, S., Riton, Y. & Simhon, A. *Longevity across generations*. (2006).

51 Houde, L., Tremblay, M. & Vézina, H. Intergenerational and Genealogical Approaches for the Study of Longevity in the Saguenay-Lac-St-Jean Population. *Hum. Nat.* **19**, 70–86 (2008).

52 Matthijs, K., Putte, B. & Vlietinck, R. The Inheritance of Longevity in a Flemish Village. *Eur. J. Popul. Eur. Démographie* **18**, 59–81 (2002).

53 Westendorp, R. G. J. & Kirkwood, T. B. L. Maternal and Paternal Lines of Familial Longevity. *Popul. An English Sel.* **13**, 223–235 (2001).

54 Danzhen You, Danan Gu & Zeng Yi. Familial Transmission of Human Longevity Among the Oldest-Old in China. *J. Appl. Gerontol.* **29**, 308–332 (2010).

55 Deluty, J. A., Atzmon, G., Crandall, J., Barzilai, N. & Milman, S. The influence of gender on inheritance of exceptional longevity. *Aging (Albany. NY).* **7**, 412–418 (2015).

56 Gavrilov, L. A. & Gavrilova, N. S. Biodemographic Study of Familial Determinants of Human Longevity. *Popul. An English Sel.* **13**, 197–221 (2001).

57 Caselli, G. *et al.* Family clustering in Sardinian longevity: A genealogical approach. *Exp. Gerontol.* **41**, 727–736 (2006).

58 Hallett, M., Human, R. a T. & Physiology, C. familial longevity: the older you are the older your father may have been. *J. Am. Geriatr. Soc.* **53**, 357–358 (2005).

59 Smith, K. R., Mineau, G. P., Garibotti, G. & Kerber, R. Effects of childhood and middle-adulthood family conditions on later-life mortality: Evidence from the Utah Population Database, 1850-2002. *Soc. Sci. Med.* **68**, 1649–1658 (2009).

60 Gavrilov, L. a. & Gavrilova, N. S. New Developments in the Biodemography of Aging and Longevity. *Gerontology* **61**, 364–371 (2014).

61 Le Bourg, É. Does reproduction decrease longevity in human beings? *Ageing Res. Rev.* **6**, 141–149 (2007).

62 Palmore, E. Predictors of successful aging. *Gerontologist* **19**, 427–431 (1979).

63 Tabatabaie, V. *et al.* Exceptional longevity is associated with decreased reproduction. *Agign* **3**, 202–205 (2011).

64 Robine, J., Cournil, A., Henon, N. & Allard, M. Have centenarians had younger parents than the others? *Exp. Gerontol.* **38**, 361–365 (2003).

65 Sun, F. et al. Extended maternal age at birth of last child and women's longevity in the Long Life Family Study. *Menopause* **22**, 26–31 (2015).

66 Helle, S., Käär, P. & Jokela, J. Human longevity and early reproduction in pre-industrial Sami populations. *J. Evol. Biol.* **15**, 803–807 (2002).

67 Jarry, V., Gagnon, A. & Bourbeau, R. *Parental Age at Birth and Longevity of Offspring in Centenarian Families: The role of biology, social interaction and culture.* (2012).

68 Doblhammer, G. & Oeppen, J. Reproduction and longevity among the British peerage: the effect of frailty and health selection. *Proc. Biol. Sci.* **270**, 1541–1547 (2003).

69 Modin, B. Birth order and mortality : a life-long follow-up of 14,200 boys and girls born in early 20th century Sweden. *Soc. Sci. Med.* **54**, 1051–1064 (2002).

70 Westendorp, R. G. J. & Kirkwood, T. B. L. Human longevity at the cost of reproductive success. *Nature* **396**, 743–746 (1998).

71 Lycett, J. E., Dunbar, R. I. M. & Voland, E. Longevity and the costs of reproduction in a historical human population. *Proc. R. Soc. London. Ser. B Biol. Sci.* **267**, 31–35 (2000).

72 Alter, G., Dribe, M. & Poppel, F. Van. Widowhood, Family Size, and Post-Reproductive Mortality: A Comparative Analysis of Three Populations in Nineteenth-Century Europe. *Demography* **44**, 785–806 (2007).

73 Harrell, C. J., Smith, K. R. & Mineau, G. P. Are Girls Good and Boys Bad for Parental Longevity? *Hum. Nat.* **19**, 56–69 (2008).

74 Hurt, L. S. *et al.* Long-term effects of reproductive history on all-cause mortality among adults in rural Bangladesh. *Stud. Fam. Plann.* **35**, 189–196 (2004).

75 McArdle, P. F. *et al.* Does having children extend life span? A genealogical study of parity and longevity in the Amish. *J. Gerontol. A. Biol. Sci. Med. Sci.* **61**, 190–195 (2006).

76 Mueller, U. Does late reproduction extend the life span? Findings from European royalty. *Popul. Dev. Rev.* **30**, 449–466 (2004).

77 Kirkwood, T. B. & Rose, M. R. Evolution of senescence: late survival sacrificed for reproduction. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **332**, 15–24 (1991).

78 Bowen, R. L. & Atwood, C. S. Living and dying for sex: A theory of aging based on the modulation of cell cycle signaling by reproductive hormones. *Gerontology* **50**, 265–290 (2004).

79 Atwood, C. S. & Bowen, R. L. The reproductive-cell cycle theory of aging: An update. *Exp. Gerontol.* **46**, 100–107 (2011).

80 Perls, T. T., Alpert, L. & Fretts, R. C. Middle-aged mothers live longer. *Nature* **389**, 133 (1997).

81 Robins, C. & Conneely, K. N. Testing evolutionary models of senescence: traditional approaches and future directions. *Hum. Genet.* **133**, 1451–1465 (2014).

82 Merchant, K. & Martorell, R. Frequent reproductive cycling: Does it lead to nutritional depletion of mothers? *Prog. Food Nutr. Sci.* **12**, 339–369 (1988).

83 Crawford, S. L. What explains the link between reproductive events and women's longevity? *Menopause* **22**, 6–8 (2015).

84 Lund, E., Arnesen, E. & Borgan, J. K. Pattern of childbearing and mortality in married women--a national prospective study from Norway. *J. Epidemiol. Community Health* **44**, 237–240 (1990).

85 Gavrilova, N. S., Gavrilov, L. A., Semyonova, V. G. & Evdokushkina, G. N. Does Exceptional Human Longevity Come with a High Cost of Infertility. *Ann. N. Y. Acad. Sci.* **1019**, 513–517 (2004).

86 Friedlander, N. The relation of lifetime reproduction to survivorship in women and men: A prospective study. *Am. J. Hum. Biol.* **8**, 771–783 (1996).

87 Dribe, M. Long-term effects of childbearing on mortality: evidence from pre-industrial Sweden. *Popul. Stud. (NY).* **58**, 297–310 (2004).

88 Helle, S., Lummaa, V. & Jokela, J. Sons reduced maternal longevity in preindustrial humans. *Science* **296**, 1085 (2002).

89 Doblhammer, G. Reproductive history and mortality later in life: a comparative study of England and Wales and Austria. *Popul. Stud. (NY).* **54**, 169–176 (2000).

90 Green, a, Beral, V. & Moser, K. Mortality in women in relation to their childbearing history. *BMJ* **297**, 391–395 (1988).

91 Kvåle, G., Heuch, I. & Nilssen, S. Parity in relation to mortality and cancer incidence: a prospective study of Norwegian women. *Int. J. Epidemiol.* **23**, 691–699 (1994).

92 Beise, J. & Voland, E. Effect of producing sons on maternal longevity in premodern populations. *Science* **298**, 317; author reply 317 (2002).

93 Philippe, P. & Yelle, L. Effect of family size on mother's longevity. *Ann. Hum. Biol.* **3**, 431–439 (1976).

94 Müller, H., Chiou, J., Carey, J. R. & Wang, J. Fertility and Life Span : Late Children Enhance Female Longevity. *Journals Gerontol. Ser. A Biol. Sci. Med. Sci.* **57**, 202–206 (2002).

95 Smith, K. R., Mineau, G. P. & Bean, L. L. Fertility and post-reproductive longevity. *Soc. Biol.* **49**, 185–205 (2002).

96 Helle, S., Lummaa, V. & Jokela, J. Are reproductive and somatic senescence coupled in humans? Late, but not early, reproduction correlated with longevity in historical Sami women. *Proc. Biol. Sci.* **272**, 29–37 (2005).

97 Yi, Z. & Vaupel, J. Association of late childbearing with healthy longevity among the oldest-old in China. *Popul. Stud. (NY).* **58**, 37–53 (2004).

98 Le bourg, E., Thon, B., Légaré, J., Desjardins, B. & Charbonneau, H. Reproductive Life of French-canadians in the 17-18th Centuries: A Search for a Trade-Off Between early Fecundity and Longevity. *Exp. Gerontol.* **28**, 217–232 (1993).

99 Crow, J. F. How much do we know about spontaneous human mutation rates? *Environ. Mol. Mutagen.* **21**, 122–9 (1993).

100 Gavrilov, L. A. & Gavrilova, N. S. human longevity and parental age at conception. *Sex Longev. Sex. Gender, Reprod. Parent.* **7**–31 (2000).

101 Gavrilov, L. A. & Gavrilova, N. S. Predictors of Exceptional Longevity: Effects of Early-Life and Midlife Conditions, and Familial Longevity. *North Am. Actuar. J.* **19**, 174–186 (2015).

102 Cutler, D. M., Lleras-Muney, A. & Vogl, T. Socioeconomic Status and Health: *Dimensions and Mechanisms. Journal of Chemical Information and Modeling (2008).* doi:10.1017/CBO9781107415324.004

103 Jasilionis, D. & Shkolnikov, V. M. Longevity and Education: A Demographic Perspective. *Gerontology* (2015). doi:10.1159/000438901

104 Meij, J. J. et al. Quality-quantity trade-off of human offspring under adverse environmental conditions. *J. Evol. Biol.* **22**, 1014–1023 (2009).

105 Temby, O. F. & Smith, K. R. the Association Between Adult Mortality Risk and Family History of Longevity: the Moderating Effects of Socioeconomic Status. *J. Biosoc. Sci.* **46**, 703–716 (2014).

106 Mazan, R. & Gagnon, A. Familial and Environmental Influences on Longevity in Historical Quebec. *Popul. (english Ed.* **62**, 271–291 (2007).

107 Gagnon, A. *et al.* Is there a trade-off between fertility and longetivity? A comparative study of women from three large historical databases accounting for mortality selection. *Am. J. Hum. Biol.* **21**, 533–540 (2009).

108 Hanson, H. A., Smith, K. R. & Zimmer, Z. Reproductive History and Later-Life Comorbidity Trajectories: A Medicare-Linked Cohort Study From the Utah Population Database. *Demography* **52**, 2021–2049 (2015).

109 Barzilai, N. *et al.* The place of genetics in ageing research. *Nat. Rev. Genet.* **13**, 589–594 (2012).

110 Akker, E. *et al.* Uncompromised 10-year survival of oldest old carrying somatic mutations in DNMT3A and TET2. *Blood* **127**, 1509–1512 (2016).

111. Gierman, H. J. *et al.* Whole-genome sequencing of the world's oldest people. *PLoS One* **9**, 1–10 (2014).