Universiteit
Leiden
The Netherlands

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

Berg, N.M.A. van den

Citation
Berg, N. M. A. van den. (2020, January 22). *Family matters: a genealogical inquiry into the familial component of longevity*. Retrieved from https://hdl.handle.net/1887/83261

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/83261

Note: To cite this publication please use the final published version (if applicable).

Cover Page

Universiteit Leiden

The handle http://hdl.handle.net/1887/83261 holds various files of this Leiden University dissertation.

Author: Berg, N.M.A. van den
Title: Family matters: a genealogical inquiry into the familial component of longevity
Issue Date: 2020-01-22

CHAPTER 1

INTRODUCTION

Progress in genealogic and genetic longevity

During the past 200 years human life expectancy at birth significantly increased in industrialized countries, with record male and female life expectancy increasing from 43/45 years in 1840 to 79/85 years in 2015[1]. At first, the life expectancy increased because of a better understanding of how to care for newborns, resulting in a strong decline in infant and childhood mortality as well as mortality of women giving birth[1–4]. After 1900, the availability of public sanitation, the understanding of hygiene and effective medical healthcare increased significantly and as a result, the group of middle aged (~50 years) and older people (~80 years) also started to live longer[1,5]. Currently, the group of older people represent the fastest growing segment of the population[6]. However, even though people started to live longer, the time they spend in good physical and cognitive health did not increase at an equal rate to the life expectancy, causing many years with chronic disabilities[7–11].

The increase in life expectancy was far too rapid to be caused by genetic change[12] and this is reflected in the low heritability (12-25%)[13–16] of age at death (lifespan) in the population at large. However, the capacity to outlive one's birth cohort members across the entire life course into extremely old age (longevity) clusters strongly within families[17–21] and is known as the familial component of longevity. This familial clustering is illustrated by genealogical studies which showed that parents, siblings[17–19,21–24], and children[19,25–29] of long-lived persons lived longer than first degree relatives of non-long-lived persons or population controls. Members of such long-lived families show a holistic nature of healthy aging, which is illustrated by their life-long decreased chance of dying compared to the general population[21,30]. In addition, they have a lower risk of coronary artery disease, cancer, hypertension and type 2 diabetes[31–33], and if affected, the onset of disease is at later ages and with shorter duration[7,8,34–36]. This phenomenon of late disease onset during a short period of time is known as the "compression of morbidity". Moreover, members of long-lived families enjoy better immune and metabolic health in middle and old age than other individuals of the same age in the general population[8,9,33,37–40]. Thus, understanding the genetic factors influencing longevity may provide novel insights into the biological mechanisms that promote health and minimize disease risk[13,41].

The identification of longevity loci, however, has been challenging and only a handful of genetic variants have been shown to associate with longevity across multiple independent studies[40–47]. The most consistent evidence has been obtained for variants in the APOE and FOXO3A genes[41–46,48], in either genome-wide association studies (GWAS) or candidate gene studies. The limited success of genetic longevity studies in identifying causal variants can be attributed to a number of issues. First, the genetic and environmental heterogeneity[30] of study populations is often large, which makes it difficult to identify specific longevity-associated genes. Second, the focus has been on the discovery of common genetic variants by GWAS approaches whereas rare variants, that can be discovered in whole genome sequencing data, have expectedly bigger functional effects on the trait and may thus be most identifiable. Third, and most relevant of all: the lack of a strict definition for human longevity[13,30], as illustrated by the large variation of longevity definitions[7,13,28,29,32,33,35,40–44,15,45–54,17,19,22,24–27]. This results in a mix of sporadically long-lived cases with those descending from a long-lived family.

Most genealogical studies define longevity as age at death, reflecting a continuous/quantitative trait[17,22,24,26,27,29,33,52] while only a few studies focus on individuals who died at a late age, for example 80, 90, or even 100 years. This is the opposite situation for genetic studies, as only a few studies define longevity as age at death or include mortality information from relatives of study participants[40,47] and, most genetic studies focus on single individuals who are extremely old[43–45,48,53,54]. In other words, they apply an age threshold to define long-lived cases, which may dilute heritable with sporadic cases. The epidemiological transition of the past 200 years, by which the lifespan increased worldwide, has resulted in many long-lived singletons for which heritable factors contributed only partly to their long lifespan. Research into these long-lived singletons showed that a wide range of individual factors, such as socio-economic status, familial resources, genes and, the living environment, associate with lifespan and longevity[55–63,55,57,61,64]. In fact, many of these factors are known to cluster in families[65–70]. However, research into the driving factors of the familial component of longevity is scarce and thus it remains to be explored which factors contribute to the longevity of members of long-lived families.

Currently, many questions regarding familial longevity remain; 1. how can we identify heritable cases for inclusion in genetic longevity studies as many individuals have become long-lived due to factors driving the epidemiological transition?, 2. to what extent is longevity passed on to subsequent generations?, 3. do men and women equally transmit the longevity trait? and, 4. how many family members should be long-lived to represent a familial longevity trait passed on from one generation to the next?. Moreover, it is unknown to what extent the familial component of longevity can be explained by genetic and/or non-genetic (social) factors, connecting to the classic "nature and nurture" dilemma. This thesis focuses on multiple sources of genealogical data to define and explore the heritable longevity trait in large scale family data, taking into account genetic as well as social factors that influence the familial component of longevity so that the results can be used for future studies.

Genealogical data to investigate the familial component of longevity

Such questions, about the familial component and the definition of longevity, can be investigated using individual-level historical genealogical data. Such genealogical data can be obtained through different data sources: 1. civil certificates (birth, marriage, and death certificates), 2. population registers, 3. parish registers, 4. census records, and 5. genealogy websites. *Table 1* explains the characteristics of these different data sources and compares their advantages and disadvantages. The table also shows the types of data sources underlying the databases used for this thesis. These datasets include the Leiden Longevity Study (LLS), the Historical Sample of the Netherlands (HSN), the HSN case/control study, the Utah Population Database (UPDB), and the LINKing System for historical family reconstruction (LINKS). The data sources can be used for life course and family reconstruction. Life course reconstruction refers to the identification process of vital events that happened during the life course of an individual, such as birth, marriage, divorce, moving from one place to another, changes in occupation, and death. Family reconstruction refers to the

identification process of relatives, such as parents, siblings, aunts and uncles, and grandparents. In the next paragraphs a short description will be provided of the different data sources and the databases used in in for this thesis.

Table 1: Overview of data sources that can be used for life course and family reconstruction

Source documents	Life course and family reconstruction method and information	Pros	Cons	Dataset
Civil certificates; birth, marriage, and death certificates	Linking certificates for single individuals and between individuals to reconstruct individual life courses and families	Very high quality information Very good representation of childhood mortality	Potential of linking errors Fragmented life course information	LINKS UPDB
Population registers	Continuous observation of individuals' life courses and information on individuals' relatives are on the cards of individuals	Very high quality information Continuous life course information	Partial coverage of childhood mortality	HSN LLS
Parish registers	Church related continuous observation of individuals' life courses and information on individuals' relatives	Pedigrees can date back more than three centuries	Usually only focus on a small geographical area	UPDB
Census records	Interval (usually 10 years) information about individuals and household members	Available in most countries. Can date back more than two centuries	Long intervals between subsequent censuses, especially for family reconstructions	UPDB HSN
Self-reported genealogies and genealogy websites	Individuals reconstruct their own families. The extent of the individual life course reconstruction depends on the source material backing up the genealogies.	Extensive pedigrees dating back very long Large numbers of pedigrees and individuals	No or limited verified demographic information such as births, deaths, marriages, and profession	UPDB LLS

LLS = Leiden Longevity Study LINKS = LINKing System for historical family reconstruction and HSN = Historical Sample of the Netherlands. Life course reconstruction refers to the reconstruction of a person's life course by, for example following a person from birth to death and observing what life events this person experiences during his/her life. Family reconstruction refers to the reconstruction of family ties, such as identifying a person's parents, siblings, children, grandparents, etc. The "dataset" column refers to the dataset that is used for the analysis during the period of this thesis.

Population registers provide a continuous observation of a person's life course[71–73]. This means that whenever a vital event, such as moving from one place to the other or the birth of a child, happens, this event is registered in the population register. Next to the registration of vital events, population registers also contain identifying information on a person's parents, siblings, and children. As such, population registers are useful to obtain life course information and family ties of a person. Of course, to obtain life course information of the identified relatives, the population register for that specific person should be obtained[73]. In the Netherlands, these registers are officially maintained by the Dutch government and as such, the life course and family information is very accurate. Data from the Dutch population register is available in the Historical Sample of the Netherlands (HSN) and the HSN is supplemented with civil certificate data[72–74].

Civil certificates provide interval life course observations[73], with the main observation moments at birth, marriage and death. Birth, marriage and death certificates all contain reference to the parental names of the person to which the certificate belongs. These names can be used to identify the civil certificates of all relatives of a person. Marriage certificates are most practical to reconstruct families because they contain the parental names of both spouses. The civil certificates contain high quality life course information and using available parental information, family reconstruction is relatively easy. However, complex, common, and ambiguous first and/or last names provide difficulties in identifying the civil certificates of individuals or their parents and may even result in identifying and connecting the wrong certificates. In the Netherlands, family and life course reconstruction based on the civil certificates has been automated using a name linking algorithm[73] and resulted in the LINKing System for historical family reconstruction (LINKS) data[75–77].

Parish registers are often similar to the Dutch population registers regarding the life course and family member information they contain. The main difference is that they are initiated and maintained by a parish church and hence, usually cover a small geographical area. In contrast to the coverage of a small geographical area, census records usually cover entire countries. Census records can provide a snapshot of a person's life course and the relatives of a person, based on the moment the census was conducted. Census records are often used to supplement life course and family reconstruction information based on other sources and moreover, with the increasing availability of data, not only census records can be combined with other sources, but all different data sources can be combined. The HSN for example includes, to a small extent, census and civil certificate information to supplement the population register information. The Utah Population database (UPDB) was initially based on records of the Mormon Church in Utah (US) and now covers the entire Utah population. The UPDB is the most extensive historical genealogical database in the world at this moment[78,79], *https://uofuhealth.utah.edu* and combines information from parish registers, civil certificates, and census information to establish high quality life course and family reconstructions. Additionally, the data are extended with information from medical records, and driver license records[78,79].

Genealogical websites provide an interesting source of family reconstructions as they may cover persons well before the introduction of the civil registry and population registers[15,16]. The pedigrees contained on genealogical websites are usually constructed by hobby-genealogists who are interested in their own family history. For (historical)

genealogical research, the pros are that the pedigrees are structured in a standardized international format (GEDCOM) and that the pedigrees may date back for many years (sometimes to before 1700). A downside of this is that there are no governmental sources before ~1800 to verify demographic and mortality information of pedigree members so that the error rate increases before 1800. Similar to the misreporting of demographic and mortality information, it is also more difficult to verify family relationships in those early years. In addition, the quality of the life course and family reconstructions may vary strongly between families as one hobby-genealogist may try to verify mortality information and family ties with the corresponding civil certificates and population registers, but another may not do this. It is also important to note that there may be a self-selection bias in such data towards non-extinct families. Living people map their own family threes and thus, extinct families may be underrepresented, which may or may not cause issues, depending on the type of research that is done with the data.

Aim and outline of this thesis

Understanding the genetic factors influencing longevity may provide novel insights into the mechanisms that promote health and minimize disease risk[13,41]. Hence, the aim of this thesis is to study the familial component of longevity by first establishing a standardized definition of longevity and subsequently investigating the intergenerational transmission of longevity. Ultimately we aim to establish a genetically enriched group (cases) and a group which represents the general population (controls) that could be included in novel genetic longevity studies. In our analyses we take into account social factors, such as socio-economic status, fertility measurements, familial resources, and living environment, which potentially contribute to the familial component of longevity. We also take the survival of spouses into account and explore the association between longevity and family size.

Data

For our investigations we use different datasets. These include the data available in the Leiden Longevity Study (LLS)[19,21], which is based on self-reported genealogies, supplemented with Dutch population registers provided by the Central Bureau of Genealogy (CBG) and Dutch Personal Records Database (PRD). We further used the LINKS and HSN data, provided by the International Institute for Social History (IISH), and the UPDB provided by the University of Utah.

The LLS was initiated in 2002 to study the genetic determinants of human longevity. Men and women could participate if they were alive and aged ≥89 and ≥91 respectively and had at least one sibling fulfilling the same criteria. From these participants, relatives were identified, including their parents, spouses, siblings, and children. The LLS consists of 421 families and covers 3 generations with a birth cohort range between 1850 and 2019. The LINKS Zeeland (from here: LINKS) data indexing began in 1995. LINKS contains around 700,00 birth, 300,00 marriage, and 600,000 death certificates providing information for around 2,000,000 persons and covering a maximum number of 7 generations. LINKS covers birth cohorts between 1750 and 1920. The UPDB data construction began in the mid-1970s with genealogy records from the archives at the Utah Family History Library and was initially based on the founding members of

the Utah population, their descendants, and then subsequently all individuals living in Utah. The UPDB contains information on more than 11 million individuals and covers a maximum of 17 generations. The UPDB covers birth cohorts between 1750 and 2019. The HSN 2010.01 release is based on a sample of birth certificates and contains complete life course information for 37,137 Dutch individuals (index persons (IPs)) born between 1849 and 1923 (32–34). These 37,137 persons were subsequently identified in the Dutch population registers and followed in the registers throughout their entire life course. The HSN 2010.01 covers a maximum number of 3 generations and covers birth cohorts between 1850 and 1922. The HSN was split into a case (persons who died ≥ 80 years) and a control (persons who died between 40 and 59 years) design for the birth cohorts 1860 - 1875 and for these groups, the third generation was extended with mortality information. Subsequently, two generations were added to the data. The final database covers 57,337 persons from 1,326 five-generational families with birth cohorts ranging between 1850 and 1995. *Table 2* provides an overview of the data sources underlying the datasets that are used for this thesis and *Figure 1* illustrates the birth cohorts included in the data analyses of this thesis.

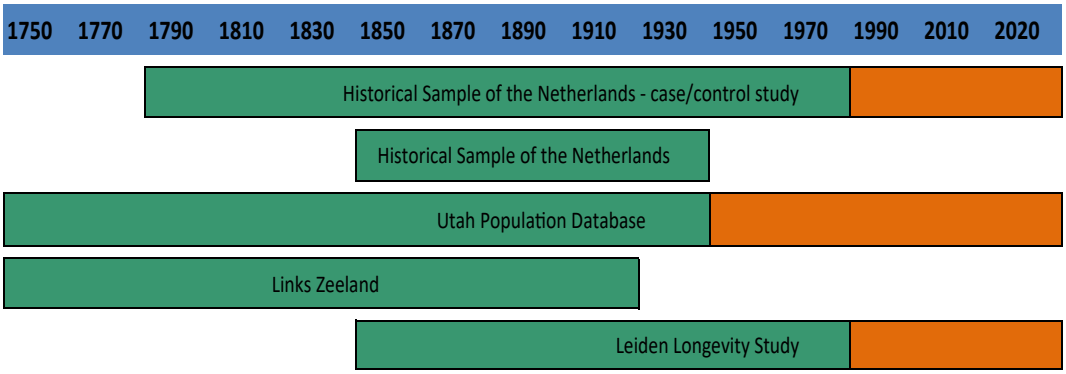

Figure 1: Overview of the databases and study periods used for this thesis.

The green colors in the time-line illustrate the included birth cohorts of all databases used during this thesis. The orange colors illustrate the birth cohorts that were not included in the analyses.

Table 2: Overview of databases that have been used for in this thesis

Dataset	Short description	Type	Birth cohorts	Continuous expanding	Source material
LLS	The LLS was initiated in 2002 and consists of 421 families, covering 3 generations.	Specific population database	First birth cohort is 1850 and the last birth cohort is 2017	No	Population registers and self-reported pedigrees
LINKS Zeeland	The LINKS data indexing began in 1995. The data currently covers over 25 million Dutch vital event records. LINKS Zeeland contains information for around 2,000,000 persons and covering a maximum number of 7 generations.	General population database	Around 1750 and around 1920	Yes	Birth, marriage, and Death certificates
UPDB	The UPDB data construction began in the mid-1970s. The UPDB contains information on more than 11 million individuals and covers a maximum of 17 generations.	General population database	Around 1750 up to 2019	Yes	Church registers, birth, marriage, and death registers, census records, and genealogy website information
HSN	The HSN was initiated around 20 years ago and contains complete life course information for 37,137 Dutch individuals The HSN 2010.01 release covers a maximum number of 3 generations.	General population database	First birth is 1850 and last birth is 1922	Yes	Population registers, birth, marriage, and death registers
HSN case/ control	The HSN case/control study was initiated in 2014 and the final database covers 57,337 persons from 1,326 five-generational families.	General population database	Around 1850 up to 1995	No	Population registers, birth, marriage, and death registers

LLS = Leiden Longevity Study, LINKS = LINKing System for historical family reconstruction and HSN = Historical Sample of the Netherlands. The "dataset" column refers to the dataset that is used for the analysis during the period of this thesis. Specific population refers to a database which reflects a specific study population, such as the LLLS which contains a directly or indirectly selected group of people, general population refers to a database which reflects a broad population, such as the HSN, which reflects the general Dutch population.

Outline

In *chapter 2* we review the relevant literature investigating the familial component of longevity. We focus on heritability studies, studies investigating the transmission of lifespan and longevity as well as lifespan and longevity inheritance patterns. We further discuss important environmental/social factors that affect individual lifespan and longevity or potentially affect the transmission of lifespan and longevity between parents and offspring. In other words, we discuss the factors that associate with individual longevity and the familial component of longevity. We emphasize the difference between lifespan and longevity traits. Lifespan generally refers to the age at death of a person whereas longevity refers to survival into extreme ages, such as 80, 90, 100 years, or an extreme survival percentile such as belonging to the top 5% birth cohort specific survivors. Finally, we discuss a strategy to study familial longevity and to identify a definition of longevity that may best represent the heritable longevity trait.

The LINKS database was recently constructed and before using these data we wanted to validate the quality of the life course and family reconstructions by comparing the LINKS data to the well-established HSN data. Hence, in *Chapter 3* we test the quality of the LINKS and HSN data by comparing life course information and family reconstruction of ~400 persons born in Zeeland who could be identified in both the HSN and LINKS data. We focus on overlap and differences in demographic information, such age at death and age at first childbirth, and family information, such as the number of identified siblings and children. Finally, we expect that some of the differences in demographic and family indicators, such as the number of children, between the two databases can be explained by how migration is represented in the two databases. In LINKS, migration during the life course to another province or country is not included. Hence, we test for migration differences as an explaining factor for discrepancies between demographic and family indicators in the HSN and LINKS.

After confirming the quality of the life span and family reconstructions in the LINKS data, we initiated a collaboration with prof. Ken Smith, who provided access to the UPDB, to investigate the definition and subsequently, the familial component of longevity. Thus, in *Chapter 4* we use three-generational demographic and mortality data from two large datasets, UPDB (US) and LINKS (Netherlands). We focus on 20,360 families who are unselected for mortality. The data contains 20,360 index persons, their parents (N=40,72), siblings (N=108,122), spouses (N=22,018), and children (N=123,599), comprising 314,819 individuals in total. We use these data to investigate which survival percentile best isolates the heritable component of longevity and we subsequently determine the importance of long-lived family members for case selection so that those insights can be used in genetic studies to identify novel longevity loci. In the analyses we include social and environmental factors, such as socio-economic status, religious denomination, number of children, birth order, and birth cohort, that may explain the intergenerational transmission of longevity. Moreover, we explore the survival of spouses marrying into longevity enriched families as an indicator for shared resources, lifestyles, and potentially socio-economic status during middle and late-life as explaining factors for the familial component of longevity.

We apply the novel survival percentile threshold based longevity definition in *Chapter 5*, where we focus on 2 generations from the Leiden Longevity Study, containing 944 long-lived siblings (participants), their parents (N=842), siblings (N=2302), and spouses (N=809) from 421 LLS families. We define longevity as belonging to the top 1% survivors of their birth cohort to investigate 1. a potential sex-specific inheritance pattern of longevity, 2. a potential survival advantage of long-lived sibships as compared to long-lived singletons and 3. whether the parents of these siblings had a life-long sustained survival advantage. Similar to *Chapter 4*, we include the spouses of the 944 LLS participants to explore mid

and late-life shared environmental/social factors contributing to the familial component of longevity.

Following-up on the longevity definition as established in chapter 4, we focus on establishing the proportion of ancestral blood relatives that should be long-lived (at least belonging to the top 10% survivors of their birth cohort) in order to observe a survival advantage in their descendants and subsequently define cases with the heritable longevity trait for inclusion in genetic studies. In *Chapter 6*, we therefore describe and explore the HSN case/control data, which is specifically compiled to cover 5 generations and connect extinct family members to their living descendants. Data construction started with 1326 families form the original HSN and was extended by acquiring population registers and population register information from the CBG and the PRD respectively. Moreover, we were granted permission by the Dutch government to obtain the current addresses of all living descendants. We use the data to investigate if longevity is transmitted for multiple generations and if the longevity effect dwindles over generations by comparing long-lived cases (died ≥ 80 years) and their descendants to population resembling controls (died between 40 and 59 years) and their offspring. Furthermore, we establish how many family members should be long-lived in order to avoid phenocopies and subsequently investigate how often long-lived parents from a long-lived family pass on their longevity to their children compared to long-lived parents from general population families. Equal to LINKS and the UPDB analyses, here we include spouses over multiple generations to explore mid and late-life environmental effects. Finally, we utilize the insights to identify a novel case and control group for future genetic and social studies into the heritable longevity trait.

References

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

2 Fogel, R. W. Aging, *Health and Public Policy.* **30**, (2004).

3 Fogel, R. W. & Costa, D. L. A Theory of Technophysio Evolution, With Some Implications for Forecasting Population, Health Care Costs, and Pension Costs. *Demography* **34**, 49 (1997).

4 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).

5 Vaupel, J. W. Biodemographic Trajectories of Longevity. *Science* (80-.). **280**, 855–860 (1998).

6 Kontis, V. *et al.* Future life expectancy in 35 industrialised countries: projections with a Bayesian model ensemble. *Lancet* **389**, 1323–1335 (2017).

7 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).

8 Christensen, K. *et al.* Physical and cognitive functioning of people older than 90 years: a comparison of two Danish cohorts born 10 years apart. *Lancet* **382**, 1507–1513 (2013).

9 Zeng, Y., Feng, Q., Hesketh, T., Christensen, K. & Vaupel, J. W. Survival, disabilities in activities of daily living, and physical and cognitive functioning among the oldest-old in China: a cohort study. *Lancet* **389**, 1619–1629 (2017).

10 Pignolo, R. J. Exceptional Human Longevity. *Mayo Clin Proc* **94**, 1–15 (2018).

11 Crimmins, E. M. Lifespan and Healthspan: Past, Present, and Promise. *Gerontologist* **55**, 901–911 (2015).

12 Burger, O., Baudisch, A. & Vaupel, J. W. Human mortality improvement in evolutionary context. *Proc. Natl. Acad. Sci.* **109**, 18210–18214 (2012).

13 van den Berg, N., Beekman, M., Smith, K. R., Janssens, A. & Slagboom, P. E. Historical demography and longevity genetics: Back to the future. *Ageing Res. Rev.* **38**, 28–39 (2017).

14 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).

15 Kaplanis, J. *et al.* Quantitative analysis of population-scale family trees with millions of relatives. *Science* **360**, 171–175 (2018).

16 Ruby, J. G. *et al.* Estimates of the Heritability of Human Longevity Are Substantially Inflated due to Assortative Mating. *Genetics* **210**, 1109–1124 (2018).

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

18 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).

19 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).

20. 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).

21 Berg, N. van den *et al.* Longevity Around the Turn of the 20th Century: Life-Long Sustained Survival Advantage for Parents of Today's Nonagenarians. *Journals Gerontol. Ser. A* **73**, 1295–1302 (2018).

22 Willcox, B. J., Willcox, D. C., He, Q., Curb, J. D. & Suzuki, M. Siblings of Okinawan centenarians share lifelong mortality advantages. *J. Gerontol. A. Biol. Sci. Med. Sci.* **61**, 345–54 (2006).

23 Perls, T. T., Bubrick, E., Wager, C. G., Vijg, J. & Kruglyak, L. Siblings of centenarians live longer. *Lancet* **351**, 1560 (1998).

24 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).

25 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).

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

27 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).

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

29 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).

30 van den Berg, N. *et al.* Longevity defined as top 10% survivors and beyond is transmitted as a quantitative genetic trait. *Nat. Commun.* **10**, 35 (2019).

31 Westendorp, R. G. J. et al. Nonagenarian Siblings and Their Offspring Display Lower Risk of Mortality and Morbidity than Sporadic Nonagenarians: The Leiden Longevity Study. *J. Am. Geriatr. Soc.* **57**, 1634–1637 (2009).

32 Terry, D. F. *et al.* Lower All-Cause, Cardiovascular, and Cancer Mortality in Centenarians' Offspring. *J. Am. Geriatr. Soc.* **52**, 2074–2076 (2004).

33 Dutta, A. et al. Longer lived parents: Protective associations with cancer incidence and overall mortality. *Journals Gerontol. - Ser. A Biol. Sci. Med. Sci.* **68**, 1409–1418 (2013).

34 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).

35 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).

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

37 Vaarhorst, A. A. M. *et al.* Lipid metabolism in long-lived families: the Leiden Longevity Study. *Age (Omaha).* **33**, 219–227 (2011).

38 Wijsman, C. A. *et al.* Familial longevity is marked by enhanced insulin sensitivity. *Aging Cell* **10**, 114–121 (2011).

39 Rozing, M. P. *et al.* Low serum free triiodothyronine levels mark familial longevity: The leiden longevity study. *Journals Gerontol. - Ser. A Biol. Sci. Med. Sci.* **65A**, 365–368 (2010).

40 Joshi, P. K. *et al.* Genome-wide meta-analysis associates HLA-DQA1/DRB1 and LPA and lifestyle factors with human longevity. *Nat. Commun.* **8**, 1–13 (2017).

41 Eline Slagboom, P., van den Berg, N. & Deelen, J. Phenome and genome based studies into human ageing and longevity: An overview. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1864**, 2742–2751 (2018).

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

43 Deelen, J. *et al.* Genome-wide association meta-analysis of human longevity identifies a novel locus conferring survival beyond 90 years of age. *Hum. Mol. Genet.* **23**, 4420–4432 (2014).

44 Broer, L. *et al.* GWAS of longevity in CHARGE consortium confirms APOE and FOXO3 candidacy. *Journals Gerontol. - Ser. A Biol. Sci. Med. Sci.* **70**, 110–118 (2015).

45 Willcox, B. J. *et al.* FOXO3A genotype is strongly associated with human longevity. *Proc. Natl. Acad. Sci.* **105**, 13987–13992 (2008).

46 Flachsbart, F. *et al.* Association of FOXO3A variation with human longevity confirmed in German centenarians. *Proc. Natl. Acad. Sci.* 106, 2700–2705 (2009).

47 Pilling, L. C. *et al.* Human longevity: 25 genetic loci associated in 389,166 UK biobank participants. *Aging (Albany. NY).* **9**, 2504–2520 (2017).

48 Sebastiani, P. *et al.* Four Genome-Wide Association Studies Identify New Extreme Longevity Variants. *J. Gerontol. A. Biol. Sci. Med. Sci.* **72**, 1453–1464 (2017).

49 Sebastiani, P. *et al.* A Family Longevity Selection Score: Ranking Sibships by Their Longevity, Size, and Availability for Study. *Am. J. Epidemiol.* **170**, 1555–1562 (2009).

50 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).

51 Boyle, E. A., Li, Y. I. & Pritchard, J. K. An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell* **169**, 1177–1186 (2017).

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

53 Flachsbart, F. *et al.* Immunochip analysis identifies association of the RAD50/IL13 region with human longevity. *Aging Cell* **15**, 585–588 (2016).

54 Zeng, Y. *et al.* Novel loci and pathways significantly associated with longevity. *Sci. Rep.* **6**, 21243 (2016).

55 Sebastiani, P. *et al.* Families Enriched for Exceptional Longevity also have Increased Health-Span: Findings from the Long Life Family Study. *Front. Public Heal.* **1**, 1–9 (2013).

56 Ailshire, J. A., Beltrán-Sánchez, H. & Crimmins, E. M. Social Characteristics and Health Status of Exceptionally Long-Lived Americans in the Health and Retirement Study. *J. Am. Geriatr. Soc.* **59**, 2241–2248 (2011).

57 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).

58 Jopp, D. S., Boerner, K., Ribeiro, O. & Rott, C. Life at Age 100: An International Research Agenda for Centenarian Studies. *J. Aging Soc. Policy* **28**, 133–147 (2016).

59 Boerner, K., Jopp, D. S., Park, M.-K. S. & Rott, C. Whom Do Centenarians Rely on for Support? Findings From the Second Heidelberg Centenarian Study. *J. Aging Soc. Policy* **28**, 165–186 (2016).

60 Seeman, T. E. *et al.* Behavioral and Psychosocial Predictors of Physical Performance: MacArthur Studies of Successful Aging. *Journals Gerontol. Ser. A Biol. Sci. Med. Sci.* **50A**, M177–M183 (1995).

61 Andersen, S. L. *et al.* Personality Factors in the Long Life Family Study. *Journals Gerontol. Ser. B* **68**, 739–749 (2013).

62 Yang, Y. C. *et al.* Social relationships and physiological determinants of longevity across the human life span. *Proc. Natl. Acad. Sci.* **113**, 578–583 (2016).

63 Hensley, B. *et al.* Life Events and Personality Predicting Loneliness Among Centenarians: Findings From the Georgia Centenarian Study. *J. Psychol.* **146**, 173–188 (2012).

64 Yashin, A. I. *et al.* "Predicting" parental longevity from offspring endophenotypes: Data from the Long Life Family Study (LLFS). *Mech. Ageing Dev.* **131**, 215–222 (2010).

65 Heath, A. C. *et al.* Education policy and the heritability of educational attainment. *Nature* **314**, 734–736 (1985).

66 Spera, C., Wentzel, K. R. & Matto, H. C. Parental Aspirations for Their Children's Educational Attainment: Relations to Ethnicity, Parental Education, Children's Academic Performance, and Parental Perceptions of School Climate. *J. Youth Adolesc.* **38**, 1140–1152 (2009).

67 Eccles, J. S. Influences of parents' education on their children's educational attainments: the role of parent and child perceptions. *London Rev. Educ.* **3**, 191–204 (2005).

68 Edvardsson, K. *et al.* Giving offspring a healthy start: parents' experiences of health promotion and lifestyle change during pregnancy and early parenthood. *BMC Public Health* **11**, 936 (2011).

69 Wickrama, K. A. S., Conger, R. D., Wallace, L. E. & Elder, G. H. The Intergenerational Transmission of Health-Risk Behaviors: Adolescent Lifestyles and Gender Moderating Effects. J. Health Soc. Behav. **40**, 258 (1999).

70 Thibault, H., Contrand, B., Saubusse, E., Baine, M. & Maurice-Tison, S. Risk factors for overweight and obesity in French adolescents: Physical activity, sedentary behavior and parental characteristics. *Nutrition* **26**, 192–200 (2010).

71 Mandemakers, K. Building life courses datasets from population registers by the historical sample of the Netherlands. in *History and Computing* **14**, 87–107 (2002).

72 Mandemakers, K. *https://socialhistory.org/en/hsn/hsn-releases. HSN 2010.01 release* (2010).

73 van den Berg, N. *et al.* Families in Comparison: An individual-level comparison of life course and family reconstructions between population and vital event registers. *SocArXiv* (2018).

74 Mandemakers, K. Historical sample of the Netherlands. In P. K. Hall, R. McCaa, & G. Thorvaldsen (Eds.). in *Handbook of International Historical Microdata for Population Research* 149–177 (2000).

75 Mandemakers, K. & Laan, F. LINKS dataset Genes Germs and Resources. WieWasWie Zeeland. Civil Certificates. (2017).

76 *https://socialhistory.org/en/hsn/linking-system-historical-family-reconstruction-links*

77 Mandemakers, K. *LINKing System for historical family reconstruction.* (2007).

78 *https://healthcare.utah.edu/huntsmancancerinstitute/research/updb/data/family-records.php*

79 Bean, L. L., May, D. L. & Skolnick, M. The Mormon Historical Demography Project. *Hist. Methods A J. Quant. Interdiscip. Hist.* **11**, 45–53 (1978).