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

Berg, N.M.A. van den

CHAPTER 7

GENERAL DISCUSSION

The complexity of longevity as a multifactorial trait

Identifying longevity loci is important because these loci likely represent key mechanisms of a life-long decreased mortality[1,2], decreased morbidity[3–5] and compression of morbidity towards the end of life[6–8]. However, the identification of longevity loci has been challenging and only a handful of genetic variants have been shown to associate with longevity across multiple independent studies[9–16]. In fact, genome-wide linkage and association studies identified only a few robust loci promoting longevity[9–17]. The most compelling evidence was obtained for alleles in the APOE and FOXO3A genes as they have been consistently identified with either genome-wide association studies (GWAS) or candidate gene studies[10–15,18].

One of the main reasons for the limited success of genetic longevity studies[12–14,18–20] is the uncertainty in defining the heritable longevity trait itself[2,21]. Given the increased life expectancy of the past 200 years due to non-genetic/social factors[22,23] (improved hygiene, nutrition and medication) there are likely many phenocopies among the long-lived cases selected for our genetic studies. In this thesis we showed that the solution may lie in the familial clustering of longevity and that the inclusion of persons with the heritable longevity trait (the persons with a high familial clustering of longevity) may provide a fruitful basis for future longevity research.

It is important to consider the strong increase in human life expectancy when investigating familial longevity using multigenerational genealogical data, as multiple generations of long-lived individuals experienced significant developments in the knowledge of good hygiene, healthy behavior and health care over the past 200 years6,[22–34] (this is known as the epidemiological transition).

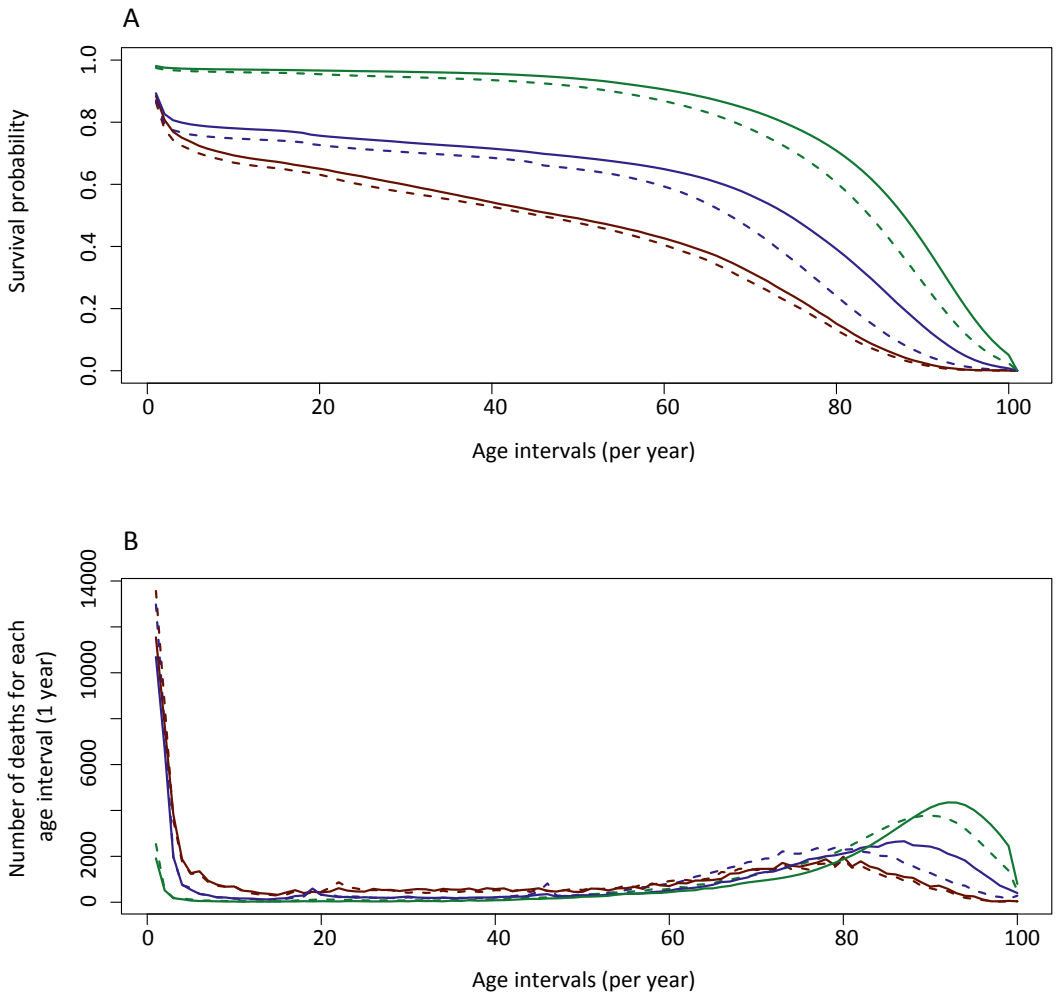

Figure 1: Increase in human lifespan between 1850 and 1950.

Panel A shows the average increase in human survival from 1850 onward in the Netherlands. Panel B shows the reduction in infant and childhood mortality and the subsequent increase in mid and later life survival. The brown lines represent the 1850 birth cohort, the blue lines represent the 1900 birth cohort, the green lines represent the 1950. The 1850 and 1900 cohorts are based on the CBS cohort lifetables and the 1950 cohort is based on the prognostic CBS cohort lifetables. The straight lines represent females and the dotted lines represent males.

Due to better knowledge of hygiene and increased food availability people were able to better cope with malnutrition of children, infections with cholera, diarrheal diseases and tuberculosis and as a result, childhood mortality decreased significantly. Next, the increased understanding of (preventive) medicine and medical health care, further improvements in hygiene, and the introduction of public sanitation caused a reduction in mid and later-life mortality[6,22–34]. Besides these developments, multiple generations of long-lived individuals experienced various important events, such as the potato blight of 1847, various outbreaks of smallpox and measles, the first world war, the almost coinciding Spanish flue of 1918, and the second world war[25,35–38]. The increase in human lifespan is illustrated in *Figure 1*. Panel A visualizes the average increase in human survival from 1850 onward in the Netherlands. It also clearly shows that the survival difference between men and

women increased over time (from a two year difference to a 6 year difference) which is speculated to be caused by more unhealthy behavior among men[39–41]. Panel B illustrates the reduction in infant and childhood mortality and the subsequent increase in mid and later life survival.

As all these changes caused a rapid increase in the human life expectancy, it can be expected that many individual non-genetic/social factors contribute to the human survival and longevity. In fact, contemporary research identified many individual factors that associate with lifespan and longevity. Among others, Socio-economic status in terms of income and occupation, educational attainment, sibship size, birth order, but also personality traits, and susceptibility to diseases such as cardio-vascular disease, cancer, and Alzheimer's disease are associated to individual survival and longevity[42–50,42,44,48,51]. These factors are likely to cluster in families as parents and their children share available resources and live in the same physical environment. In fact, research showed that children may inherit personality traits from their parents, either through genetic predispositions or socialization, but also for example their parents' socio-economic status[52–57]. It is however unknown to what extent these factors contribute to the clustering of longevity in members of long-lived families.

The aim of this thesis was to study the familial component of longevity by first establishing a standardized definition of heritable longevity and subsequently investigating its intergenerational transmission characteristics, its life course influence on survival and the interrelation with other, environmental or familial factors. We ultimately aimed to establish a genetically enriched group (cases) and a group which represents the general population (controls) who could be included in future genetic longevity studies.

We used unselected three-generational demographic and mortality data from the Utah Population Database (UPDB, US) and the LINKing System for historical family reconstruction (LINKS, Netherlands) to investigate which survival percentile best isolates the genetic component of longevity. We subsequently determined 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. We applied our knowledge about survival percentiles to the Leiden Longevity Study (LLS) where we investigated 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 sibships had a life-long sustained survival advantage. Next, we used the Historical Sample of the Netherlands (HSN) case/control study to establish how many family members of a potential long-lived case should be long-lived in order to avoid phenocopies. Finally, we utilized the insights to identify in the HSN data a novel case and control group for future genetic studies in which we connect deceased ancestors to living family members and compared these to a group of sporadically long-lived persons. Besides these main aims we discuss the survival of spouses marrying into long-lived families, the effect of environmental/social factors on individual as well as familial longevity, the interplay between longevity and family size, and sex specific transmission patterns observed in some of the data.

Discussion of our findings

Main results

In **chapter 2** we reviewed the relevant studies investigating the familial component of longevity. We focused on heritability studies, studies investigating the transmission of lifespan and longevity as well as lifespan and longevity inheritance patterns. We further discussed important environmental/social covariates that affect individual lifespan and longevity and/or potentially affect the transmission of lifespan and longevity between parents and offspring. We emphasized the importance of distinguishing between lifespan and longevity because currently, longevity is often confused with lifespan. Lifespan generally refers to the age at death of a person whereas longevity refers to survival into extreme ages beyond an arbitrarily chosen threshold, such as 80, 90, 100 years, or an extreme survival percentile such as belonging to the top 10%, 5%, of 1% birth cohort specific survivors. We concluded that heritability estimates for lifespan vary slightly but consistently indicate that 12-25% of the variance in lifespan is due to additive genetic effects. The large impact of environmental factors on the average lifespan likely surpasses that of genetic ones and a large range of factors determine early death. In contrast to the number of studies into the heritability of lifespan, studies into the heritability of longevity are scarce and report inconsistent heritability results. Moreover, there are indications that the heritability increases with a more strict cutoff of lifespan towards more extreme ages of survival[1,58–66]. Studies focusing on the transmission pattern of lifespan and longevity are both inconsistent but, for lifespan studies there are indications of a female transmission pattern. We conclude that environmental/social factors, such as socio-economic status, sibship size, maternal age at first and last birth and birth order should be taken into account when investigating the familial component of lifespan and potentially also for longevity. We further conclude that novel research is needed to estimate the heritability of longevity and establish a longevity transmission pattern. Because there is no consensus of how longevity should be defined, we first discuss a strategy to identify a definition of longevity that best represents the heritable component of the trait.

To investigate the familial component of longevity in large genealogical data in the absence of study related selections such as the inclusion of alive persons who survived into old age, we constructed the LINKS data together with the International Institute of Social History (IISH) and the Radboud University (RU). Because of the novel character of the LINKS data, we first set out to validate the life course and family reconstruction quality by comparing the LINKS data with the already existing HSN data. Thus, in **chapter 3**, we compared indicators of fertility, marriage, mortality, and measurements of occupational status of ~400 individuals identified in both databases and concluded that life course and family reconstructions in the HSN and LINKS reflect each other well. As we expected, LINKS provides more complete family information on siblings and parents, whereas the HSN provides more complete life course information, especially for individuals who migrated out of Zeeland. We also observed that the number of children was very similar between the ~400 persons identified in both the HSN an LINKS. This coincides with the very complete life course information in the HSN which accurately captures the births of children. We conclude that life course and family reconstructions based on linked, fragmented observations, such as in LINKS, on individuals constitute a reliable alternative to such reconstructions based on continuous observations from population or parish registers.

After verifying the quality of the LINKS data, we continued to investigate the definition of heritable longevity and the familial clustering of longevity using the LINKS data and the Utah Population Database (UPDB), which combined represent the largest genealogical database with verified (mortality) information in the world. In **chapter 4** we used three-generational mortality data from the UPDB and LINKS, and studied 20,360 families who were unselected for mortality. We focused on 20,360 index persons, their parents (N=40,72), siblings (N=108,122), spouses (N=22,018), and children (N=123,599), comprising a total of 314,819 individuals. We investigated which survival percentile best isolates the heritable component of longevity and we subsequently determined 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. We further studied the non-genetic/social 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 explored 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. In addition, it is important to note that we indirectly investigated social and living environmental influences on the familial component of longevity by comparing Utah and Zeeland. Utah and Zeeland distinctly differed in their physical environment, living conditions, and subsequent mortality patterns. Our analyses provided strong evidence that longevity is transmitted as a quantitative genetic trait among survivors up to the top 10% of their birth cohort. We subsequently showed a survival advantage, mounting to 31%, for individuals with top 10% surviving first and second-degree relatives in both databases and across generations, even in the presence of non-long-lived parents. Further results showed that, among others, socio-economic status, sibship size, birth order did not affect the association between parental or sibling longevity and the survival of the index persons. Some factors, such as socio-economic status, birth year, and religious denomination did affect the individual survival of the index person, but as mentioned, independently of the parental and sibling effects. No evidence was observed that spouses marrying into a longevity enriched family also showed a survival benefit. This will be discussed in more detail further on. Interestingly, the Hazard Ratios, reflecting the survival benefit of index persons with 1 or 2 compared to 0 long-lived parents or siblings were remarkably similar between the UPDB and LINKS. This similarity provides a strong indication that the familial component of longevity is very limitedly affected by effects of the physical environment and for example migration patterns. Finally, to guide future genetic studies, we suggest to base case selection on top 10% survivors of their birth cohort with equally long-lived first and second-degree family members.

In **chapter 5**, we applied the new survival percentile threshold based longevity definition to the Leiden Longevity Study (LLS) where we studied the 944 participating long-lived siblings and their relatives 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. Family longevity scores were estimated to explore whether human longevity is transmitted preferentially through the maternal or paternal line. Standardized mortality ratio's (SMRs) were estimated to investigate whether long-lived siblings have a survival advantage compared to long-lived singletons and we investigated if parents of long-lived siblings harbor a life-long sustained survival advantage compared to the general Dutch population by estimating lifetime SMRs (L-SMRs). We observed that sibships with long-lived mothers and non-long-lived fathers had 0.41 (P=0.024) less observed deaths than sibships with long-lived fathers and non-long-lived mothers and 0.48 (P=0.008) less observed deaths than sibships with both parents non-long-

lived. Participants had 18.6% less deaths compared to matched singletons and parents had a life-long sustained survival advantage (L-SMR=0.510 and 0.688). In conclusion, genetic longevity studies may incorporate the testing of a maternal transmission pattern (further discussed later on) and potential genes involved appeared to beneficially influence the entire life-course of individuals.

In **chapter 4** we addressed the issue of the uncertainty in defining the longevity trait itself and observed that the survival percentile threshold that best reflects the genetic component of longevity is at the top 10% survivors of their birth cohort and beyond. Moreover, we investigated the familial component of longevity and observed that the survival advantage of family members increased with each additional long-lived family member. In **chapter 6** we followed-up on the longevity definition as established in chapter 4. We investigated if longevity is transmitted for multiple generations and whether the longevity effect diminishes over generations. We did this by comparing long-lived cases (died ≥ 80 years) and their descendants to population resembling controls (died between 40 and 59 years) and their descendants. Furthermore, we developed the Longevity Relatives Count (LRC) score to establish how many family members should be long-lived in order to avoid phenocopies. We subsequently investigated 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.

Our analyses included 37,825 persons from 1,326 three-generational families in the HSN case/control study. The analyses in the HSN case/control dataset provide strong evidence that longevity is transmitted for at least 2 subsequent generations if at least 20% of all relatives are long-lived, but preferably 30%. Moreover, the family based cases seem to be at least partially genetically enriched for longevity, as birth year, sibship size, and sex did not affect the transmission of longevity. The evidence for genetic enrichment is strengthened by the fact that their spouses resembled the family based controls as well as the general population in their average survival. Moreover, other studies, as outlined in **chapter 4**, did not obtain evidence that other non-genetic factors, such as religion and socio-economic status could explain the familial component of longevity. Finally, 27% of the F3 descendants showed a survival pattern similar to the general population even though they had at least one long-lived parent. Hence the parents of theses 27% F3 descendants were sporadically long-lived as they did not transmit their longevity. In summary, to select individuals that are enriched for the heritable longevity trait, case should be selected on the basis of being long-lived themselves and having at least 30% long-lived ancestors.

We now have a much more clear understanding about the familial component of longevity. Most importantly, we know that longevity is transmitted as a quantitative genetic trait among survivors up to the top 10% of their birth cohort as long-lived blood relatives independently and additively contribute to the survival advantage of index persons. Long-lived study participants with a family history of longevity have a lifelong sustained survival advantage and their spouses seem to resemble the general population. In line with other studies, we showed that the association between parental longevity and the survival of their offspring is not affected by non-genetic/social factors such as socio-economic status and sibship size. In addition, by using the LRC sore we determined that at least 30% of an individual's relatives should be within the top 10% survivors of their birth cohort before a survival advantage is observed. Moreover, individuals without long-lived relatives represent the general population and may thus be considered as phenocopies even when they do become long-lived. These results provide novel opportunities for future research into longevity.

Secular trends in longevity research

As mentioned earlier, when investigating multiple generations of long-lived persons it is important to take into account that these persons experienced important developments in knowledge about hygiene, healthy lifestyle, medical health care and related technological advancements. It is also important to acknowledge that there were epidemic periods and periods of war which affected mortality. In addition, the life expectancy difference between men and women increased over the last 200 years. In demographic research, these changes over time are known as secular (mortality) trends. A consequence of these secular trends, for longevity research is for example that being 90 years old nowadays is not nearly as special as it was around 1800. Hence, to investigate familial longevity with data spanning more than 200 years, it is important to take these secular trends into account.

An approach to incorporate these secular trends in statistical models is to standardize the measurements for mortality, e.g. age at death, of study participants to that of their birth cohort members who experienced the same developments and epidemic hazards during their life. We used cohort lifetables to calculate birth cohort and sex specific survival percentiles (for example, belonging to the top 10%, 5%, or 1% survivors) so that the mortality of a person is measured relative to his or her birth cohort members who experienced the same secular trends. This approach requires a reference population on which the lifetables are based and has as a main advantage that the survival percentiles are calculated in exactly the same way every time. This ensures a fair comparison between study participants of different birth cohorts and different study populations. For our research, we used both Dutch and Swedish lifetables which are consistent with the lifetables of multiple industrialized societies[67]. Some alternative approaches have been developed[1,59,68–70] which come down to regressing out study population specific environmental effects (that reflect secular trends as good as possible, for example a person's birth cohort) associated with individual mortality and analyzing the residuals as a measure of mortality. There are some clear drawbacks of these study population specific residual methods compared to the lifetable standardization method we used. The residual methods account only for the environmental factors that are included in the statistical model and are thus dependent on what is measured in a specific study and the arbitrary choices of a researcher on how to model these factors. Moreover, because the included environmental factors are specific for a study and depend on what is measured in a study, comparative research between different populations is difficult. The lifetable method provides an advantage because by comparing to birth cohort members it is possible to adjust for all secular trends over time that are shared by members of a specific birth cohort. This includes many environmental factors that are usually not observed in specific studies, for example, relating to the improved living conditions or health care system over time. The lifetable model also allows for fair comparisons across study populations, as illustrated in **chapter 4**. One drawback of comparing different study populations, however, is that the environment factors accounted for in the reference population are unlikely exactly the same in the different study populations. Taking the pros and cons of the different methods into account, we prefer the lifetable method over the residual method.

Spouses seem to resemble the general population

In **chapter 3, 4, and 5** we investigated the survival of spouses marrying with long-lived persons or into longevity enriched families. All our results indicated that the spouses had a survival pattern equal to the general population, except for the spouses investigated in the LINKS study. The LINKS data contains an overrepresentation of persons who stayed in

Zeeland because we could not identify those who migrated to another province or abroad, as is described in **chapter 3**. Moreover, Zeeland had a low number of outmigration[71] which is described in **chapter 4**. The combination between the overrepresentation of stayers and the general pattern of low outmigration potentially caused high levels of relatedness among study participants in the LINKS data and in fact, unpublished results based on the LINKS data confirm this. Hence, it is likely that in Zeeland, spouses and long-lived persons were often (distantly) related to each other and thus shared some of the genetic component contributing to longevity. As a result, the observed survival benefit of spouses marrying into a longevity enriched family in LINKS, is likely caused by their relatedness instead of a shared mid and later-life lifestyle. Other studies showed diverse results regarding a possible survival benefit for spouses of long-lived persons[1,61,65,72,73]. In the Long Life Family study, Pedersen et al. (2017) identified a survival benefit for spouses of longevous siblings. The authors compared the spouses to sex and birth cohort matched controls and suggest assortative mating as an explanation for the observed survival benefit of the spouses[65]. A Quebec study, focused on the spouses of 806 centenarians, also reported a survival benefit[72] and a study of Southern Italy demonstrated that male nonagenarians outlived their spouses, whereas this was not the case for female nonagenarians[73]. In the future, more research is necessary to find out whether spouses marrying to long-lived persons or into a longevity enriched family were already predisposed with a survival advantage, gained a survival advantage, or in fact, resemble the general population.

Sex specific inheritance pattern of longevity

We investigated sex specific longevity transmission effects. Such effects can be divided into 1. Longevity is transmitted stronger via the mother than the father or vice versa, 2. sons or daughters who are more susceptible to parental (either the mother or the father) transmission of longevity, or 3. a combination between the two, for example mothers could transmit longevity more frequently to daughters than to sons. In **chapter 5**, when investigating the LLS, we observed evidence for a maternal transmission pattern with equal distribution to sons and daughters. We discuss two possible mechanisms that may cause this maternal transmission pattern: 1. the transmission of a beneficial genetic component via mitochondrial inheritance, as the mitochondrial DNA is only inherited via mothers and mitochondria play a vital role in many metabolic processes which have been associated to aging and longevity, 2. It might be possible that these long-lived mothers had babies with a high birth weight. In the end of 1800 a high birth weight provided a significant advantage in coping with the often harsh environmental/social circumstances, such as food scarcity and epidemics.

We however did not find evidence for a maternal transmission pattern or a higher level of susceptibility for sons or daughters in the UPDB, LINKS (**chapter 3**), and the HSN case/control study (**chapter 6**). Previous studies focused mainly on a sex specific inheritance pattern of lifespan (**chapter 2**) and had mixed conclusions[74–76]. Several explanations for the mixed results between the LLS and the UPDB, LINKS, and HSN case/control study can be possible. A first explanation concerns the specific time period that we observed in the LLS (1875 - 1941, *Figure 1* of the introduction). The period between 1875 - 1941 was characterized by repeated epidemics such as the smallpox, the measles, the potato blight, and the Spanish flue, creating strong infant and childhood mortality peaks[25,35–38]. This first explanation might be strengthened by the fact that the LLS sample was not random and it might be that many persons were born in places with a high epidemic impact, revealing the maternal transmission effects. Alternatively, the measurement that we used to test for the sex specific transmission pattern in the LLS was binary. This meant that a person was defined as long-lived (top 1% survivor) or alternatively as non-long-lived

(not top 1% survivor). By defining our groups like this, we ignored the survival percentile distance between parents. Consider for example a set of children in family A. If their mother belonged to the top 1% survivors she was considered long-lived. If their father belonged to the top 2% survivors (and thus not to the top 1%) he was not considered long-lived. A father within the 98th percentile of his birth cohort is still very much able to transmit his longevity as we observed in **chapter 4**. The distance between both parents is in this example 1 percentage point. Now consider a set of children in family B. Here their mother also belonged to the top 1% survivors of her birth cohort. However, their father belonged to the top 60% survivors of his birth cohort. The distance here is 59 percentage points. It is not too difficult to imagine that the described distances between fathers and mothers could be unequally distributed in the LLS and may thus have driven the observed maternal transmission pattern. This example points to a more general methodological issue in the literature when analyzing sex specific transmission patterns in longevity, which is generally defined as a binary trait. For future research it would be interesting to use methodology that can incorporate the continuous distribution between the binary longevity cutoffs of parents. In fact, for the LLS we did some preliminary analyses by defining non-long-lived as belonging to the bottom 85% survivors. So far this has not lead to different conclusions about the observed maternal inheritance pattern in the LLS.

In **chapter 2** we addressed fertility measurements, such as the number of children a mother has, or maternal age at first and last birth. We used these measures to investigate their interplay with the intergenerational transmission of longevity. In all our data; the LINKS, UPDB, LLS, and HSN we observed that persons from a smaller sibship had a lower hazard of dying than individuals from a larger sibship. Nonetheless, sibship size did not affect the transmission of longevity to a subsequent generation. In addition, in the LLS we observed a better survival and smaller sibship sizes for children with a long-lived mother and a non-long-lived father than the other way around or for children without any long-lived parents. This implied that long-lived mothers had less children and that these children also lived longer. Replication of these findings in data from less selected study populations would be interesting. We however have not yet explored this exact relation in the other studies.

Identifying long-lived families

One of our main goals was to identify a group of families who were (genetically) enriched for longevity and identify their living descendants who are interesting to include in genetic research. We describe a strategy to identify descendants from such long-lived families using the HSN data in **chapter 6**. The HSN data is interesting to identify individuals from long-lived families because it connects living persons to their deceased ancestors. The HSN however, does not contain very broad pedigrees and misses mortality information for relevant groups in the first and second generation, such as parents, siblings and spouses. Moreover, mortality information is also incomplete for spouses in the other generations and no relatives are included from the spousal family lines. In other words, the pedigrees in the HSN have sufficient depth, in numbers of generations but are very narrow in terms of known relatives, especially for potential inclusion of families into a genetic longevity study.

The use of the Longevity Relatives Count score (LRC) described in **chapter 6** is related to the narrow pedigrees. Based on the HSN case/control study we could build and test the LRC score only for the proband line (compared to the spousal line) with a maximum of 2 generations of deceased relatives (*Figure 1* in **chapter 6**). The observed results look promising but

there is a need to test the score on both the proband and spousal family lines (relatives of both parents). In addition to this, we currently constructed the score based on top 10% surviving relatives and observed that 27% of the F3 descendants showed a survival pattern similar to the general population even though they had at least one long-lived parent. It remains to be investigated if similar levels will be obtained when building the LRC score based on top 5% or 1% survivors, although we expect this to be the case based on the results of **chapter 4**. In **chapter 4** we showed a survival advantage with each additional long-lived relative, when defining long-lived as belonging to either the top 10% and top 5% survivors. Finally, the LRC score would benefit from validation in an independent dataset such as the LINKS data. The LINKS data could be used to estimate a survival difference between individuals with an LRC score of 0 (controls) and those with an LRC score ≥ 0.30 (cases) but will be difficult to use for the identification of potential phenocopies due to the high level of relatedness for individuals in the database and extreme mortality, especially early in life. As a result, SMRs, which can used to estimate whether a group of individuals follow a mortality pattern similar to the general population, cannot be accurately estimated.

In **chapter 4, 5, and 6** we observed a strong familial clustering of longevity within specific families and used non-genetic/social covariates, pedigree information and, genetic assumptions to distinguish between potential genetic and environmental influences to the familial component within the long-lived families. We obtained evidence for potential genetic influences to the familial component of longevity and these genetic influences are illustrated by the transmission of longevity, even if parents themselves did not become long-lived but had long-lived relatives, such as siblings or parents. Likewise, we observed that an additive increase in the number of parents, siblings, or aunts and uncles is associated with an increase in the survival of study participants and the children of study participants. This additive pattern is not necessarily expected if the findings are due to other, non-genetic, factors that cluster within families (for example wealth). This evidence is strengthened by the fact that similar additive associations were identified for study participants and children of study participants without long-lived parents but with long-lived siblings or aunts and uncles (where the latter generally share less environmental influences with the IPs). Further evidence for the transmission of a genetic component was shown by the fact that none of the tested environmental/social factors, including socio-economic status, sibship size, birth year, twin birth, religious denomination, and birth order, affected the associations between parental/sibling longevity and the survival of their children. These findings are in line with other studies using historical pedigree data[2,77,78] and with unpublished results based on the LINKS data (Forthcoming: Mourits et al, 2019). In addition, the fact that we observed very similar results between the different databases used for our analyses, which cover populations with vastly different environmentally related mortality regimes, significantly adds to the generalizability of our observations regarding the genetic component to human longevity.

Nevertheless, it might be possible that 1. the familial component of longevity is explained by other, unobserved, environmental factors, 2. that some of the definitions of historical factors do not capture the same underlying concept as their contemporary counterpart, 3. That adding the environmental/social factors as covariates to the model in order to explain the parent-offspring association does not cover the full extent of the complexity of familial longevity as for example sibling effect may be independent of parental effects. Regarding the third point, the parent-offspring association captures familial longevity components that might be transmitted from parents to their offspring. There may however be other parts to this

familial component, such as sibship effects, that are not captured by the parent-offspring association. An example may be the competition for resources between siblings in times of scarcity. This competition aspect may in fact explain a part of the familial component of longevity that is reflected in siblings and not in the parent-offspring association. Alternatively, the influences of environmental/social covariates may only become present when sufficient contrast is present between long-lived and non-long lived families. For example, socio-economic status may explain the difference between long and short-lived families but might not explain the difference between long-lived and non-long-lived families as the non-long-lived group can still contain very old persons. An example regarding point two concerns socio-economic status, which in historical data is based on HISCO, a profession based measure or a variant to that[79,80], whereas in contemporary data socio-economic status is often measured in terms of a combination between income, educational attainment, and profession[81–83]. Furthermore, regarding point one, factors such as educational attainment, living environment, social networks, but also lifestyle, eating habits, and activity pattern cluster in families[52–57] and thus may explain a part of the familial component of longevity. Moreover, an important study in Science showed that having little money changes the human mindset to a state of short-term thinking, in terms of financial planning, healthy lifestyle, etc.[84]. Hence, growing up in a state of poverty can significantly influence a person's long-term decision making, affecting that person's health, and at least gives a difficult start in life. Future research in existing data may focus on more detailed analyses of the familial component of longevity. In addition, future research in more contemporary populations, supplemented with extensive pedigree information is needed to gain more insight in the role such factors play in explaining the familial component of longevity.

Future perspectives

Opportunities for longevity research with (big) genealogical data

With the increase of available digitized data, new opportunities have opened up for the analysis of big genealogical data. In the Netherlands, the LINKS project is currently complete for the province of Zeeland and the data is described in **chapter 3 and 4**. The LINKS project is continuing for the entire Netherlands and soon all provinces will be covered in the database, providing researchers with extensive pedigrees and solving the problem of missing migration to other provinces in the LINKS Zeeland version. In addition, worldwide efforts of large genealogy websites such as Geni[85,86] and Ancestry[87] provide a commercial platform with infrastructure for persons to map their family tree and provide a DNA sample to identify unknown relatives. Both the extended LINKS data and the genealogy websites provide novel opportunities to study demographic aspects of longevity and other phenotypes. In fact, the Geni and Ancestry pedigrees have already been used to estimate the heritability of lifespan based on millions of individuals[85,87]. In addition, combining genetic information of living individuals with pedigree data of their deceased and living ancestors has been used to successfully solve cold case murders in the United States[86]. Moreover, combined genetic and pedigree data opens up new research opportunities, as individuals from specific families can be identified based on their family history of a specific trait, such as longevity.

The increase of available digitized data also opens opportunities for linking different data sources. An example, as discussed above concerns combining pedigree and genetic data, but there are also opportunities to link different data sources that are used to reconstruct pedigrees to each other. This allows the cross-checking and improving the quality of pedigree and life course databases, based on a single source. An example of this is discussed in **chapter 3**. In addition, pedigrees in data such as the HSN could be extended by connecting the LINKS data as discussed in **chapter 3**. In the future it may be possible to not only link contemporary genetic data, but also more elaborate data, for example about income, socio-economic status, different health parameters, such as blood pressure, but also social networks, cause of death, lifestyle, and living conditions. By combining all these data sources, rich datasets can be created to provide novel insights into the familial component of longevity and many other traits.

Demographic (longevity) research

In this dissertation we focused mainly on investigating the familial component of longevity by inquiring into the influence of environmental/social factors on individual and familial longevity. Moreover, we studied the definition of the longevity trait itself and used our insights to construct a method, the LRC score, to identify individuals from long-lived families. Now that we have a better understanding about which factors contribute to longevity, how to define longevity and which families to investigate it is possible to use the existing and novel large scale genealogical data to investigate the heritability of longevity in the right families, since this has only been done for lifespan in the general population up to now8[5,87–94]. Following this, it is also possible to test how much of the variance in longevity is due to additive genetic effects, or for example dominant or epistatic effects. Separating the variance of a trait can be done with variance components analysis and was already done with twin studies for lifespan[91,92] but is still open for longevity. In addition to the variance components analysis it is possible to test different transmission patterns of longevity. Testing longevity transmission

patterns can be done with segregation analysis[95–100], but up to now this has not been done for either lifespan or longevity as it requires extensive pedigrees. Segregation analysis allows the estimation of how a phenotype is transmitted to subsequent generations by testing how well the observed transmission fits that of a known genetic pattern. One assumed genetic transmission pattern can be the transmission of a dominant autosomal trait, but many options can be tested, even a non-genetically/socially driven transmission pattern.

Another application of large scale high quality demographic data, such as the LINKS and the UPDB, is to address general demographic research questions. One of the pressing general demographic longevity questions is that of the existence of a decrease in death rates at later ages or the existence of a mortality plateau[101,102]. The existence of a mortality plateau implies that the risk of dying does no longer increase after a certain age. According to some, this implies that the aging process has stopped, since aging goes hand in hand with exponentially increasing death rates during life[103,104]. In animal studies a mortality plateau was already observed[23] but in humans the observations have not led to consistent conclusions[101,102]. A recent study even argues that administrative mistakes in ages at death increase at higher ages and that the accumulation of these mistakes can mimic a mortality leveling-off, or even a mortality plateau in humans[105]. In the Netherlands the civil registration contains a very low level of mistakes as shown in **chapter 3** and in various other studies[106]. Similarly, the mortality data in the UPDB was cross-checked with multiple sources, ensuring high quality mortality data. The low level of mistakes in the Dutch civil registration and the prospect of coverage for the entire Netherlands open up new opportunities to use Dutch data, as well as the UPDB, to investigate the existence of a mortality leveling-off, or mortality plateau at extreme ages in humans.

Using the LRC score, described in **chapter 6**, we were able to identify individuals who likely descent from genetically enriched families for longevity (family cases). Moreover, we identified individuals without any long-lived relatives (family controls) and individuals with up to 20% long-lived relatives but with a mortality pattern that resembles the general population (potential phenocopies) and who are unlikely to be genetically enriched for longevity. For future research it is interesting to connect large scale genealogical data to more contemporary demographic data that contains information on social and economic indicators to investigate what factors cause the familial aggregation of longevity in the potential phenocopy individuals. Moreover, the individuals who become long-lived but have no long-lived relatives at all are interesting to study, as they will likely have acquired their longevity by mean of a healthy lifestyle. We will discuss the potential for genetic research with the family based cases and controls in the next section.

Genetic longevity research

Genetic longevity research focused on the identification of genetic variants with little success[9–16]. Such studies mainly focused on singletons that survived beyond a threshold of for example, 80, 90, or 100 years, and sometimes a survival percentile is used, such as belonging to the top 10%, 5%, or 1% oldest persons of their birth cohort[1,21]. In addition, as in the UK Biobank, middle aged cases are studied based on the longevity of their parents. Loci are sometimes identified in such studies which seem to represent lifespan associated genes, rather than (protective) longevity loci. One of the main reasons for the limited success in genetic longevity studies is the uncertainty in defining the heritable longevity phenotype2 since a large (unknown) part of the population in the last 200 years reaches a high age without representing familial longevity

(phenocopies). It was unknown at what survival percentile longevity becomes heritable in unselected multigenerational datasets (**chapter 4**) and how many family members should be long-lived in order to avoid phenocopies (**chapter 6**). We observed that longevity becomes heritable beyond the top 10% survivors (**chapter 4**) when at least 20%, but ideally 30% of all family members are also within the top 10% survivors (**chapter 6**). We applied this knowledge to construct the LRC score and estimated that 27% of the F3 descendants showed a survival pattern similar to the general population even though they had at least one long-lived parent (**chapter 6**). Hence, in this dissertation, we attempted to improve the inconsistent definition of longevity as an important factor to gain more success in identifying longevity variants. We constructed a longevity definition and identification strategy to select the largest number of persons with a likely genetic enrichment for longevity and we observed that for such definition, the environmental/familial factors seem to play a limited role in reaching the long lived status.

The most consistent evidence has been obtained for genetic variants in APOE and FOXO3A genes[10–15,18], in either genome-wide association studies (GWAS) or candidate gene studies. Such studies use a case-control design, in which allele frequencies among long-lived cases are compared with those in controls. In addition to the misclassification of cases and controls due to the lack of a clear longevity definition, the nature of the used study designs and methods, in GWAS and candidate gene studies only allowed the identification of common genetic variants[107]. The assumed genetic architecture in such studies is that many common variants in the population have small effects on the trait. However, even though longevity studies are likely to often include phenocopies, more genetic longevity variants should have been identify given the very large sample sizes of current meta GWA studies (personal communication on a worldwide longevity GWAS study). Hence, the small proportion of explained heritability by the currently identified loci suggests that rare variants may potentially play a role in the longevity phenotype. Apart from that, the observation that, in comparison to controls, members of long-lived families carry the same numbers of disease risk alleles as other populations of elderly, while showing lower prevalence of age-related disease, may indicate that these families carry protecting factors1[08].

Another interesting possibility concerns the current availability of large genealogical datasets, as described in the "future perspectives" section and combining these data with genetic data. Combining broad genealogic and genetic data provides the possibility to use the distant relatives approach for longevity. In this approach, distantly related long-lived relatives from long-lived families can be examined for the common and rare genetic variants that they share in common (identical by descent, IBD). Distant long-lived relatives in a family are expected to have acquired their longevity by the same genetic variant. For genetic studies it is convenient that such relatives have less DNA in common than for example siblings because it becomes easier to identify the genetic variants responsible for longevity. Moreover, the distant relatives approach has been successfully applied for other complex traits such as oral clefts[109], thoracicaorticaneurysms[110], and osteoarthritis[111]. Second degree nieces and nephews or more distantly related relatives are considered distant relatives and they are illustrated by the fourth (green) generation in *Figure 2* and the "2C" notation in *Figure 3*. These second degree nieces and nephews share on average 238 Centimorgan (CM) DNA strands. More distantly related family members will thus share on average less than 238 CM DNA strands *(Figure 3)*.

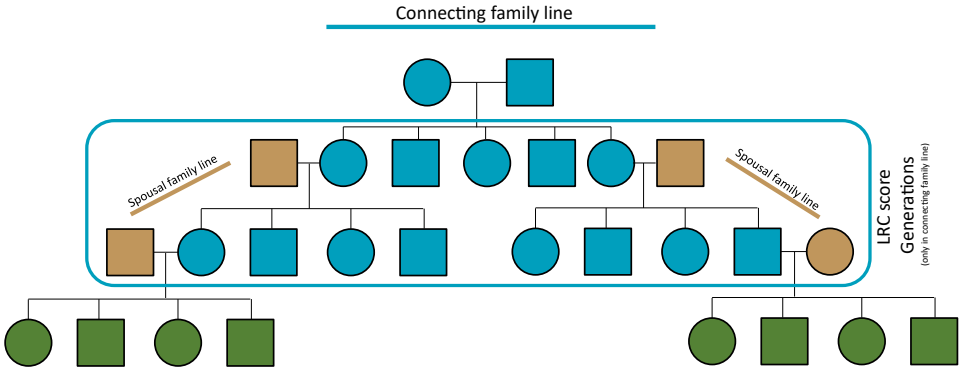

Figure 2: Example pedigree illustrating long-lived families (LRC ≥ 0.30) and distant relatives.

The fourth generation (green color) represents distant relatives (second degree nieces and nephews). A fifth generation (and further) would indicate even more distant relatives. The ancestral (blue and brown colors) generations can be used to identify long-lived families by means of the LRC score.

Here, the intersection between the current availability of big genealogical data, the LRC score to identify the best genetically enriched individuals for longevity, and the need to focus on both common and rare longevity variants using whole genome sequencing data provides novel opportunities to identify the most interesting families for a new genetic study into longevity variants. In addition, current GWA studies focusing primarily on singletons can be extended with familial information in order to rule out misclassified cases who may now obscure the identification of longevity loci in GWAS.

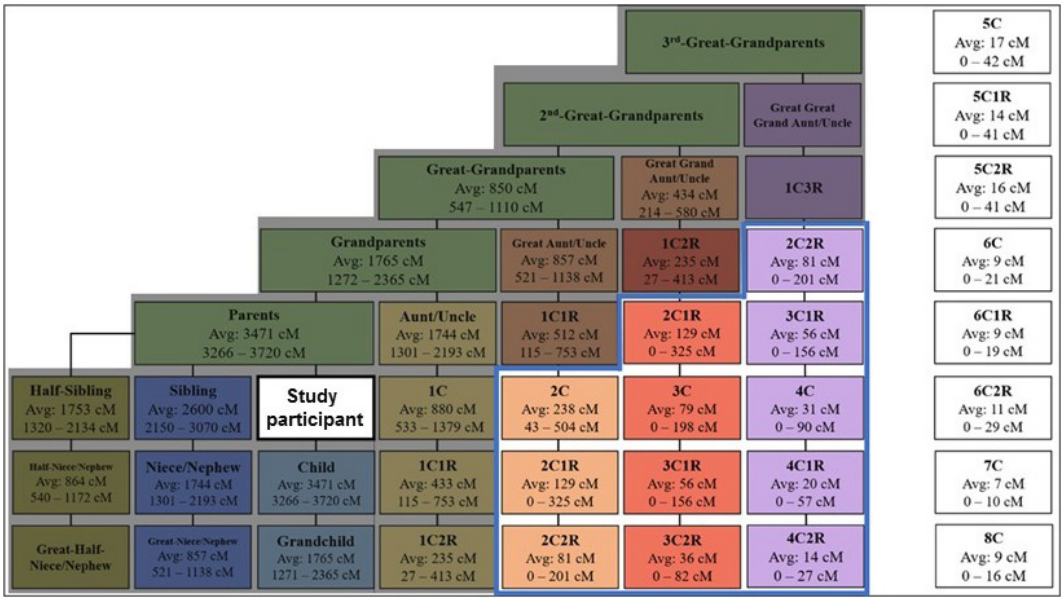

Figure 3: DNA sharing between relatives.

Adjusted from the Genetic Genealogist (https://thegeneticgenealogist.com). G=Great, N=Niece or Nephew, C=Cousin, R=Removed. For example: 3C2R = third cousin 2 times removed (2 generations away).

Final remark

Even though genetic longevity research in humans has been very difficult, the results of this dissertation show that with the selection of cases (and controls) from the proper families new opportunities open up for genetic as well as social research into human longevity.

References

1 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).

2 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).

3 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).

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

5 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).

6 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).

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 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).

9 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).

10 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).

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

12 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).

13 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).

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

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

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

17 Partridge, L., Deelen, J. & Slagboom, P. E. Facing up to the global challenges of ageing. *Nature* **561**, 45–56 (2018).

18 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).

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

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

21 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).

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

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

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

25 Rutten, W. *De vreselijkste aller harpijen: Pokkenepidemieën en pokkenbestrijding in Nederland in de 18e en 19e eeuw.* ('t Goy-Houten: HES Uitgeves B.V., 1997).

26 Walhout, E. C. Is breast best? Evaluating breastfeeding patterns and causes of infant death in a Dutch province in the period 1875–1900. *Hist. Fam.* **15**, 76–90 (2010).

27 Kunst, A. E., Looman, C. W. N. & Mackenbach, J. P. Socio-economic mortality differences in the Netherlands in 1950–1984: A regional study of cause-specific mortality. *Soc. Sci. Med.* **31**, 141–152 (1990).

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

29 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).

30 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).

31 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).

32 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).

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

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

35 van dijk, I. K. & Mandemakers, K. Like mother, like daughter. Intergenerational transmission of infant mortality clustering in Zeeland, the Netherlands, 1833-1912. *Hist. Life Course Stud.* **1**–26 (2018).

36 Hoogerhuis, O. W. *Baren op Beveland. Vruchtbaarheid en zuigelingensterfte in Goes en omliggende dorpen gedurende de 19e eeuw.* (Wageningen Afd. Agrarische Geschiedenis Wageningen Universiteit 2003, 2003).

37 Hofstee, E. W. *De demografische ontwikkeling van Nederland sinds 1800.* (Haalem: Fibula-Van-Dishoeck, 1974).

38 Ekamper, P., Duin, C. van, Poppel, F. van & Mandemakers, K. Heat Waves and Cold Spells and their Effect on Mortality: an Analysis of Micro-data for the Netherlands in the Nineteenth and Twentieth Centuries. *Ann. Demogr. Hist. (Paris).* **120**, 55 (2010).

39 Janssen, F. & van Poppel, F. The Adoption of Smoking and Its Effect on the Mortality Gender Gap in Netherlands: A Historical Perspective. *Biomed Res. Int.* 2015, 1–12 (2015).

40 Lindahl-Jacobsen, R. *et al.* The male–female health-survival paradox and sex differences in cohort life expectancy in Utah, Denmark, and Sweden 1850–1910. Ann. *Epidemiol.* **23**, 161–166 (2013).

41 Beltrán-Sánchez, H., Finch, C. E. & Crimmins, E. M. Twentieth century surge of excess adult male mortality. *Proc. Natl. Acad. Sci.* **112**, 8993–8998 (2015).

42 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).

43 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).

44 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).

45 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).

46 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).

47 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).

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

49 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).

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

51 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).

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

53 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).

54 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).

55 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).

56 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).

57 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).

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

59 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).

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

61 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).

62 Hjelmborg, J. vB *et al.* Genetic influence on human lifespan and longevity. *Hum. Genet.* **119**, 312–321 (2006).

63 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).

64 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).

65 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).

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

67 Lindahl-Jacobsen, R. *et al.* The male–female health-survival paradox and sex differences in cohort life expectancy in Utah, Denmark, and Sweden 1850–1910. *Ann. Epidemiol.* **23**, 161–166 (2013).

68 Rozing, M. P. *et al.* Familial longevity is associated with decreased thyroid function. *J. Clin. Endocrinol. Metab.* **95**, 4979–4984 (2010).

69 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).

70 Arbeeva, L. S. *et al.* How Well Does the Family Longevity Selection Score Work: A Validation Test Using the Utah Population Database. *Front. Public Heal.* **6**, (2018).

71 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).

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

73 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).

74 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).

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

76 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).

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

78 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).

79 Leeuwen, M., Maas, H. & Miles, A. HISCO:
Historical international standard classification of occupations. (Leiden Universithy Press, 2002).

80 Mandemakers, K. *et al.* HSN standardized HISCO-coded and classified occupational titles. (2018).

81 Ganzeboom, H. B. G., De Graaf, P. M. & Treiman, D. J. A standard international socio-economic index of occupational status. *Soc. Sci. Res.* **21**, 1–56 (1992).

82 Braveman, P. A. *et al.* Socioeconomic Status in Health Research. *JAMA* **294**, 2879 (2005).

83 Shavers, V. L. Measurement of socioeconomic status in health disparities research.
J. Natl. Med. Assoc. **99**, 1013–23 (2007).

84 Shah, A. K., Mullainathan, S. & Shafir, E. Some Consequences of Having Too Little. *Science (80-.).* **338**, 682–685 (2012).

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

86 Erlich, Y., Shor, T., Pe'er, I. & Carmi, S. Identity inference of genomic data using long-range familial searches.
Science **4832**, eaau4832 (2018).

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

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

89 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).

90 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).

91 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).

92 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).

93 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).

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

95 King, M. C. *et al.* Genetic Epidemiology of Breast Cancer and Associated Cancers in High-Risk Families. I. Segregation Analysis. *JNCI J. Natl. Cancer Inst.* **71**, 463–467 (1983).

96 Williams, W. R., Anderson, D. E. & Rao, D. C. Genetic epidemiology of breast cancer: Segregation analysis of 200 Danish pedigrees. *Genet. Epidemiol.* **1**, 7–20 (1984).

97 Win, A. K. *et al.* Prevalence and Penetrance of Major Genes and Polygenes for Colorectal Cancer. Cancer Epidemiol. *Biomarkers Prev.* **26**, 404–412 (2017).

98 Newman, B., Austin, M. a, Lee, M. & King, M. C. Inheritance of human breast cancer: evidence for autosomal dominant transmission in high-risk families. *Proc. Natl. Acad. Sci. U. S. A.* **85**, 3044–3048 (1988).

99 Carter, B. S., Beaty, T. H., Steinberg, G. D., Childs, B. & Walsh, P. C. Mendelian inheritance of familial prostate cancer. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 3367–71 (1992).

100 Lalouel, J. M. & Morton, N. E. Complex segregation analysis with pointers. *Hum. Hered.* **31**, 312–21 (1981).

101 Barbi, E., Lagona, F., Marsili, M., Vaupel, J. W. & Wachter, K. W. The plateau of human mortality: Demography of longevity pioneers. **360**, 1459–1461 (2018).

102 Dong, X., Milholland, B. & Vijg, J. Evidence for a limit to human lifespan. *Nature* **538**, 257–259 (2016).

103 Weitz, J. S. & Fraser, H. B. Explaining mortality rate plateaus. *Proc. Natl. Acad. Sci.* **98**, 15383–15386 (2001).

104 Chen, H. -y., Zajitschek, F. & Maklakov, A. A. Why ageing stops: heterogeneity explains late-life mortality deceleration in nematodes. *Biol. Lett.* **9**, 20130217–20130217 (2013).

105 Newman, S. J. Errors as a primary cause of late-life mortality deceleration and plateaus. *PLOS Biol.* **16**, e2006776 (2018).

106 Janssens, A. *Family and social change: The household as a process in an industrializing community.* (Cambridge University Press, 2002).

107 Bloss, C. S., Pawlikowska, L. & Schork, N. J. Contemporary human genetic strategies in aging research. *Ageing Res. Rev.* **10**, 191–200 (2011).

108 Beekman, M. *et al.* Genome-wide association study (GWAS)-indentified disease risk alleles do not compromise human longevity. *Proc. Natl. Acad. Sci. USA* **107**, 18046–18049 (2010).

109 Bureau, A. *et al.* Whole Exome Sequencing of Distant Relatives in Multiplex Families Implicates Rare Variants in Candidate Genes for Oral Clefts. *Genetics* **197**, 1039–1044 (2014).

110 Milewicz, D. M., Regalado, E. S., Shendure, J., Nickerson, D. A. & Guo, D. Successes and challenges of using whole exome sequencing to identify novel genes underlying an inherited predisposition for thoracic aortic aneurysms and acute aortic dissections. *Trends Cardiovasc. Med.* **24**, 53–60 (2014).

111 Ramos, Y. F. M. *et al.* A gain of function mutation in TNFRSF11B encoding osteoprotegerin causes osteoarthritis with chondrocalcinosis. *Ann. Rheum. Dis.* **74**, 1756–1762 (2015).