Death and the Family. High mortality families and the Life coures, The Netherlands, 1812-1912

van Dijk, Ingrid

2019

Document Version:
Publisher's PDF, also known as Version of record

DEATH AND THE FAMILY

High mortality families and the life course, The Netherlands, 1812–1912

Ingrid Kirsten van Dijk

DEATH AND THE FAMILY

High mortality families and the life course, The Netherlands, 1812–1912

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op woensdag 12 juni 2019
om 12:30 uur precies

door

Ingrid Kirsten van Dijk
geboren op 17 november 1989
te Groningen

ISBN
978-94-028-1478-1

Design/lay-out
Promotie In Zicht, Arnhem

Print
Ipskamp Printing, Enschede

Table of contents

Abstract 7

Part I 9
1. Introduction 11
2. Theory 15
   2.1 The role of the family in high infant and child mortality 18
   2.2 Health in later life 20
   2.3 The second generation 24
   2.4 Research questions, hypotheses, and unresolved problems 29
3. Setting the scene: Zeeland, 1812–1912 30
4. The data 35
   4.1 The Zeeland population: the LINKing System for historical family reconstitution (LINKS) 35
   4.2 The Historical Sample of the Netherlands (HSN) 41
   4.3 The Utah Population Database (UPDB) 42
5. A short note on methodology 44
6. Results 46
   6.1 Families in comparison: an individual-level comparison of life-course and family reconstructions between population and vital event registers 46
   6.2 Early-life mortality clustering in families: a literature review 47
   6.3 Like mother, like daughter: intergenerational transmission of mortality clustering in Zeeland, the Netherlands, 1833–1912 49
   6.4 Repeating a lethal experience: mother's and father's exposure to sibling mortality and offspring mortality 52
   6.5 The long harm of childhood? Childhood exposure to mortality and subsequent risk of adult mortality in Utah and the Netherlands 53
7. Concluding discussion 55

Part II 67
Paper I. Families in Comparison: An Individual-Level Comparison of Life-Course and Family Reconstructions Between Population and Vital Event Registers 69
Paper II. Early-Life Mortality Clustering in Families: A Literature Review 101
Paper III. Like Mother, Like Daughter: Intergenerational Transmission of Mortality Clustering in Zeeland, the Netherlands, 1833–1912 131
Paper IV. Repeating a Lethal Experience: Mother's and Father's Exposure to Sibling Mortality and Offspring Mortality 159
Paper V. The Long Harm of Childhood? Childhood Exposure to Mortality and the Subsequent Risk of Adult Mortality in Utah and the Netherlands 195

Part III 219
Summary: “Death and the Family” 221
Nederlandstalige samenvatting *[Dutch Summary]* 225
References 229
List of Abbreviations 241
List of Figures and Boxes 243
List of Tables 245
List of Publications 247
Funding and acknowledgments 248
A word of thanks 249
Curriculum Vitae 251

Abstract

Families play a key role in the survival of infants and children through their vulnerable years. Throughout the 19th century, infant and child mortality were excessively high in the province of Zeeland. This work sets out to assess the intergenerational effects on health and mortality in 19th century Zeeland. To that end, a large database on the Dutch province of Zeeland—LINKS—is employed. For several research questions, comparisons are made to another historical population: the population of Utah, US. This work is the first to explicitly address mortality exposure in the family of origin, to take a comprehensive look at the consequences over the life course, and to assess the explanatory mechanisms that are at work in explaining intergenerational family relations in high mortality.

Early-life deaths were clustered in high-mortality families across the province. The elevated mortality was not limited to one generation or to childhood mortality. A larger number of infant and child deaths occurred in the second generation as well, and surviving individuals tended to die earlier than individuals from low-mortality families, especially if their mother originated from a high-mortality family of origin. The evidence contradicts mechanisms of mortality selection—where relatively frail individuals would be selected out, and surviving individuals would live for a relatively long time—but supports the notion of scarring. According to these theories, the health of surviving individuals who have been exposed to (infectious) disease and high mortality is negatively affected by exposure to disease and mortality. Indeed, this study demonstrates that part of the consistencies in high mortality across generations and the life course can be explained by effects that result from exposure to mortality within the family of origin. Another part can be attributed to characteristics of families that were transmitted from generation to generation, including family socio-economic status and demographic characteristics. Declining infant and child mortality and increasing life expectancy at the end of the 19th century and in the 20th century are likely to be related to declining exposures to mortality in families and in early life.

Part I

1. Introduction

In Hontenisse in June 1853, Anna Maria, aged 25, married Theodorus, a miller, aged 29. In the 16 years that followed, the couple had 13 children. First, four girls arrived: Louisa Anna Maria (1854), Maria Anna (1855), Maria Elisabeth (1856), and Rosalia (1858)—but she died as a one year old. Two boys followed and then another girl, who was again named Rosalia, like her older sister who did not survive. In 1864, Petrus was born, but passed away before his second birthday. In 1865, Jan Francies—named after his maternal grandfather—was born. The next child, born in 1866, was again called Petrus. Unlike the second Rosalia, who made it through the dangerous childhood years, the second Petrus died as a two year old, living only a little longer than his namesake. Three more children followed for the couple as they reached their middle age—the last one, a girl, was born when the mother was already over 40 years old.

The large number of young-age child deaths in this family was not exceptional in the province of Zeeland at the time, but was the sad reality for many families. At the beginning of the 19th century, infant and child mortality in the province of Zeeland, the Netherlands was very high in comparison to other regions in Europe and around the world. Communicable diseases and food- and waterborne diseases causing diarrhea were predominant in the sky-high mortality rates. In some municipalities and years, one third of the newborns did not reach their first birthday and half of them did not live to their fifth birthday. Indeed, high mortality was not limited to the family of Anna Maria and Theodorus themselves: both originated from families in which several siblings did not make it through their childhood. Anna Maria was born into the family of an innkeeper in a large family in Hontenisse. She had 13 siblings, of which four died in infancy and one in childhood. Theodorus originated from a farmer's family and had 13 brothers and sisters, and, like his wife, lost several siblings in infancy and early childhood.

At the end of the 19th century, initially slowly, and then rapidly accelerating, child mortality started to decline. Mortality peaks related to epidemics became increasingly rare, and levels of infant and child mortality became subdued over time. Declining levels of infant and child mortality were initially the main driver of increasing life expectancy in this period. Insights into the changing nature of infant and child mortality in the 19th century are therefore key in understanding how modern health and life expectancy were shaped. To understand these childhood mortality patterns and their changes over time, contemporaries and academics today mainly focus on the individual child and his or her characteristics. Of what cause did they die, and at what age exactly? Where did they originate from and where did they live? Were they boys or girls?

But the investigation should also be aimed at the families in which these children lived. How can it be explained that such a large number of lives were lost in families such as Anna Maria's and Theodorus's families of origin? And why did it happen again when they had children? Were these families poor, and unable to feed and clothe their children properly? That seems unlikely: as innkeepers, millers, and farmers, they may not have been rich, but they were certainly not poor in comparison to their contemporaries. Maybe the parents were overburdened with the care of many young children simultaneously, or childcare habits played a role? Or was it just bad luck?

Historical demographical research and studies focusing on developing-world populations have shown that a relatively small group of families experiences a large share of the infant and child mortality. In other words, infant and child mortality clusters in a select group of families in the population. Although a large number of causes have been explored in earlier research, the exact reasons why early-life mortality is concentrated in a relatively limited group of families remain unclear. That families matter for survival is not surprising as such: after all, siblings share their early-life environment, parents, and family background, all of which may contribute to their likelihood of dying young. Furthermore, young children are highly dependent on their parents and other relatives for their health and survival. Thus, many of the characteristics that are relevant for the survival of children and infants are not found at the level of the individual child but at the level of the family, including biological and genetic characteristics, but also socioeconomic, cultural, and behavioral characteristics. Behavior, including child-caring habits, is learned within families, and affects the likelihood of survival of infants and young children.

Not all children from these high-mortality families died in childhood, but the lucky survivors may have been harmed by their childhood experiences. The factors that are important in causing high infant and child mortality in their families of origin may have a lasting effect on their health and survival, and be related to increased mortality in adulthood. Insufficient nutrition and exposure to infectious disease in families of origin may link together high infant and child mortality in the family and adult mortality, as there exists a relation between adversity in childhood and health in later life. Decreasing infant and child mortality in birth cohorts has been related to an increasing life expectancy of these cohorts, as exposure to infant and child diseases has declined over time, reducing the negative effects of such exposures on life expectancy. As families play a major role in shaping early-life mortality, decreasing exposure to sibling mortality in families of origin could provide a crucial link between mortality in early and later life. Thus, the lucky survivors from high-mortality families could still face the consequences of their childhood environments on their health and survival in adulthood. These interrelated effects on life expectancy may still be at work for

people today, and research on these phenomena may provide us with clues on the mechanisms that link together health in earlier and later life.

Clues on the origins of mortality clustering in families and the effects of exposure to high mortality may be provided by studying intergenerational relations in high infant and child mortality. Studies have shown that elevated infant and child mortality is not limited to one generation, but contains an intergenerational component. Individuals who originate from high infant mortality families, but who themselves survive into adulthood and are able to reproduce, experience in their turn higher mortality among their own infants compared to individuals who grew up in low-mortality families (Lindkvist & Broström, 2006; Vandezande, 2012). In that sense, the family of Anna Maria and Theodorus was no exception. Comparative work has shown that for five historical populations in Europe, a grandmother experiencing a single infant death has daughters whose infants have a 10% increased risk of death before their first birthday. Second-generation mothers whose mother experienced two or more infant deaths face an about 30% higher risk of death for their infants (Quaranta et al., 2017). An intergenerational approach may provide insights into the role of family histories of high mortality that may contribute to early-life mortality risk (Vandezande, 2012).

It is important to note that sociologists, historical demographers, and other social science researchers are principally interested in the socioeconomic, behavioral, and contextual roots of more direct, physiological causes of disease and mortality. In the medical literature, morbidity and mortality are largely studied from an individual perspective, with a focus on direct and indirect physiological causes of disease and mortality, including biological and genetic causes. Social research, including historical and demographic research, does not usually center on the medical causes of death; rather, historians and social scientists aim to determine why some individuals are more vulnerable than others. Often, the causes of death are communicable diseases, to which exposure was relatively common in historical and developing-world populations. Here, diseases such as diarrhea and measles constitute a major cause of death (Preston & Haines, 2014). Therefore, researchers are looking for characteristics that constitute risk factors for exposure and vulnerability to diseases and events that are detrimental for young children. These characteristics, then, are the social, economic, biological, and environmental roots of more direct causes of death (Mosley & Chen, 1984; Sastry, 1997).

Several explanations are possible for family relations in mortality risk, which can roughly be divided into approaches that focus on shared characteristics and causal explanations related to mortality exposure. First, grandparents and parents may share certain characteristics—including social and behavioral characteristics and

genes—that are related to high infant mortality, causing high mortality in multiple generations (Vandezande, 2012). Multigenerational poverty is one of those characteristics, which could connect high mortality in the family of origin to higher mortality among one's own offspring and a poor life expectancy in adulthood. Second, exposure to mortality in the family of origin itself may affect life courses after childhood. For instance, biodemographic theories emphasize that exposure to illness and death in earlier life scars one's own health and may create stress, which then, in turn, leads to reduced life expectancy in adulthood. Life-history theory claims that exposure to high-mortality environments in childhood affects the reproductive career such that there is a stronger emphasis on earlier and faster reproduction, leading to decreased birth intervals and younger age at first reproduction (Nettle, Coall, & Dickins, 2011; Störmer & Lummaa, 2014). There will be lower investment in individual children, possibly leading to higher infant and child mortality and adverse health outcomes in adulthood.

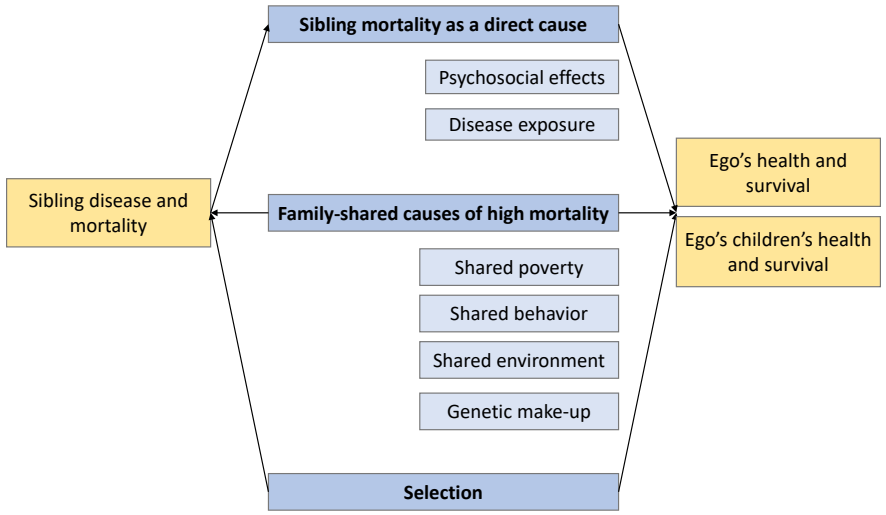
In this dissertation, I investigate how the life courses of survivors from high-mortality families deviate from those who grow up in more beneficial conditions. The central aim of this study is to assess whether increased mortality runs in families and, if so, how increased mortality over the life course—within generations—and in families—across generations—can be explained. To that end, I establish a framework on explaining these deviations, and identify which mechanisms may affect the lives of surviving siblings. I distinguish between family-shared characteristics that cause high mortality in one generation—for instance, behavioral characteristics and poverty patterns—and specific exposures to childhood adversity that may lead to increased mortality within and across generations—for instance, due to scarred health and altered fertility patterns. These mechanisms may not only affect child survival, but also survival in adulthood, and the mortality chances of children in the next generation. The questions that are central to this dissertation are complex and wide-ranging. In contrast to many earlier studies on early-life effects on health, individual life-course data are employed. This allows for a focus on both individual life courses and on family-shared experiences. To that end, a large database on the Dutch province of Zeeland—LINKS—is employed. For several research questions, comparisons are made to another historical population: the population of Utah, US. Furthermore, survival analysis is used, which enables the assessment of the complex of relations between exposure to disease and death in early life, health in adulthood, and survival of offspring.

The dissertation is based on a collection of papers that explore the causes and consequences of high mortality in families. Are there factors at the family level that can explain why infant and child mortality occurs in one family but not another? What happens with the infants and children who have been exposed to sibling mortality and survive into adulthood? Is their health and survival affected? And what happens with their own children, if they have any? By focusing on these questions, the family is placed at the center of research into the health and survival of infants, children, and adults. The papers employ theories from a wide array of disciplines from the social sciences: approaches to health and well-being over the life course from sociology, demography, and biology. They collectively contribute to the understanding of how health and survival are shaped within families, how early-life environments affect life courses of individuals, and how the family contributed to change over time in health, survival, and mortality in the course of the 19th and early 20th centuries. A closer look at the mechanisms connecting mortality in families and across generations generates crucial insights into processes of early-life exposures affecting health, morbidity, and mortality. This has implications for research and theory on mortality and health, not only for historical but also for modern-day populations that experience rapid change in disease and mortality exposure. This work is the first to explicitly address mortality exposure in the family of origin, to take a comprehensive look at the consequences over the life course, and to assess the explanatory mechanisms that are at work in explaining family relations in high mortality.

2. Theory

Mortality rates of infants, children, and adults are related within families, showing that individual survival chances are strongly affected by family backgrounds. Life spans and longevity are partially heritable and shared between siblings (for an overview, see Van den Berg et al., 2017). Mortality is not only related within families in adulthood and for long-living individuals, but also for infants and young children (Edvinsson & Janssens, 2012). Infant and child mortality are concentrated in a subset of families in the population, the so-called high-mortality families. Individuals that originate from such high-mortality families are affected in their health and survival beyond childhood, as their adult mortality rates are elevated. And even in the second generation—among the children of the children from high-mortality families of origin who survive into adulthood—mortality among infants and young children is increased. Why would sibling mortality be related to the survival of other individuals in the same families, both in childhood and in later life? What does it mean to be exposed to sibling deaths in infancy or childhood? Why does exposure to sibling mortality have such a detrimental influence on siblings, life expectancy in adulthood, and offspring survival of individuals?

Figure 1 Mechanisms that link together sibling mortality and the mortality risk of ego and ego's children



Mortality patterns are related within families due to effects of scarring after sibling death, out-selection of relatively robust individuals, and family-shared factors that affect all siblings simultaneously (Figure 1 gives an overview of the mechanisms that explain interrelations in mortality in families). First, scarring effects occur through exposure to sibling death and have a negative effect on the health and survival of individuals. These negative effects can result from infection with disease, stress through exposure to sibling deaths, and other psychosocial effects of growing up in a high-mortality context that affect health negatively. Exposure to sibling death will therefore lead to an increased mortality risk of exposed individuals in infancy, childhood, or adulthood. Second, selection effects work through the weeding out of relatively frail individuals. As a consequence, relatively more robust individuals survive and the remaining siblings experience relatively more favorable death rates. Finally, siblings may be exposed to similar conditions in their families—such as inadequate nutrition and childcare, or growing up in an unhealthy context—that are related to increased mortality and affect them independently. These family-based characteristics work through more proximate determinants of death that directly affect morbidity and mortality (Mosley & Chen, 1984) and affect all family members—infants, children, and surviving adults in later life—who share these characteristics collectively. In contrast, out-selection does not affect the health of surviving siblings but, rather, affects their shared health and life expectancy by removing their weakest

link. With regard to effects of scarring on the health of survivors and the second generation, these scarring effects affect the remaining siblings collectively after the death of a sibling, but do not shed light on the mechanisms that led to the death of the first sibling.

In itself, "scarring" is a black-box explanation that does not shed light on the causes of reduced survival chances of individuals who experienced the death of a sibling. In most studies, no distinction is made between sibling death due to similar causes—as may happen, for instance, due to infectious disease spreading within families—and later sibling deaths as a consequence of earlier infection or stress and other psychosocial effects related to bereavement. The latter effects may also occur indirectly, through parents and other relatives, for siblings who did not experience their sibling death themselves. Parents may be depressed after the death of a child, or birth intervals may be shorter. After cessation of breastfeeding, fecundity is restored more quickly, and also parents may desire to replace the child quickly after a child death. Short birth intervals are related to maternal depletion and increased maternal and child mortality through pregnancy complications, preterm delivery, and low birth weight (Curtis, Diamond, & McDonald, 1993; Omariba, Rajulton, & Beaujot, 2008). For mortality in adulthood, scarring processes may play a role too. Research has shown some support for the notion that in birth cohorts, scarring processes are related to infant and child deaths due to infectious disease because of mechanisms related to the burden of infection and inflammation (Finch & Crimmins, 2004; Crimmins & Finch, 2006). Scarring can also occur through within-family exposure to disease, death, and adversity, which affects the remaining siblings. A large number of authors have argued that early-life mortality clustering results from scarring processes within families, which negatively affect siblings' survival after the death of one of their brothers and sisters (Arulampalam & Bhalotra, 2006, 2008). These effects may also be found later in childhood and in adulthood. Indeed, Alter, Oris, and Brostrom (2001) find that sibling mortality is related to increased mortality before age 15. At the same time, they find evidence for acquired immunity besides scarring processes within families, as mortality in the older age groups is relatively low.

The effects of acquired immunity are, however, difficult to distinguish from the effects of mortality selection in the sibling set. The concept of mortality selection refers to the notion that relatively weaker individuals may be weeded out of the population by relatively high (childhood) mortality in some birth cohorts (Preston, Hill, & Drevenstedt, 1998; Preston & Haines, 2014). In families, this may mean that relatively higher mortality in some families mainly affects the relatively frail children in the sibling set, and that relatively healthy siblings survive. The surviving individuals are likely to be stronger and healthier, have better physical abilities and abilities to work, and possibly

have a higher than average number of children. At the core of this idea lies the assumption that the peers or siblings who would have experienced more negative outcomes did not survive their infancy or childhood.

Family-shared characteristics put all children at risk of early death simultaneously. Family-shared factors include, for example, poverty, housing conditions, an unsanitary or unsafe living environment, childcare habits, and genes. Although these factors could contribute to the death of only one or a few siblings, these factors affect all siblings simultaneously, and lasting effects of these disadvantageous conditions in childhood may be found in adulthood (see Figure 1). Furthermore, through processes of social reproduction and socialization, in adulthood many individuals reflect the characteristics of their families of origin, as they may work in similar occupations and exercise similar marriage, reproductive, and child-caring behaviors. In turn, these characteristics may affect their health and well-being in adulthood, as well as the survival of their own children. In addition to these within-family mechanisms, the role of genes should be acknowledged. Some aspects of health and human tendencies in behavior are (partially) heritable and shaped in the interaction between genes and the environment, and should thus be attributed to their genetic roots rather than family-shared behaviors and characteristics.

To summarize, mortality patterns that are related within families may result from family-shared characteristics and scarring processes that damage surviving siblings, and may be due to out-selection of frail individuals. The consequences of these mechanisms for child survival in families, survival in adulthood, and cross-generational mortality patterns, as well as earlier results from the literature, are discussed in the following three sections.

2.1 The role of the family in high infant and child mortality

It has been shown that families play an important role in both infant (before the first birthday) and child (before the fifth birthday) mortality. Over the past 25 years, there has been a steady interest in this phenomenon, which is known in the literature as infant and child death clustering, the phenomenon that young-age mortality is concentrated in a subset of families in the population. Indeed, the likelihood that families experience infant and child mortality is not equally divided among all families in the population (Das Gupta, 1990, 1997; Edvinsson, Brändström, Rogers, & Broström, 2005). In some families, more infant and child deaths occur than expected given their size, showing that the family matters as far as sibling mortality is concerned. After initial observations of this phenomenon (Brändström, 1984; Das Gupta, 1990), there has been an increasing volume of literature on this topic, which has tried to describe concentration within families.

Most work on death clustering is rather descriptive in nature, assessing whether and how many young-age deaths are concentrated in a subset of families, and how large that subset is, but some studies attempt to pinpoint the causes of increased risk of mortality within the high-mortality families. As was discussed above, factors at the family level that affect survival of all siblings simultaneously include low-quality drinking water, low quantities of food or low-quality food, caretaking behaviors, and factors such as child labor. The health of a whole sibling group may be affected by these characteristics. Some siblings may die in infancy or childhood as a result of malnutrition, increased susceptibility to disease, or other factors. Other siblings may survive their childhood, but their life expectancy and life courses may be affected in later life by this childhood adversity. Thus, sibling death may constitute a red flag for a household under strain.

One of the factors that affects the whole sibling group simultaneously is family socio-economic status (SES). Earlier research has shown that there existed a socioeconomic gradient in infant mortality in the Netherlands, with lower infant mortality among the middle and higher classes than the lower classes. It is hypothesized that this gradient in infant mortality increased over the course of the 19th century, possibly due to healthier living environments for the upper classes (Van Poppel, Jonker, & Mandemakers, 2005) and a slow spread of knowledge about health and access to medical innovations from the upper classes to the lower classes (Antonovsky & Bernstein, 1977). Some evidence has been found for a relationship between socio-economic status and death clustering, as variability in child mortality was larger between low-educated and low social class women in India than higher-educated women (Das Gupta, 1990, 1997) and mortality clustering was lower among share-croppers in 19th century Italy than among landless laborers and nonrural workers (Scalone, Agati, Angeli, & Donno, 2017).

Family characteristics do not, however, necessarily affect all children in the family in a similar manner. The burden of a lack of resources, due to family poverty, may be unequally divided between boys and girls, at the expense of the girls. Earlier studies on sex-specific mortality rates in 19th century Western Europe found excess mortality among girls for certain age groups (Devos, 2000; Poppel, Schellekens, & Walhout, 2009; Janssens, Messelink, & Need, 2010). Although life expectancy for women was higher than for men from an early stage of the demographic transition onward, that survival advantage was inconsistent through the life course. For adolescent girls and adult women in the fertile age range,mortality was often higher than for men, due to maternal mortality and excess mortality due to infectious disease (Devos, 2000). Similarly, for the oldest women and men, in many populations a female survival advantage first developed during the late 19th century (Mourits, 2017). As the female

immune system is more effective than that of males, excess female mortality in the adolescent years has been interpreted as evidence for cultural behaviors and gender preferences within families, that ran at the expense of girls. For instance, women were eating later than men in many places in Europe and were given less and lower-quality food at meals. After the onset of menarche, girls need a higher intake of protein; if their food intake did not match their needs, this may have led to a higher vulnerability to infectious disease, leading to excess mortality (Devos, 2000). Repeated high mortality in families—mortality clustering—may point toward gender-selective neglect within some families. At the same time, for the Netherlands, thus far, evidence has been mixed with regard to the age patterns at which girls were vulnerable.

In addition to the gender of the affected children, indications for the causes of mortality clustering in families may be sought in the timing of deaths. Post-neonatal death may be related more strongly to exogenous causes (Lalou, 1997; Beise & Voland, 2002) such as kin support (Sear & Mace, 2008; Sear & Coall, 2011) and childcare (Edvinsson et al., 2005), while perinatal death is assumed to be related more strongly to endogenous causes such as pregnancy-related factors, health problems, and genetic characteristics (Lalou, 1997; Reid, 2001; Vandezande, 2012; Pakot, 2015). For instance, Reid (2001, 2002) found that a woman's history of child deaths was related to further post-neonatal mortality of her infants, especially that caused by wasting, diarrhea, and respiratory diseases. At the same time, her history of stillbirths and miscarriages affected the likelihood of future repeated stillbirths and miscarriages, possibly due to damaged maternal health (Lalou, 1997). This implies that early-life mortality clustering in some women is related to endogenous or perinatal causes, whereas in other women it is related to exogenous causes, including behavioral factors.

2.2 Health in later life

The clustering of early-life mortality in families and its potential antecedents—negative conditions in high-mortality families—as well as the exposure to disease and mortality of individuals from such families of origin, may have effects beyond infancy and childhood, and may continue to affect the life course in adulthood. Earlier research has shown that there exists an interrelation between cohort mortality levels in infancy and childhood and survival in adulthood and old age. During the epidemiological transition, the increase in expected life span followed 50 years after the decline in childhood mortality, which has led some authors to believe that reductions in exposure to early-life disease were the driving force behind increased life expectancies at older ages (Finch & Crimmins, 2004; Almond, 2006; Crimmins & Finch, 2006; Almond & Currie, 2011). These childhood mortality levels have further been related to a lessening burden of infectious disease on children, leading to an

increasing availability of nutrition for cellular growth. In turn, increasing body stature and weight have led to increasing opportunities to work, better health in adulthood, and better health in the second generation (Fogel, 2004; Quaranta, 2013). These links between childhood and adult health and survival have been fiercely disputed in the literature, and the evidence has been mixed. Research for rural Sweden has shown that high infant and child mortality within birth cohorts is related to lower life expectancies of survivors of these cohorts (Bengtsson & Lindström, 2000; Costa, 2000; Quaranta, 2014), while in other cases no such relationship has been found (Gagnon & Mazan, 2009; Cohen, Tillinghast, & Canudas-Romo, 2010; Hayward, Rigby, & Lummaa, 2016).

The interrelation between childhood and adult survival in birth cohorts has been attributed to a variety of mechanisms. First, bacterial and viral infections are related to inflammatory immune responses and stunting in adulthood, indicators of adverse health (Finch & Crimmins, 2004). Inflammatory immune responses and inadequate development of vital organs and the immune system may lead to reduced life expectancy (Barker et al., 2002; Lindeboom, Portrait, & Van den Berg, 2010; Doblhammer, Van den Berg, & Lumey, 2013) through increased risks of stroke, cardiovascular disease (Caruso, Candore, Colonna-Romano, Lio, & Franceschi, 2005), metabolic diseases (Crimmins & Finch, 2006), and other causes of death in adulthood (Vasto et al., 2007). Furthermore, childhood infections are linked to a variety of adverse health outcomes in later life: cardiovascular and respiratory diseases, cancer, diabetes, and organ damage (Finch & Crimmins, 2004). Also, a permanent dysfunction of the immune system has been proposed (Bengtsson & Broström, 2009). Finch and Crimmins also identify nutritional deprivation as one of the potential links, as in malnourished children—who are more likely to die in childhood—cells and organs do not develop properly and therefore they have a higher chance of developing diseases later in life.

For birth-cohort consistencies in childhood mortality and adult life spans, Preston, Hill, and Drevenstedt (1998) grouped explanations into four broad mechanisms: mechanisms working through scarring, acquired immunity, selection, and correlated environments between childhood and adulthood. Scarring and acquired immunity have a direct, physiological effect, which affects death probabilities negatively in the case of acquired immunity and positively in the case of scarring. Furthermore, selection and correlated environments have indirect, associational effects, leading to relatively robust survivors of high-mortality birth-cohort groups. Selection effects lead to decreased mortality chances of survivors, as they are relatively stronger than the individuals who were weeded out of the cohort. Correlated environments are relevant as a confounder, as the underlying origin of consistencies in mortality levels

is not a direct cause but a common cause related to better health of infants and children on the one hand, and adults on the other.

Exactly which environmental characteristics are relevant for child and adult survival, as well as for improved survival during the demographic transition, has been fiercely contested in the context of the debate on the origins of the demographic transition. Proposed explanations include, but are not limited to, modernizing economies, increasing access to and availability of nutrition and medical care, better public health and hygiene, and decreasing virulence of epidemics. The underlying cause of improvements in life expectancy over the 19th century is therefore not necessarily decreasing exposure to infant and childhood diseases, but may also be sought in characteristics that affect infant and child as well as adult survival simultaneously. Cohort factors—which simultaneously affect a group of individuals born in the same year—have been identified as relevant in conjunction with period factors—which affect all individuals who are alive at a certain moment, regardless of their age. McKeown (1976) emphasized a role for an increasing availability of nutrition, after rejecting explanations based on medical technology and improvements in public health. Fogel and his colleagues maintain that not only improvements in nutritional status but also the interaction with declining infections and increasing availability of nutrients for cellular growth contributed to increasing body health, size, and survival (Fogel & Costa, 1997; Fogel, 2004; Floud, Fogel, Harris, & Hong, 2011). In turn, this led to increased earning capacities in adulthood and increased resources available for the next generation. From generation to generation, therefore, health, body size, and earning capacities increased further. However, the work of Fogel and his colleagues was based on body heights in several countries, and some of their data have been criticized for being overly pessimistic with regard to human stature in the 18th and 19th centuries (Öberg, 2018).

Most of the research into the connection between mortality exposure and survival in old age has been conducted using birth cohorts. In the demographic and epidemiological literature on old-age mortality, relatively little attention has been paid to the role of families in shaping early-life conditions of individuals and mediating exposure of infants and young children to mortality. However, it is unlikely that within populations, all individuals from one birth cohort are exposed similarly to infectious disease and deaths. Strong geographic and familial differentiation may exist. Individual experiences with exposure to mortality are therefore affected to a high degree by family contexts, rather than by cohort-level mortality rates, and will differ strongly over the population. Furthermore, infants and young children are highly dependent on their families, and therefore families are the context in which children's early-life health experiences are established. In the light of the finding that deaths and disease are concentrated in certain families, the question arises as to how families mediate exposure to cohort mortality, and what the effect of sibling deaths in infancy and childhood is on life expectancy. Surviving siblings are likely to have been exposed to sibling mortality and to have suffered from undernutrition (Hanson, Smith, & Hasstedt, 2013). Thus, the role of the family in shaping early-life survival and death underlines the fact that mortality experiences should be narrowed down from cohort exposure to individual-level mortality exposures or, more precisely, to exposure to mortality within the family of siblings, which may leave a lasting imprint on the health of surviving individuals.

Furthermore, by considering mortality in families rather than in birth cohorts, it is possible to distinguish more precisely *why* there exists an interrelation between mortality at young and older ages. Can this interrelation be attributed to infectious disease exposure within families? Then, mechanisms of scarring and declining exposure to childhood diseases play a role in increasing life expectancy in the past. Are there differences in adult survival of individuals originating from families with varying levels of infant and child mortality? Do these exist independently from exposure to disease and death of siblings? Then it is likely that between-family differences in nutritional patterns and behavioral factors, such as child-caring habits, play an important role in early-life effects on health and survival. If differences are mainly found within families, between siblings who were and who were not exposed to sibling disease and death, it is likely that factors related to infectious disease and death exposure dominate nutritional and behavioral factors that are shared within a sibling group.

Some earlier research has focused on the interrelation between sibling health and adult life courses and has shown that there is a lasting impact of sibling mortality on their other siblings' life courses in adulthood. Alter et al. (2001) analyzed sibling mortality to assess the role of health behaviors in explaining excess mortality in early and later life. They found that sibling mortality affects mortality among other siblings during childhood but diminishes after age 15, with no effects after age 55. This evidence is consistent with acquired immunity after age 30. Yu et al. (2017) found that in modern-day Scandinavia, childhood exposure to sibling mortality leads to increased likelihood of death between birth and age 37, which suggests that sibling deaths have a lasting influence on health and mortality. At the same time, the effects were strongest shortly after the event. Individuals could not be followed beyond age 37. Norton et al. (2017) found that deaths in the family of origin during childhood predict systemic inflammation in later life, suggesting that stress responses link bereavement to adult life expectancy. McFarland et al. (2018) suggest that stressful events in early life, including death of individuals close to ego, such as siblings, are related to cellular ageing, indicated by telomere length for individuals aged 22–68. It

is unclear, however, whether such effects can also be found into old age. Other questions remain with regard to the exact causal mechanisms, comparability across populations, and the window of exposure.

Exposure to infectious disease, sibling deaths, or other childhood adversity has a stronger effect on later health and well-being in some stages of life than in others. There is no consensus in the literature about the age at which exposure to mortality affects later health outcomes: when individuals are *in utero*, infants, young children, or older children. Even the months before conception may be relevant, as maternal disease and nutritional status are likely to affect the health of the fetus (Quaranta, 2013). Famines are related to shorter life expectancies for newborns at the time of the famine (Lindeboom et al., 2010), but also for children who are exposed *in utero* during mortality peaks (Barker et al., 2002; Bengtsson & Lindström, 2003; Doblhammer, 2003). Barker and his colleagues (2002) maintain that fetal undernourishment and low birth weight are related to cardiovascular disease and type 2 diabetes via catch-up growth. Bruckner and Catalano (2009) found no effects of cohort infant mortality on adult mortality rates and hypothesize that the critical period of children's exposure to disease is later in childhood. Other researchers have hypothesized that exposure to mortality is significant up to age 5 (Fridlizius, 1989) or 15 (Crimmins & Finch, 2006). Possibly, periods of relatively strong growth in particular are relevant to later life survival. A systematic exploration of the stages during which early-life adversity may affect later life health and survival is, however, lacking in the literature.

Finally, men's and women's health and survival in relation to sibling mortality exposure should be examined separately. Men have higher mortality rates than women for sociocultural and biological reasons, including behavioral patterns, the protective role of estrogen, the suppressing effect of testosterone on immunity, and an unprotected X chromosome (Lindahl-Jacobsen et al., 2013). The effect of early-life conditions on survival may therefore be stronger for men than for women (Smith, Mineau, Garibotti, & Kerber, 2009; Doblhammer et al., 2013). Furthermore, in historical populations there were different expectations of girls and boys with regard to socioeconomic contributions to the household as well as caretaking tasks—including taking care of ill brothers and sisters—leading to specific exposures to sibling disease and workplace hazards for boys and girls.

2.3 The second generation

Exposure to high mortality could not only affect one's own life expectancy negatively, but also have a detrimental effect on the survival of one's offspring. Infant and child mortality levels that are related across generations may originate from mechanisms similar to that for the relation between childhood and adulthood health. On the one hand, individuals' families of origin and their own families in adulthood may share characteristics that put children at increased risk of mortality in both generations. These characteristics are transmitted across generations, through processes of socialization—cultural characteristics, norms, and behavior—and socioeconomic status through social reproduction. On the other hand, exposure to sibling mortality may affect psychosocial characteristics and the health of surviving individuals in such a way that their children are less likely to survive their childhood.

Poverty is a structural characteristic of certain families across generations, which is repeated through processes of social reproduction. In several consecutive generations, poverty-related adversity may then lead to an increased likelihood of infant and child mortality. Evidence from the literature shows that poverty is a structural problem across generations, which is repeated through processes of social reproduction. The poverty of an ancestor affects the social status of offspring up to the third and the fourth generations, even after controlling for the social status of fathers and uncles (Knigge, 2016). Individuals originating from poor households may therefore experience increased mortality among their own children, and also face a lower life expectancy.

Furthermore, parental preference for one of the sexes or a specific treatment of boys and girls may lead to an increased chance of mortality in early life for one of the sexes. Because such preferences are shaped through processes of socialization, it would be expected that gender preferences would occur in several generations of the same family. As a consequence of sex-specific neglect, increased mortality is likely across generations of one family, and would affect girls rather than boys.

Besides family characteristics shared between generations, high-mortality environments in one's childhood may also independently affect the likelihood of child mortality in subsequent generations. That could occur through two mechanisms: effects on health of the parents who were exposed to mortality, disease, and adverse conditions in their families of origin, and the socio-psychological effects on the life course of growing up in a high-mortality environment. In earlier work on intergenerational continuities in high infant mortality, this group of mechanisms has largely been overlooked.

First, individuals' health and women's reproductive health can be negatively affected as a result of exposure to mortality and disease in their young years. Damaged childbearing capacities of women may result in reproductive problems, stillbirths, and child loss (Lalou, 1997). If unhealthier men and women have children, in turn, their children's health can be negatively affected by their parent's bad health and reduced life expectancy. Parental ill health and early death affects their children's health and survival negatively, because of economic stress, psychological stress,

and grieving, or—in the case of ill health—a reduced ability for the financial and physical care of young children. These adverse childhood environments could, in turn, negatively affect the health of these children in adulthood. Thus, adverse health may be linked across generations. Research into the effect of parental mortality on child survival has shown that a maternal death in particular has a detrimental effect on her offspring's survival, especially for young children (for an overview, see Rosenbaum-Feldbrügge, 2018). For a paternal death, a more mixed picture emerges. Effects on child survival are smaller than for maternal death (Beekink, Van Poppel, & Liefbroer, 1999; Sear & Mace, 2008) and in many populations effects of paternal death on child survival are small or nonexistent (Derosas, 2002; Sear & Mace, 2008). These differences between fathers and mothers are likely to be caused by different roles within the household in historical populations, with mothers as the main caretaker and fathers as the main earner of the household. In addition, widows and widowers apply differential coping strategies after widowhood, such as remarriage, splitting up families, and migrating toward relatives (Rosenbaum-Feldbrügge, 2018), affecting their children differently. Finally, in historical populations, widows may have more easily qualified for poor relief and other assistance to their families than widowers, hence reducing the impact of partner death on offspring survival.

Second, life-history theory holds that there exist psychosocial effects of growing up in adverse environments, including high-mortality environments, that affect reproductive careers. In turn, the likelihood of survival of the children of parents from risky environments may be negatively affected. The effects of adverse early-life environments on the life course remain hardly ever studied, beyond direct health effects through processes of scarring. The major exception is evolutionary biology, where mortality exposure in childhood and its outcomes during the life course are studied from an evolutionary angle. Here, reproductive careers are studied from the perspective of life-history theory. Here, the degree of security and predictability in the childhood environment is assumed to be related to psychosocial mechanisms that, upon reaching the reproductive years, lead to diversity in timing of reproduction and investment in children. Such mechanisms include sexual risk-taking behavior.

Life-history theory claims that exposure to childhood adversity affects life histories, including the speed of reproductive careers and the level of infant mortality. This diversity in reproductive careers has been evolutionarily shaped, referring to the likelihood of successful reproduction. Through subconscious processes, individuals judge the potential length of their lives to decide how resources are allocated toward reproduction and long-term strategies, solving the allocation issue of resources and time. Thus, the bodies of individuals from less secure environments develop more quickly, with a younger onset of puberty and for girls menarche, leading to an earlier

start to sexual activities; an earlier start of reproduction; higher fertility, and lower investment in individual offspring (Chisholm, 1993; Nettle, Coall, & Dickins, 2011; Störmer & Lummaa, 2014; Webster, Graber, Gesselman, Crosier, & Schember, 2014). Furthermore, preferences with regard to timing of reproduction and risk-taking may also be affected (for further references, see Störmer and Voland, 2014).

Cues about danger in the environment are, in addition to father absence and mortality in the population, also taken from mortality in the sibling set. Hardships also include, for instance, food shortages and famines; climatic events and disasters; and stressful early-life environments related to unstable places of residence and housing, parental absence or death, and sibling mortality. For modern populations, low socioeconomic status environments are considered possible precursors to faster life histories. Störmer and Lummaa (2014) linked high mortality in the family, rather than individual experiences with mortality, to both male and female age at first birth and age at first marriage. They found that all sibling mortality within the sibling set was a precursor of these outcomes, rather than individual-level exposures to sibling mortality. As a consequence of this increased fertility and shortened birth intervals, there will be a larger group of offspring and relatively fewer resources to invest in each individual child, which may lead to elevated infant and child mortality among the offspring of affected individuals (Chisholm, 1993; Störmer & Lummaa, 2014).

The larger sibling groups that are predicted through earlier onset of reproduction and faster reproduction can be an adverse environment in which children have to grow up. In historical populations in which household resources are scarce, larger sibling groups are correlated with increased competition over resources among the siblings (Stradford, van Poppel, & Lumey, 2017). At the same time, this does not necessarily affect all children negatively: the distribution of resources between the children is decided upon within the household (Riswick, 2018), and children beyond a certain age contribute to the household as well as consuming part of its resources (Hammel, 2005; Hagen, Barrett, & Price, 2006). Furthermore, in larger sibling sets, overcrowding and poor hygiene contribute to poor health of children (Hatton & Martin, 2010). Kippen & Walters (2012) found that, in particular, the number of brothers and sisters in the household affected mortality in the first month of life. It is unlikely that in the first month competition over nutrition plays a large role; during these months, maternal depletion and transmission of diseases between siblings is more important in high infant and child mortality in large households.

To summarize, exposure to high mortality may affect not only the chances of survival in childhood and adulthood, but also the survival of the second generation. These effects may work through health effects on caretaking and earning capacities as well

as through psychosocial effects on reproduction and investments in individual offspring. As was discussed above, not only exposure mechanisms but also family-shared characteristics may affect the survival of the second generation. In this study, both groups of mechanisms are explored in depth.

Box 1 An overview of the hypotheses on exposure to mortality in families and effects across life courses and generations

Mechanism	Link across life course and/or between first and second generation	Exposure related?	Mother (M) or father (F) linked?	Timing related?	Other indicators	Papers
Scarring	Damaged health of parents results in decreased caretaking capacities	Yes	M as main caretaker, F as main provider	No	Early death of parents	2–5
Damaged reproductive health	Damaged reproductive health results in problematic pregnancies and childbearing	Yes	M	Puerperal period		3, 4
Life histories	Fast reproduction and short birth intervals lead to higher mortality and adverse health	Unclear	M as main caretaker	No	Age parent at first birth; short birth intervals	3, 4
Poverty	Low socioeconomic status is related to mortality and health and transmitted between generations	No	F as main provider	No	SES of grandparents and parents	3–5
Behavior	Caretaking shortfalls and habits cause high mortality in multiple generations	No	M as main caretaker	No		5
Genetic deficits	Genetic diseases	No	M and F	Cross-generational similarities?		

2.4 Research questions, hypotheses, and unresolved problems

The dissertation is based on a set of papers that are guided by a series of research questions:

1. To what extent do grandparents and mothers experience similar mortality among their infants and children offspring? And to what extent does that apply to grandparents and fathers?
2. How can intergenerational correlations in infant and child mortality be explained?
3. How does exposure to mortality in the family of origin affect mortality in adulthood and in the second generation, and how can this relation be explained?

A number of hypotheses are formulated:

Hypothesis 1. Exposure to high mortality in the family of origin has distinct effects across the life course, including increased offspring mortality and heightened mortality in old age,

Hypothesis 2. Family characteristics such as poverty play a role in increased mortality across generations of some families in the Netherlands,

Hypothesis 3. Increased mortality in old age of individuals exposed to high mortality in their sibling set can be attributed to mechanisms of inflammatory immune responses,

Hypothesis 4. Increased insults—via increasing exposure to adversity—are related to stronger effects on mortality in adulthood,

Hypothesis 5. Both family-shared characteristics and increased mortality among disease-exposed siblings explain part of the relation between sibling mortality and mortality in the second generation.

Hypothesis 6. Selection and scarring affect child and adult mortality in certain families, and are visible in mortality crossovers in adulthood of survivors from high-mortality families of origin.

Hypothesis 7. Evolutionary mechanisms explain part of the relation between childhood mortality exposure and mortality in the second generation.

These questions will be answered in the series of papers that constitute the main body of this dissertation (see Part II). The conclusions, the shortcomings of the papers, and the implications for the literature will be discussed after a short discussion of the results below.

3. Setting the scene: Zeeland, 1812–1912

Zeeland is a largely Protestant province located in the southwestern coastal region of the Netherlands, and consists of islands and peninsulas (see Figure 2). In the 19th century, the province was predominantly rural and agricultural. Industrialization had a relatively limited impact on the province, as the scale and geographic distribution remained modest. The process started relatively late, as it only took off after 1900 (Priester, 1998; Blom & Lamberts, 2006). In the province, there were only a few large towns, including Vlissingen, Goes, and Middelburg, where the population density rose slowly over the 19th century. At the same time, these towns were rather small, whereas the surrounding countryside was relatively densely populated (Ekamper, Duin, Van Poppel, & Mandemakers, 2010). The phenomenon of the urban graveyard—unhealthy living conditions in cities that caused high mortality, whereby high in-migration helped to help sustain labor potential in urbanized regions (Reher, 2001; Woods, 2003)—was therefore largely unknown.

The economy was mainly dependent on market-oriented, capital-intensive farming, focused on growing wheat, flax, and madder on fertile sea clay (van Cruyningen, 2005). Although the majority of the population belonged to the unskilled agricultural-worker class, wealthy farmers constituted a small but prosperous share of the population (Bras & Neven, 2007). Among the agricultural workers, women's labor was relatively common. Married and unmarried women, as well as children, worked as field laborers on farms, especially during the busy summer months. As a consequence, young children were sometimes left at home, unattended by adults (van Cruyningen, 2005). At the end of the 19th century, the agricultural sector went through a prolonged crisis, which hit commercial farmers in particular hard.

Over most of the 19th century, the life prospects of infants and children were short, with a child mortality rate reaching more than 40% in some municipalities and years (Van Poppel & Mandemakers, 2002; Hoogerhuis, 2003). Furthermore, the onset of the demographic transition occurred relatively late in comparison to other countries (Hofstee, 1981). Infant mortality began to decline after 1880, but remained high in an international perspective until the early decades of the 20th century (Van Poppel et al., 2005). At the time, the high infant mortality rate in Zeeland was blamed on women who did not breastfeed their children, as they were working outside the home. Indeed, it appears that breastfeeding habits played an important role in explaining the high infant mortality (Hoogerhuis, 2003).

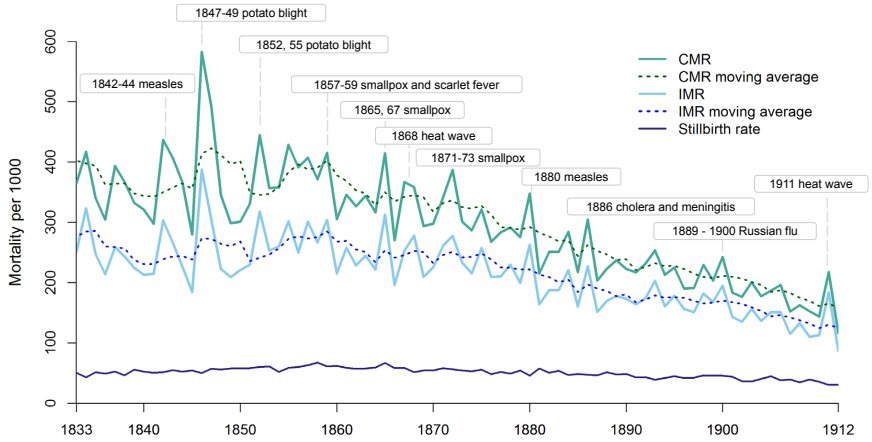
Figure 2 A map of the province of Zeeland, 1866

Source: Kuyper, 1868.

Other factors that played a role in high mortality in Zeeland included the quality of drinking water and the occurrence of malaria. Surface water in Zeeland was often unfit for human consumption due to salinization. Water in canals and cities often did not constitute a safe alternative, and its use led to a high prevalence of waterborne diseases. Until water pipes were constructed in the province, from 1910 onward, in most municipalities rainwater was used as the main source of drinking water. However, saved rainwater could spoil and be infected with bacteria, and the amount was not always sufficient. Especially during warm summers, there could be a lack of safe drinking water. *Malaria tertiana*, with a milder disease profile than *Malaria tropica*, was endemic in coastal regions of the North and Northeast of Europe, including Germany, Sweden, Finland, and the Netherlands. Although the disease started to decline after the 19th century, it played a role in illnesses, especially in summer, until the early 20th century (Hoogerhuis, 2003). In the Netherlands, the last case was confirmed in the

summer of 1959, and in 1970 the Netherlands was declared malaria free by the World Health Organization (WHO), as one of the last countries in Europe (Van der Kaaden, 2003).

Figure 3 Stillbirths and infant and child mortality in Zeeland, 1833–1912



Sources: Own calculations based on LINKS Zeeland 2017_01 (Mandemakers & Laan, 2017). Moving averages represent the rolling average mortality rates over five-year periods. For peak mortality years, the following literature has been consulted: Hofstee (1974), De Meere (1982), Hoogerhuis (2003), Ekamper et al. (2010), and Krantenbank Zeeland [Newspaper Archive Zeeland] for the relevant years.

In conjunction with these contextual elements, earlier research has found an important role for sociocultural characteristics in shaping mortality and life (Hoogerhuis, 2003; Van Poppel et al., 2005). Families therefore play a key role in infant and child survival, and for understanding long-term mortality change it is crucial to use historical micro-data. The role of families is highlighted by the concept of child mortality clustering. Early-life mortality clustering is defined here as an excess concentration of early-life deaths in a subset of families. In other words, there are more families experiencing a high number of child deaths than could be expected given the number of families of that size in the population. Here, I will illustrate the infant mortality clustering in the province of Zeeland.

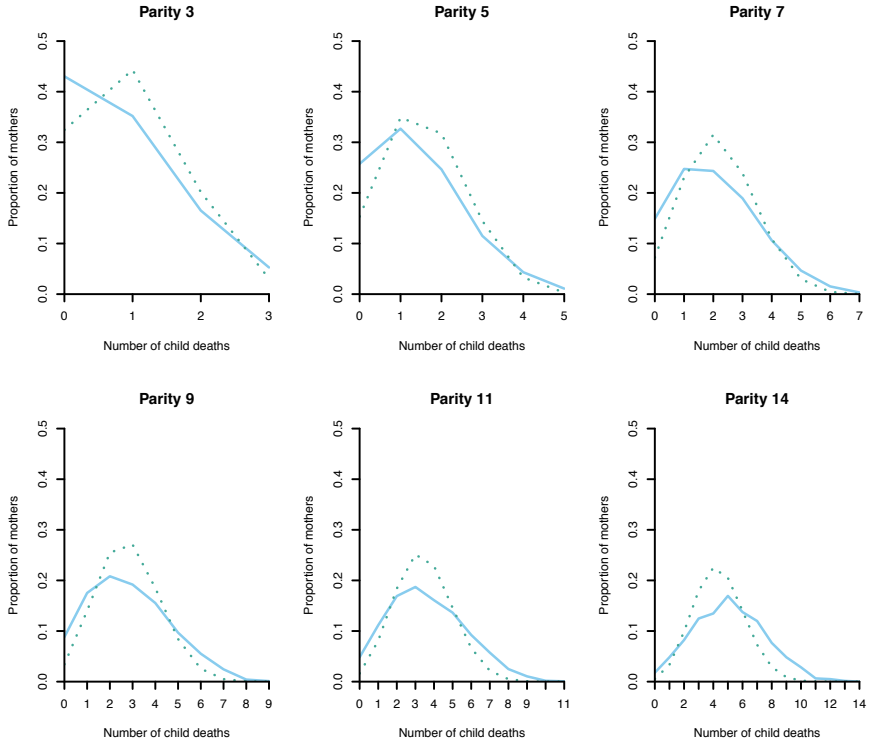
In Zeeland, some families experienced a higher infant and child mortality than would be expected given their size and the infant mortality rates in the population. Other families experienced much lower mortality than population averages would predict. That fact, as such, is unsurprising: chance predicts that some families will experience more child deaths than others, even if they all have the same number of children and every child has the same chance of dying before their fifth birthday. The fact that families differ in the number of children that they lose before a certain age only becomes remarkable if the concentration in families is stronger than chance would predict. And that is exactly what was found for the province of Zeeland: there was a stronger concentration of deaths in some families than chance would predict. And thus, the family into which a child was born determined to a large extent whether or not that child survived until its first and fifth birthday.

In the literature, several approaches have been used to assess the role of the family in shaping infant and child mortality, and the concentration of infant and child mortality in families. These can be subdivided into four strands: approaches that focus on estimating the effect of scarring on siblings, threshold models that assess whether a specific level of mortality has been reached in a family, methods that focus on variation between families within populations, and methods that focus on mortality clustering in the population (for an extended discussion, see Part II, Paper II, and Paragraph 6.2). For the province of Zeeland, I used three different methodologies, including assessing whether there existed mother-level variance in infant and child mortality that is larger than chance would predict. Here, another approach is shown. Following earlier work (Zaba & David, 1996), I show the likelihood that a family of a certain size experiences a certain number of infant deaths. Mortality clustering in the population should then be visible by a higher number of families, the so-called high-risk families, experiencing many more infant deaths than expected based on the calculations. On the other hand, there should be fewer families than expected experiencing a more average number of deaths. This pattern is taken to represent mortality clustering in families in the population.

Here, I show these results for families with three, five, seven, nine, 11, and 14 children, respectively. The figures show that for families of these sizes, there are more families experiencing zero deaths or a few deaths than expected based on the number of families of this size and the population mortality rates. At the same time, there are fewer families experiencing average numbers of deaths. Finally, more families than expected experience a large number of child deaths. This pattern is in line with those that were expected for populations in which child deaths cluster in a subset of families.

To assess the long-term effects of adversity in early life and exposure to sibling mortality on survival in adulthood and survival in the second generation, I also employ threshold models, using the number of deaths in the family of origin—controlling for family size—to investigate the effect of sibling mortality on survival in adulthood. I further include the number of deaths, to see whether a higher exposure was related to stronger effects in adulthood and on the second-generation children.

Figure 4 The clustering of child deaths in families with three, five, seven, nine, 11, and 14 children in Zeeland, 1833–1912



Notes: The dotted lines represent the expected number of infant deaths; the filled lines represent the observed number of infant deaths. For all family sizes, the figures show that there are more families experiencing zero deaths or a few deaths, fewer families experiencing average numbers of deaths, and more families experiencing a large number of deaths. This pattern is an indication for mortality clustering.

4. The data

The body of the project is based on LINKS, a historical family reconstruction database from the Netherlands, the province of Zeeland. For a paper focusing on the data quality, comparisons are conducted with an established life-course reconstruction database, the Historical Sample of the Netherlands. For comparative work, I employ the Utah Population Database.

4.1 The Zeeland population: the LINKing System for historical family reconstruction (LINKS)

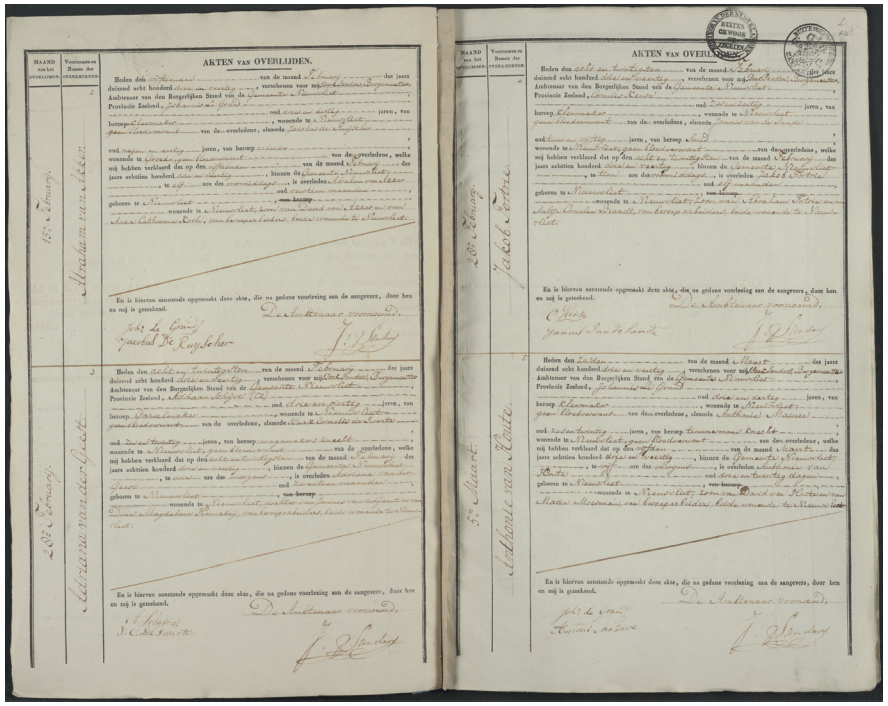
LINKS is the outcome of a project with the aim of reconstructing all 19th and 20th century families in the Netherlands, based on civil certificates. The project is unique for the Netherlands, and its scope and size are matched by only a few historical demographic projects in the world. The LINKS family reconstructions are based on digitized indexes of civil certificates from the 19th and 20th centuries. For over 20 years, numerous volunteers have been working to build these indexes, known as *WieWasWie* ("WhoWasWho"). For the province of Zeeland, the reconstruction of families is completed, and the database contains almost 2 million persons from the province of Zeeland, including almost 1 million births and deaths and 200,000 marriage certificates.

Mandatory vital event registration was introduced in the Netherlands in 1812 by the French. The registers of births, marriages, and deaths become public with a delay of 100, 75, and 50 years, respectively. Indexes of the registers have been released for the period 1812–1912 for births; 1812–1937 for marriages; and 1812–1962 for deaths. Unlike population registers, civil registers do not contain information about religion and household composition. For some municipalities and years, home addresses were registered, but these are not part of the WieWasWie project on which LINKS is based. The indexes of the civil certificates contain information on names of individuals who were born, married, and deceased, and also the names of parents, witnesses to the vital event, the municipality in which the vital event took place, the ages of the individuals and sometimes of the parents, and occupational titles.

To construct the LINKS dataset, civil certificates of individuals were linked together on first and last names of individuals. By linking individuals, the database has reached a large scope, enabling the identification of kin over a large geographic area and in several generations. In this case, vital events from families occurring in Zeeland over time can be followed for up to eight generations. LINKS has made possible research on historical life courses, families, and populations. The project has enabled researchers to ask new research questions, due to its large scope and the large number of life

courses and wide kinship networks that are contained in the database. Up to the present time, the database has been used for over 50 publications (IISG, 2018) on a broad range of topics, including the role of grandfathers in shaping socioeconomic status (Knigge, 2016), women in the labor market (Schulz, Maas, & van Leeuwen, 2014), maternal mortality (Ory & Van Poppel, 2013), intergenerational transmission of fertility (Bras, Van Bavel, & Mandemakers, 2013), and numerous other topics. The major innovation of LINKS is the large number of families that can be followed from generation to generation and between municipalities.

Figure 5 Examples of death certificates



Source: Burgerlijke Stand Gemeente Nieuwvliet.
Notes: The example death certificates concern the deaths of four children, at age 14, age 17, age 11 months and age 23 days and further include the place and date of death, and the names, ages, occupations, and place of residence of the parents.

Within the database, individuals were linked together into families in four steps. First, marriage certificates of children were linked to marriage certificates of parents, using the names and ages on the certificates. Second, death records were linked to birth records. Third, birth and death records were linked to the marriage certificate of the parents. Fourth, birth and death records were linked with marriage certificates of individuals. Small inconsistencies and double links—when certificates were linked to several other certificates—were resolved by selecting exact links or by not accepting established matches. To avoid false matches, matching was limited to within the time frame on which events could logically occur. For instance, if names of parents matched but the age of the possible mother indicated that she was outside her fertile years, the potential link was not established.

LINKS contains individuals for whom follow-up is available, as well as unlinked information on individuals. Follow-up on individuals is sometimes lacking because of out-migration and temporary or seasonal migration, leading to oversight of the birth of children, marriages, and, in the case of permanent out-migration, deaths. Persons who migrated between municipalities in Zeeland could be followed over time; however, individuals who left the province of Zeeland were lost from observation. Furthermore, toward the end of the research period, an increasing number of events will have occurred outside the time frame that has been indexed for LINKS. Matches between indexed certificates are sometimes unsuccessful, for instance, due to name changes or substantial spelling differences between civil certificates. Upon marriage, birth certificates were shown to the administration, and therefore this issue exists for deaths rather than for other vital events. Births to unmarried mothers are not linked, because of the large number of potential mothers based on the name of only one parent. Furthermore, individuals who died in old age often do not have their parents mentioned on their death certificates, as the names of their parents were not always known to those who registered their deaths with the municipality. This creates additional problems for linking these individuals to their birth and marriage certificates and linking them into their family network.

These potential issues may affect the results of studies based on LINKS and have to be taken into account upon designing and executing studies employing the database, especially for selecting relevant cases for analysis. Authors of earlier studies based on LINKS have reflected on the quality of the database, primarily oriented toward the quality of its sources and potential omission of demographic groups such as single mothers and migrants. Although data-quality checks—such as explorations of sex ratios at birth over time, percentages of cases with follow-up, the population share that marries, and other demographic indicators, as well as checks of consistency of individual life courses—have undoubtedly been performed, these are largely lacking

from the (published) literature. It therefore remains largely unknown how well families and life courses have been reconstructed, and what the scale and type of possible deviations from the reality might be. Furthermore, it is unknown how traditional historical demographical database using church registers or population registers as a source compare to databases built from fragmentary observations of individuals—including LINKS. Nevertheless, despite these potential issues for the reconstruction of families and life courses, a comparison to other databases on the Netherlands conducted within this project has shown that family and life-course reconstructions in LINKS are fairly reliable with regard to family size, age at first birth, and other demographic indicators (Van den Berg et al., 2017).

One of the indications of the quality of the data contained in LINKS is the number of persons found on the indexes of birth, marriage, and death certificates that have been linked to other certificates. An overview of the numbers of individuals in the LINKS indexes and databases can be found in Table 1, and an overview of the linking success for the successive generations is shown in table 2. Generation 0 concerns individuals for whom no parents can be identified. Usually, parents' names are included on any civil certificate. Thus, individuals in generation 0 are almost exclusively parents of individuals for whom a birth, marriage, or death is observed. Their children constitute generation 1 and higher. Generations are counted in both the male and female lines and in the longest possible line; that is, the line defined by the number of subsequent generations that can be identified in any line, for all persons found on the indexes from 1812 onward. Thus, the second generation consists of individuals who are linked to their parent's marriage, and their grandparents are known through the marriage certificate of their parents. For a generation 3 individual, there is at least one grandparental marriage linked to their parent's marriage, in either the paternal or maternal line. For individuals who have been linked to the marriage of their parents (generation 2 and upward), there is a follow-up observation for almost four out of five cases (Table 2, Follow-up observation, generations 2–8). For all individuals who are found in any role on any birth, marriage, or death certificate, there is another observation matched for more than 80% of the cases.

As the data are based on indexes of vital event certificates, the end of observation of marriages and deaths in 1937 and 1962, respectively, means that individuals who were childless in 1912, unmarried in 1937, and alive in 1962 were only observed at birth; that is, follow-up is lacking. The cohorts for which observations are most complete can be found in Table 3. As was discussed above, individuals who out-migrate from Zeeland before vital events occur are not found in the civil certificate indexes of Zeeland, and for some individuals a link between their follow-up vital event certificates and birth certificates could not be established. In Table 3, an overview of linking success

for several birth, marriage, or death cohorts can be found. In particular, life courses from earlier-born individuals (1812–62) and births that were successfully linked to the parental marriage cohort of 1812–62 appear to be relatively complete. Marriages have a very high follow-up, with other observations linked in more than 90% of the cases.

Table 1 An overview of the numbers of cases in the LINKS data files

Certificates	N persons
Birth and death certificates[1]	980,538
Marriage certificates[2]	193,793
Unique persons[3]	1,939,916

[1] All persons identified in the indexes of birth and death certificates, linked together when possible.
[2] All marriage certificates.
[3] All unique individuals identified in one or more of the sets of indexes of civil certificates.

Table 2 An overview of the generations and dates of birth, marriage, or death

	N	Birth date	Number of siblings	Marriage date	Own child	Death date	Follow-up observation
Generation 0	889,704	0.00	0.00	0.00	0.90	0.00	0.90
Generation 1	461,813	0.99	1.12	0.32	0.20	0.65	0.81
Generations 2–8	588,389	0.99	5.08	0.37	0.23	0.60	0.78
Total	1,939,916	0.54	1.98	0.19	0.53	0.34	0.84

Source: LINKS, February 2017.
Notes: On birth, marriage, and death certificates, parents are named and in the indexes of the certificates they receive an index number. Persons with an observed vital event are generation 1; their parents—without observations of vital events—are generation 0. Persons with an observed vital event and a linked parent for whom a vital event is observed are generation 2.

Socioeconomic statuses of individuals are measured in concordance with vital events occurring to them or to their offspring. The life-course transitions upon which occupation is measured include individuals' own marriage or remarriage and/or death and the births, deaths, and marriages of their children. Thus, for individuals who marry and have children, there are more observations on socioeconomic status than for individuals with relatively few or no offspring. Individuals who remained single throughout their lives often have only one observation on socioeconomic status, upon

Table 3 An overview of succesful linking of civil certificates by cohort

	N	Birth certificate	Death certificate	Marriage certificate	Linked parent	Linked child(ren)	Number of siblings	Any follow-up observation
Birth cohort								
1812–1862	316,231		0.68	0.32	0.65	0.27	2.2	0.76
1812–1912	681,374	–	0.57	0.35	0.76	0.23	2.7	0.73
1832–1912	565,543		0.54	0.35	0.81	0.22	3.0	0.72
Death cohort								
1812–1912	467,730	0.63	–	0.20	0.51	0.17	1.6	0.72
1812–1962	610,285	0.65		0.29	0.56	0.24	1.8	0.73
1912–1962	139,463	0.74		0.57	0.76	0.45	1.6	0.77
Marriage cohort								
1812–1862	111,994	0.44	0.66		0.27	0.78	1.7	0.90
1812–1912	251,316	0.67	0.64	–	0.54	0.78	2.0	0.93
1832–1912	210,596	0.79	0.64		0.64	0.80	2.0	0.94
Marriage cohort of linked parent								
1812–1862	256,694	0.93	0.73	0.34		0.29	5.8	0.82
1812–1912	551,786	0.94	0.59	0.37	–	0.23	5.5	0.77
1832–1912	462,209	0.94	0.55	0.37		0.22	5.5	0.76

Source: LINKS, February 2017.

death. At the same time, for death certificates of individuals dying in old age, there is not always an observed occupation, as not all individuals were still economically active at advanced ages. Often, the occupation "retired" or "none" is registered for individuals dying at older ages.

In historical research, it is often only the male socioeconomic status or employment status that is taken into account. Women's work is often not included, but not because women were not employed. During industrialization, many women worked in factories and in textile mills. Other women worked in a more traditional occupation—for instance, as a domestic servant—and among farmers and farm laborers women worked on the farm, especially in the busy summer months, and contributed to the household income. However, female occupations are often not registered in documents such as civil certificates or household registers. Therefore, usually, only the socioeconomic status of the male is relied on. Here, in line with earlier work, both the socioeconomic status of the household of origin and the household in adulthood are included, and we also use the male's socioeconomic status for their wives. Moreover, we distinguish between farmer and nonfarmer occupations, due to the different nature of women's labor on farms, access to nutrition, and potential confusion due to lack of detail on the specific agrarian occupation in the civil certificates.

4.2 The Historical Sample of the Netherlands (HSN)

Whereas the aim of the LINKS project has mostly been to reconstruct families and identify family members, connecting lives of individuals into a family network, the main focus of the HSN has been to reconstruct individual lives and households. The Historical Sample of the Netherlands was drawn from birth certificates, after which individuals (research persons, or RPs) have been followed throughout their lives in the municipal population registers. Information from other sources was added, including information from death and marriage registers. The sample focuses on individuals and not on families, although these individuals are followed in a source that centers on households. The consequence is that although information on first-degree relatives as well as nonrelated household members is included in the database, this information is collected from the viewpoint of the research person, or RP—the sampled individual—and usually stops when this individual moves away from the household or after his or her death.

The population registers of the Netherlands are based at the municipal level. The population registers kept a record of each household at their addresses, and individuals were expected to register with their municipality upon moving. In addition, at the start of every ten-year population-register period, each household was visited at home, upon which the employment and religion of the head of the household were registered too.

Changes in the household composition were usually registered either because persons had moved away or through the registration of births and deaths of household members. In the HSN, information about the household is registered, including resident siblings, parents, children, and other related and nonrelated household members, as long as the RP is present in that household. Over time, changes in the composition of the household are registered, as individuals move into and out of households and deaths and births occur. Relationships between the members of the household are included in the registers, albeit from the perspective of the head of the household, often the male spouse of the family. From this information, relationships between other persons present in the household can be reconstructed.

For the purpose of this project, only the portion of the HSN that originates from the population registers is used. In addition, the focus is on Zeeland-born individuals who were born in the period 1863–72, to guarantee maximum overlap between the two databases. For this period, observations from LINKS are relatively complete, and follow-up until death is relatively complete for both databases.

4.3 The Utah Population Database (UPDB)

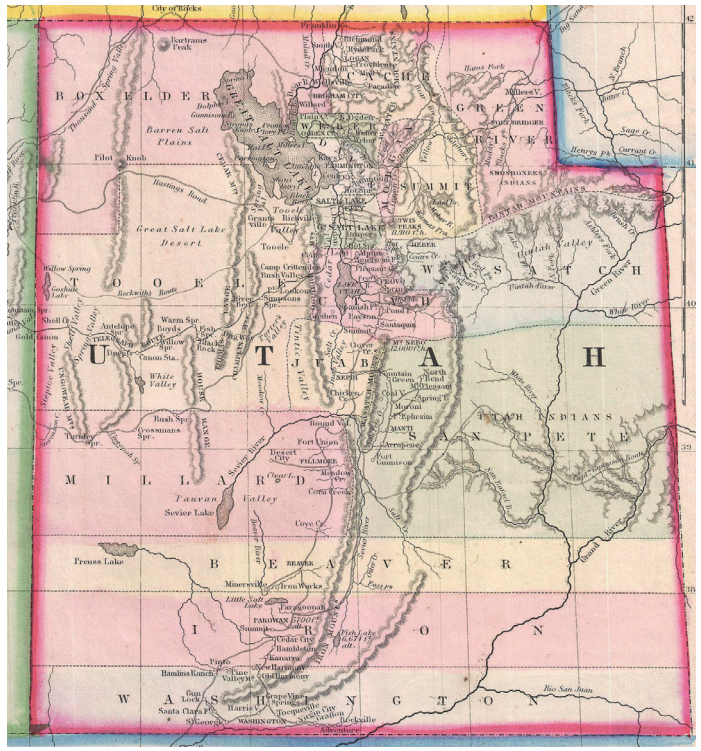
For two of the papers, comparisons are conducted between the results for the province of Zeeland and the State of Utah, US. Here, analyses are based on both LINKS and data taken from the UPDB. The UPDB is a high-quality family-reconstitution database, containing life courses of the Mormon and non-Mormon settler population of Utah. A portion of the UPDB is based on family group sheets on which genealogies of migrants and their relatives to Utah are recorded, as well as their descendants. The families have been linked together into intergenerational family trees with information from other official administrative and medical sources linked in, including birth and death certificates and US Census records. The full database includes over 11 million individuals. Both individuals with and without an affiliation with The Church of Jesus Christ of Latter-day Saints (LDS, or Mormons) are included in the database. The Utah Population Database is comparable to LINKS in its structure and setup, as both databases contain a large number of individuals and multigenerational pedigrees with life-course information linked to spatial information. These features make them highly suitable for assessing both contextual and familial effects on adult mortality.

From 1904 onward, for the UPDB, cause-of-death information is available for persons whose death certificates were found in the state of Utah. On the death certificates, an obligatory mention is made of the cause of death. Although the reliability of historical cause-of-death information is disputed, as conceptions of diseases have changed over time (Risse, 1997), a careful examination of these causes of death may provide us with a deeper understanding of the events that link together early life experiences,

events during the life course, and age at death. The cause of death recorded on Utah death certificates was encoded using International Classification of Diseases (ICD) codes. The work based on the UPDB also capitalizes on the censuses that have recently been added to the UPDB, providing information on an additional group of inhabitants of the state of Utah and information on their place of residence, their socioeconomic status, and other characteristics.

For the comparative work between Utah and Zeeland, some differences in historical context should be kept in mind. The UPDB concerns the state of Utah, a western state in the US, which was settled by migrants from other parts of the US, Europe, and Canada after 1847. Initially, most of the migrants were associated with The Church of Jesus Christ of Latter-day Saints (i.e., the Mormon church), but over time the share of non-Mormons and inactive Mormons has grown. Despite poverty and harsh conditions in the early years of settlement, the population in Utah was relatively

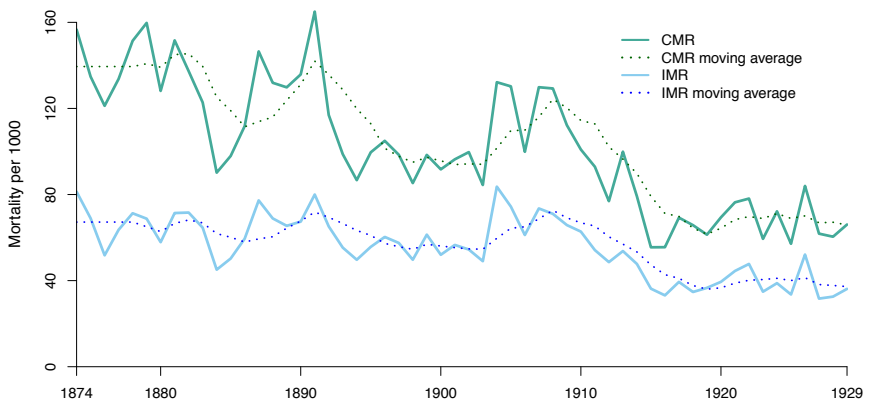
Figure 6 A map of the state of Utah, 1868



Source: County map of Nevada and Utah, Mitchell, S.A, Jr. (1868).

healthy in comparison to Europe, with relatively low infant and child mortality and high life expectancy. Utah has a desert climate, with most precipitation occurring as snowfall during winter; in summer, high-quality drinking water is available in most of the state as runoff from the mountains. During the settlement of the state, several famines and epidemics reduced life expectancy; over time, living conditions improved.

Figure 7 Infant and child mortality in Utah, 1874–1929



Source: Own calculations based on UPDB. Moving averages represent the rolling average mortality rates over five-year periods.

5. A short note on methodology

In the preceding sections, the mechanisms that affect health and survival within families have been described. But how can these effects be distinguished from each other? Whereas scarring effects are likely to affect siblings who were exposed to sibling mortality in particular, effects of family-shared characteristics are likely to affect all siblings in the sibling set simultaneously. It is possible to use the differences between siblings who were and siblings who were not exposed to certain events to disentangle the effect of exposure to these events on survival. By comparing exposed and unexposed siblings to each other, we can control for unobserved factors that may affect both the likelihood of experiencing sibling mortality and siblings' own life expectancy.

The papers that are described here are principally based on quantitative methods, especially survival analysis. The exception is the first paper, which is a review of the literature concerning clustering of infant and child mortality in families. For this paper, a systematic overview of the literature on mortality clustering has been conducted. Literature on the topic of mortality clustering among infants and children has been collected, after which the results have been compared to create an overview of the state of the art on the topic of infant and child mortality within families. For the other papers, quantitative methodologies have been used.

All of the quantitative papers use Cox proportional hazard methods. Proportional hazard models are a type of survival model: they analyze the time that passes until an event occurs and relate that to covariates that multiply the hazard rate. Two different types of Cox proportional hazard models are used: clustered models and stratified (fixed-effect) Cox models. Both types of regression analysis take into account that the included observations are correlated. That is, because infants and children are highly dependent on their families for survival, and because in some families the risk of early death is greater than in other families, the usual assumption from the Cox models that observations are independent is violated. Therefore, in the stratified models that are employed, the standard errors in the models are corrected for the correlated observations (clustered models).

In the second group of models that are applied here, the fixed-effects models, brothers and sisters are compared. This means that unobserved heterogeneity originating from the family of origin is taken into account. In most regression models, observed heterogeneity is taken into account: characteristics of families that may affect the relations of interest are controlled for. Examples include family socioeconomic status, or the number of children in a family. By taking these observed characteristics into account, we hope to estimate a more "true" interpretation of the effects of interest. By using fixed-effects models, we go a step further: not only observed heterogeneity between families is taken into account, but also unobserved characteristics that are shared within families. To that end, siblings who share a family background but may differ in other relevant characteristics—such as exposure to disease and mortality, and marital status—are compared in the analyses. In that way, unobserved family characteristics are controlled for. Example characteristics that we do not observe include, but are not limited to, religion and caretaking behavior. Thus, a stricter test of the mechanisms that link together mortality in families is conducted.

It should be noted that mechanisms that link together mortality exposure and adult survival or children's mortality indirectly—for instance, through trauma and behavioral change of the parents of deceased siblings—will be underestimated using models

based on within-family comparisons. These models will only bring to light differences within the sibling set, and not the factors that affect all siblings simultaneously. Thus, the estimates will be relatively conservative.

6. Results

6.1 Families in comparison: an individual-level comparison of life-course and family reconstructions between population and vital event registers

In the first paper, a closer look has been taken at the LINKS database on the population of Zeeland. Although more than 50 papers have been written on the LINKS database (IISG, 2018), a systematic exploration of its data quality was lacking in the literature. To test the potential of LINKS as a tool in conducting historical demographic research, the database was compared to an established database on the Dutch population: the Historical Sample of the Netherlands (HSN). The analysis capitalized on the sampling strategy of the HSN. Although the body of the information in the HSN is based on the population registers, in which individuals are followed over time and between municipalities and provinces of the Netherlands, the sample on which the project is based was drawn from the birth registers, so that the individuals followed in the HSN were a representative sample of the births in the Netherlands. Because LINKS is based on civil certificates, document numbers for all Zeeland-born individuals in the HSN match those for Zeeland-born individuals in LINKS. As the databases partially observe different years, a ten-year birth cohort from the middle of the observation period of LINKS is used. A total of 495 persons from the 1863–72 birth cohort are identified in both the HSN and LINKS.

For these individuals, their life courses and families are reconstructed based on information present in the two databases, which is derived from the population registers (HSN) and the civil registers (LINKS). Observations in the population registers are partially continuous—with regard to place of residence—and partially more regularly updated than observations from the civil registers, as every ten years, in a process akin to the censuses, house visits were used to update population registers. In the civil registers, observations only occur in concordance with vital events, such as births, deaths, and marriages that are linked to individuals because they either concern them or one of their children. Other major concerns with regard to family reconstruction based on civil registers are issues of false linking—due to similar names and/or places of birth and marriage—and missed links due to changed or misspelled names, the cessation of follow-up due to out-migration of Zeeland-born individuals, and the exclusion of illegitimate births. Issues with research based on the population registers include the oversight of brothers and sisters who never lived together with the sampled research persons and the failure to include stillbirths.

Indicators of fertility, marriage, mortality, and measurements of occupational status of individuals found in both databases and their siblings are compared, based on LINKS on the one hand and on the HSN on the other. Two strategies are used. First, all research persons are compared between the databases in their demographic indicators. Second, the same indicators are used, but not for *all* individuals—rather, for *the same* individuals in both databases, to see whether life courses are reconstructed in a similar way. It turns out that there is more parent information available in the HSN than in LINKS. However, considerably more siblings and spouses are identified in LINKS, as siblings who died young are often not included in the HSN. In cases where a marriage is observed in LINKS—through a marriage certificate—sibling information is usually more complete than in the HSN. At the same time, out-migration is not part of the LINKS data. Within-Zeeland migration has been shown to not affect the linking success, but for out-migrants that will not be the case, and they have been shown to downwardly bias the ages at death found in LINKS: individuals reaching adult ages and out-migrating from Zeeland are less likely to be linked to their birth certificate. Thus, vital events occurring outside the province of Zeeland will be an essential addition to future versions of LINKS. Furthermore, there are more individuals without a known socioeconomic status in LINKS than in the HSN, suggesting that the linking of population-register occupation information can be a valuable addition to the LINKS-based family reconstructions.

Despite these issues with specific subsets of individuals in HSN on the one hand and LINKS on the other, the life-course reconstructions and family reconstructions reflect each other relatively well, if information is known. We conclude that it is necessary to consider the questions for which LINKS and the HSN are used with care: LINKS provides more complete family information on siblings and parents, whereas the HSN provides more complete life-course information, especially for individuals who migrate out of Zeeland.

6.2 Early-life mortality clustering in families: a literature review

Research on contemporary and historical populations has shown that infant and child mortality tend to cluster in a limited number of high-mortality families that experience a far higher number of deaths than the average family in the population. Since the first literature came out attributing a large share of infant and child mortality to these high-mortality families, there has been an increasing interest in the phenomenon in a variety of social sciences. This paper was the first to provide a review of the phenomenon of mortality clustering.

In the literature, four approaches to mortality clustering are commonly used—approaches that use sibling deaths, mother heterogeneity, a set threshold, and differences between observed and expected numbers of deaths in the population—to assess whether mortality clustering occurs. In the first approach, the death of a sibling is incorporated in models to assess the effect on survival of siblings. Sometimes, the correlated death risks of the other siblings are accounted for. An alternative approach focuses on mother heterogeneity in the number of children that are lost in infancy or childhood, using the variance of the random intercept in multilevel models, usually in survival analysis. However, in both approaches, it is not taken into account that in larger families, the risk that at least one child will die is greater than in smaller families.

In a third approach, a threshold approach is applied to distinguish between low- and high-mortality families, after which logistic regression analysis is used to model explanatory characteristics. The threshold is defined quite varyingly: from criteria such as sibling loss, multiple child loss, or a combination of size of the family and population death rates. Although the latter approach is quite intuitive and takes fertility and mortality patterns in the population into account, there exists an operationalization effect that may bias results. For instance, in a population with a 25% mortality rate, families that lose more than half of their children before their fifth birthday could be defined as high-mortality families in Zeeland. Families with ten children have to lose five children to fulfil this criterion; for a family with four children, the loss of only two children suffices to fulfil that criterion. Thus, the criterion is relatively more strict for larger families than for smaller families. Fourth and finally, some papers focus on the difference between observed and expected numbers of deaths in a population. This approach explicitly takes the binomial chance distribution into account—larger families are more likely to experience deaths—and can, depending on the approach, give insights into the family-level characteristics that create these patterns. At the same time, they often do not allow for the inclusion of child-level characteristics.

The diversity of conceptual and methodological approaches is joined to a set of theories that concern clustering in infant and child mortality. Exposure to infectious disease and death may result in damage to the bodies and immune systems of surviving siblings, increasing the likelihood of further deaths. Depression and stress among bereaved parents may have detrimental effects on remaining children. In natural-fertility populations, birth intervals are shorter after the death of an infant because fecundity is restored more quickly after the cessation of breastfeeding; the decision to replace the lost child may also contribute to shortened birth intervals. Maternal depletion—which is related to pregnancy complications, preterm delivery, and low birth weight—and resource competition between siblings in larger households may therefore contribute to mortality clustering. For modern-day India, there is some proof for a role for sex preferences.

Furthermore, life-history theory predicts that individuals exposed to high mortality will reproduce early, have shorter birth intervals, and invest less in their children, which is essentially predicting mortality clustering among mothers exposed to high-mortality environments (Störmer & Lummaa, 2014). Other explanations have been sought in socioeconomic status and childcare practices. Inadequate feeding patterns of mothers, especially poor mothers, may have resulted in repeated problems during delivery, repeated stillbirths, and high maternal mortality. Das Gupta (1997) found more diversity in parenting behavior among lower-educated mothers, suggesting that maternal care and socioeconomic status overlap to some extent. As childcare plays a large role in mortality clustering, it is not surprising that the death of a parent and remarriage of the parent matter for child survival.

Some of the earlier work on mortality clustering has considered the extent to which mortality clustering runs in families beyond a single generation. Does intergenerational transmission of mortality clustering exist, and if so, what mechanism can be attributed? It was found that, indeed, there exists some intergenerational transmission in both the maternal and the paternal line, but the explanatory mechanism has remained rather unclear. Do the families live in adverse contexts and endure poverty for multiple generations? Or do specific mechanisms related to exposure to mortality shape mortality across generations in these families? Then, the second-generation and adult mortality rates are affected through different mechanisms than the first generation, unlike explanations related to characteristics that exist across generations. These issues are similar to that of studying mortality clustering in general: Are general explanations for mortality predominant, and is early-life mortality concentrated in certain families because these families are large, or have bad luck? Or are there specific explanations for infant and child mortality in these high-mortality families?

6.3 Like mother, like daughter: intergenerational transmission of mortality clustering in Zeeland, the Netherlands, 1833–1912

In this paper, a distinction is made between two types of explanations for intergenerational continuities in infant and child mortality in families: explanations that work through shared characteristics of grandmothers and mothers, or through exposure to mortality in the family of origin, affecting the life course and child survival. On the one hand, characteristics of families that are relevant for infant and child survival in several generations, such as socioeconomic status, childcare behaviors, and others, may run in families. These characteristics may occur in several generations as the result of

a process of social reproduction, and affect infant and child survival across these generations. On the other hand, exposure to mortality itself can also lead to increased mortality in several generations. The health of surviving children may be damaged, including the reproductive health of surviving daughters, caretaking capabilities, and earning capacities.

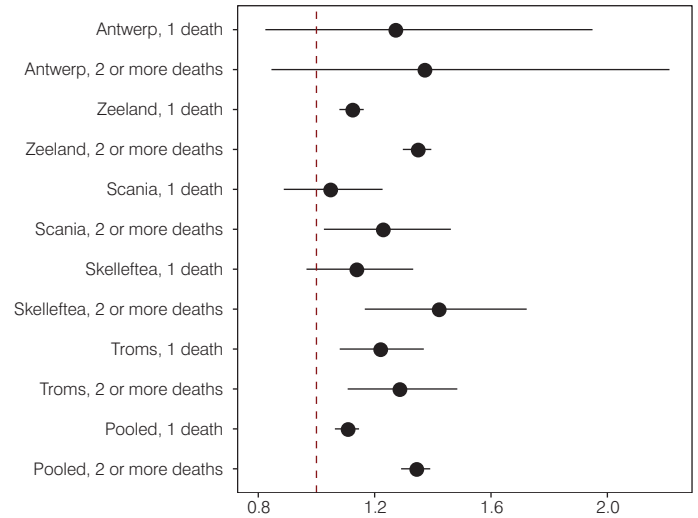
Alternatively, evolutionary biology emphasizes that exposure to childhood adversity affects life histories, including reproductive careers and infant mortality. Individuals who are exposed to childhood hardships, such as food shortages, or stressful events in childhood, including mortality of siblings, are expected to employ fast life strategies for reproduction, to increase the chances of successful reproduction in risky environments. These exposures are therefore related to a lower age at menarche, a lower age at first marriage, a lower age at first birth, and shorter-spaced births of children (Nettle et al., 2011; Störmer & Lummaa, 2014). Furthermore, as part of this strategy, individuals are expected to invest fewer resources in individual children and experience higher infant mortality (Chisholm, 1993; Störmer & Lummaa, 2014). Through the reproductive career of the mother, grandmothers' experiences of infant mortality may be linked to survival of their grandchildren.

A combination of two modeling strategies was employed. First, I checked whether or not there is indeed a stronger concentration in families than can be expected given the binomial chance distribution, family sizes, and mortality levels in the population (see Section 6.2). It was found that, indeed, that was the case: there are more families experiencing a large number of deaths than can be expected. Next, survival models accounting for mother heterogeneity were employed to assess the mechanisms behind this concentration. The risk of infant death was considerably greater if the grandmother lost one infant—12%—and this increased to a 34% higher risk if the grandmother lost two of her children in infancy. These figures are controlled for the demographic characteristics of the family (the family size of the grandmother and mother, the birth order of the mother, and the age of the mother). Without controlling for these characteristics, the risk of losing an infant is even further elevated among mothers born to grandmothers who experience multiple infant deaths.

The effects do not appear to be especially relevant for boys versus girls, which suggests that in Zeeland sex-specific neglect or biological mechanisms affecting one of the sexes in particular do not play a role. Infants from high socioeconomic status groups have a better survival prospect, and the grandparents of infants matter too: over and above the effect of having prosperous parents, the prosperity of grandparents also matters for infant survival. At the same time, these socioeconomic status characteristics do not explain why the infant mortality risk is elevated in some families over the generations. This also applies to parental survival, which does not appear to play a central role in linking together the survival of infants across generations.

The work that has been done for this paper was part of a larger project together with research groups from Umeå and Lund (Sweden), Leuven (Belgium), and Troms (Norway), in which the same approach has been chosen and the same modeling strategy applied for the first parts of the papers, after which the focus has been on more specific explanations for each region. In all regions, there is indeed intergenerational transmission of mortality clustering. These effects are stronger for grandmothers who lost two or more children than for grandmothers who lost one child in infancy. Furthermore, the effects are not explained by the demographic characteristics of the family. Because of the similar procedures for data selection, differences in results between populations caused by differences in data construction between populations can be refuted, although the source material may still introduce differences in included cases.

Figure 8 The effect of maternal infant deaths on the death risk of her daughter's infant children in five historical European populations



Source: Quaranta et al. 2017.
Note: Pooled estimates are based on a pooled database containing observations for five populations (Antwerp, Zeeland, Scania, Skelleftea, and Troms) for individuals born between 1845 and 1901.

6.4 Repeating a lethal experience: mother's and father's exposure to sibling mortality and offspring mortality

In this paper, I take a closer look at intergenerational correlations in the likelihood of early-life mortality. The paper innovates in comparison to the earlier literature in several ways. First, I take a closer look at causal mechanisms by conducting within-family comparisons that focus on exposure to mortality. The results are compared to analyses that focus on between-family comparisons. Second, in addition to infant mortality, I take a closer look at all under-five child mortality. Third, whereas in earlier work the focus lay on correlations between maternal grandmothers and mothers, here, fathers are included alongside mothers. Results are compared between historical populations from Utah (US) and Zeeland (the Netherlands).

Broadly, there are two groups of mechanisms that may play a role in multigenerational continuity in high mortality. First, family characteristics that are related to infant and child mortality levels are shared between generations, including socioeconomic status, demographic characteristics such as fertility, and social and behavioral characteristics (Anderton, Tsuya, Bean, & Mineau, 1987; Lyngstad & Prskawetz, 2010; Wolfinger, 2011; Carvalho, 2012), as well as genes. Second, exposure to mortality in the first generation may causally affect young-age mortality in the second generation through alterations in exposed individuals' behavior and preferences, and effects on their health and survival. While the first group of explanations predicts an intergenerational relation between all deaths occurring in the sibling group and mortality in the second generation, the second view on intergenerational effects of infant and child mortality suggests that it is *exposure* to sibling mortality that is relevant in predicting mortality in the second generation. In this paper, I take a closer look at intergenerational consistencies in early-life mortality, with the aim of disentangling the mechanisms behind them.

To distinguish between the mechanisms that may affect intergenerational continuities in infant and child mortality, I use a combination of approaches. I look at the role of the father versus the role of the mother in affecting offspring mortality; test whether socioeconomic and demographic characteristics of the family explain the increased likelihood of offspring mortality; and conduct within-family comparisons to assess the effect of exposure to sibling mortality on offspring's survival. Within-family comparisons control for unobserved heterogeneity, but are also more likely to underestimate the total effect of sibling deaths, as some deaths may affect the whole family simultaneously.

We find that the mother's and father's exposure to sibling mortality, in all stages of their childhood, is significantly and negatively related to the chances of survival of their offspring between birth and age 5. With increasing numbers of parents' sibling deaths, the likelihood of mortality among their children increases further, providing evidence for a dose–response pattern. Mother effects tend to be stronger than father effects in both Zeeland and Utah, but for Utah the differences are not significant. Furthermore, the mother and father effects are additive but not multiplicative. That is, father's and mother's sibling deaths contribute to higher child mortality in models in which they are included simultaneously as well as in analyses where they are considered separately, but the effects do not interact: the effects add up, but neither exacerbate nor compensate each other.

The effects are partially explained by the sociodemographic characteristics of the household, most importantly by birth order of the child, the birth interval, and the young age at first reproduction of the mother and the father. Similarly, within-family comparisons, which account for unobserved family characteristics and assess the effect of exposure rather than all sibling mortality in the household, show that maternal exposure affects child survival more consistently than paternal exposure. For fathers, only strong exposure to sibling mortality that occurs after early childhood affects the survival chances of their children. For men and women alike, effects are found for both infancy (age 0–1) and early childhood (age 1–5), but parental mortality exposure affects survival of infants more strongly than survival of children.

This study is the first to show that in addition to mother´s experiences with sibling mortality, father´s experiences with sibling mortality are significantly related to offspring survival between birth and age 1, and between age 1 and age 5. Earlier research looked exclusively at infant mortality. Here, I also find evidence for an intergenerational link in child mortality, illustrating that mortality chances are not only related to processes of vulnerability soon after birth, but also by exposures and deaths after infancy. At the same time, effects are stronger in infancy than in early childhood, between the first birthday and age 5. Furthermore, although child mortality on both the father's and the mother's side is related to increased mortality rates for their offspring, maternal effects are significantly stronger than paternal effects, indicating that mother-linked mechanisms are likely at play in the intergenerational transmission of infant and young-child mortality.

6.5 The long harm of childhood? Childhood exposure to mortality and subsequent risk of adult mortality in Utah and the Netherlands

In this paper, I take a closer look at the adult survival of individuals from high-mortality households. The findings are as follows. In line with earlier work by Finch and Crimmins (Finch & Crimmins, 2004; Crimmins & Finch, 2006), I show that there is an association between exposure to mortality in communities and birth cohorts and

subsequent adult mortality rates. Especially for infant and all-child mortality, effects on adult mortality rates are found, which is in line with some earlier research (Schellekens & van Poppel, 2016) but contradicts other findings (Crimmins & Finch, 2006).

Exposure to sibling deaths is related to later-life stress (Norton, Hatch, Munger, & Smith, 2017) and increases mortality in early and late childhood and early adulthood (Yu et al., 2017). In line with these results, it is found that sibling deaths are related to reduced survival at adult ages. Indeed, exposure to sibling deaths has a stronger effect on adult mortality than community-level exposure to mortality. After taking family-level characteristics and individual-level characteristics into account, exposure to sibling deaths has a lasting effect on adult mortality rates. Community-level mortality has a lasting effect on adult mortality too, but the effects are weaker and are reduced by including family characteristics.

No evidence is found that weak mortality crises have a selection effect that overwhelms scarring effects. On the other hand, scarring effects appear to dominate selection effects, even if there is only a single exposure to sibling death (Hatton, 2011). Despite the potential ameliorating effects of immunization and access to modern health technologies and care, including antibiotics, even a single exposure to a sibling death in Utah is related to adult mortality rates. In Zeeland—where child mortality was very high in the period that we study—exposure to multiple sibling deaths is related to adult survival, while exposure to only one sibling death is not significantly related to the adult mortality risk. Possibly, in environments with high mortality, the event of a sibling death does not necessarily contribute much to adult mortality risks, as individuals are likely to be exposed to deaths in childhood at any rate. In addition, considering that adult life expectancies were lower in Zeeland and adult mortality rates were high, the cumulative damage done by exposures is stronger in Zeeland than in Utah, which may have led us to find lower effects of specific exposures of events such as sibling deaths.

In contrast to earlier work (Hayward, Rigby, & Lummaa, 2016), which found no link between exposure to infectious disease and all-cause mortality, deaths due to cardiovascular disease, cancer, and stroke using a birth-cohort approach, here, it is found that exposure to all-cause sibling deaths is related to all-cause mortality, cardiovascular disease, cancer, diseases of the respiratory, digestive, and nervous systems (women), and metabolic disease (women). Thus, exposure to all-cause early-life mortality among siblings is related to an array of causes of death, and appears to be a general risk factor for earlier adult mortality. In addition, the risk of death due to these diseases follows a dose–response pattern. The extent to which individuals are exposed to sibling mortality matters for adult survival: the higher the number of sibling deaths, the stronger is the effect on adult all-cause mortality and cause-specific mortality. On the other hand, exposure to sibling deaths due to infection is related to cardiovascular disease and metabolic diseases for men and women. Thus, deaths from infections appear to be specifically related to chronic inflammatory responses. Sibling all-cause mortality is related to a wider array of causes of death, suggesting that there may be a mechanism at play beyond inflammatory immune responses relating early-life exposure to mortality to adult mortality rates.

7. Concluding discussion

In the Dutch province of Zeeland, early-life mortality was distributed unequally over families in the population. A large number of infant and child deaths were concentrated in a select group of high-mortality families. Within these families, a much larger number of infants and young children died before their first or fifth birthday than could be expected given the infant and child mortality rates in the province at the time. That families play such an important role in the survival of infants and small children is not necessarily surprising, as infants and young children are strongly dependent on their families for their health and well-being. What is surprising, however, is that for the surviving children from these high-mortality families, the effects of their families on health and survival was not limited to childhood. Health and survival were affected negatively beyond the childhood years, and the influence of the familial component on mortality was so strong that even *their* offspring had an increased likelihood of death in infancy and early childhood. Thus, individuals exposed to childhood adversity in their families of origin faced a life-long reduction in health and survival. Furthermore, concentration of childhood mortality in families can at least partially be attributed to family histories of high mortality across generations.

Consistencies in family mortality patterns across the life course and generations can be attributed to two groups of mechanisms. On the one hand, across generations families share characteristics that increase their susceptibility for infant and child mortality. Mortality and health are not solely personal: to a large extent, they are socially, economically, and culturally determined, and many of the factors that play a key role in determining health and survival affect the whole family. These characteristics are transmitted across generations, through processes of social reproduction—socioeconomic status—and socialization—cultural characteristics, norms, and behavior that are transmitted to the next generation. Thus, individuals originating from families with high infant and child mortality live shorter lives and their children have an increased risk of early death due to continued poverty or the continuation of behavioral characteristics that are adverse for health and survival.

On the other hand, in conjunction with mechanisms related to shared characteristics, it is also *exposure* to sibling mortality and disease itself that explains its continued effects across generations. Surviving siblings may have been affected in their health and reproductive behavior, which then in turn affected the survival of the next generation. Evidence for exposure-related effects is found in within-family comparisons. Here, two or more siblings are compared in their adult life expectancy and in the likelihood of their children to survive childhood. These siblings came from the same household, but differed in their exposure to sibling mortality: of the sibling set, one or more was exposed to sibling mortality, and the other was not. These siblings were then compared with regard to the chances of their children surviving their first years of life. Siblings who were exposed to mortality in their childhood had children whose chances of mortality in the first five years of life were greater than for children of siblings without such exposure. This finding illustrates that it is not just the characteristics of the family, but also exposure to mortality itself that affects survival chances in adulthood and the second generation.

The link between mortality exposure in the family of origin and full life-course mortality in adulthood was explored for the first time in the current work. Earlier research has shown that there is a link between the childhood mortality of a cohort and the survival of that cohort in adulthood. That link has been fiercely disputed in the literature, with some evidence in favor of a link between high mortality in the cohort and life expectancy and in other cases no such evidence. Most of the research into the connection between mortality exposure in childhood and survival in old age has been studied in birth cohorts. However, young children are highly dependent on their families, and therefore exposures to childhood adversity are mainly shaped within their families of origin.

Sibling mortality to which individuals were exposed in their families of origin affected adult mortality rates well into adulthood, between age 18 and age 85; and the more sibling mortality individuals experienced during childhood, the higher the relative mortality rates were in adulthood. Further insights can be generated by investigating the causes of death of siblings who died in childhood as well as the causes of death in adulthood of their surviving siblings. Exposure to infectious disease in childhood in particular may have a negative effect on survival in adulthood due to a chronic inflammatory response. For the population of Utah, the causes of death are known. It is therefore possible to distinguish between sibling deaths due to infectious disease and due to other causes. Exposure to sibling death due to infection is specifically related to chronic inflammatory responses and cardiovascular disease and diabetes. At the same time, all sibling deaths—both those due to infection and those due to other causes—are also found to be related to an array of causes of death and

functions as a general risk factor of mortality, in line with ideas of scarring effects of disease and mortality exposure.

Thus, results for outcomes at older ages show that at least part of the scarring processes in families is related to sibling infectious disease exposure. These are effects that cannot be explained by family shared characteristics, because differences in survival were observed between children who had already been born when their siblings passed away and children who had not yet been born. These findings suggest a causal effect related to exposure, rather than to an indirect effect working through parents or family-unobserved characteristics affecting the mortality of all children embedded in a household. Furthermore, no evidence was found that mortality selection occurs, in the sense that mortality exposure only negatively affects relatively frail individuals within families or cohorts, as survival did not improve in any period of adulthood for individuals who had experienced higher mortality in their families of origin. Rather, the evidence is in line with scarring mechanisms, which could work through stress but also through inflammatory immune responses. Of course, this does not exclude some degree of selection within these families, as scarring and selection effects may co-exist: while relatively frail individuals are selected out from the set of siblings, the health of surviving children is damaged by disease and mortality exposure. Thus, the net effect of mortality in the sibling set on survival is negative rather than positive.

Sibling mortality scars the health of surviving siblings, who have a higher mortality risk in adulthood. For young children, similar results have been described in the literature: sibling mortality has a scarring effect on the health of their infant brothers and sisters. Due to methodological constraints, research was often limited to the effects of infant mortality on their younger brothers and sisters. Here, it is shown that the health of older brothers and sisters was also affected by exposure to sibling health. It is likely that this is also part of the explanation why deaths cluster in certain households. The health of all children in those families is affected by exposure to infectious disease, and possibly also by parental stress, trauma, or psychological detachment following child loss.

Exposure to mortality and disease in childhood does not, however, only affect healthy life expectancy and mortality in adulthood. Life-history theory, originating from evolutionary biology, argues that individuals originating from environments with higher mortality risk experience earlier menarche, start participation in sexual behavior earlier, and commence reproduction younger than peers from more favorable, stable backgrounds. Cues about danger in the environment are, amongst others, taken from the absence of fathers, mortality rates in the population, and

sibling mortality. For example, Störmer and Lummaa (2014) linked high mortality in the family to both brothers' and sisters' age at first marriage and age at first birth. These responses to unsafe conditions in childhood are not a rational choice; rather, they are the result of subconscious processes that affect characteristics such as sexual risk-taking. These tendencies may have developed in an evolutionary context in which, in high-mortality environments, faster and riskier life-history strategies are linked to greater reproductive success. Thus, on average, in risky environments, reproducing, even in unsafe conditions, was more beneficial than postponing reproduction. The developed tendencies are, however, not necessarily beneficial in a more modern context. Younger and faster reproduction, as well as investment in quantity rather than quality of offspring, was often related to higher mortality among one's children. Short birth intervals are related to maternal depletion syndrome, with adverse consequences for both mother and child; and increased competition over scarce resources—including parental time, food, and medical attention—in young, large families may be detrimental to the survival and health of children.

Here, evidence was found for an effect of exposure to adversity in childhood and earlier and faster reproduction, which in turn contributed to higher infant and child mortality. Mothers and fathers originating from high-mortality families experienced higher mortality among their offspring. This relationship is partially explained by a younger parental age at first birth (regardless of the age at birth of the index child), the birth order of the child, and the birth interval to the preceding child. Fathers and mothers from high-mortality families of origin had slightly shorter birth intervals, had on average more children, and had commenced reproduction younger, and these characteristics contributed to a higher likelihood of early mortality among their offspring.

Alternatively, in a subset of the population, family-shared sociocultural mechanisms could link together high child mortality across generations. Women from disadvantaged environments, characterized by early fertility and high mortality, may be socialized to pursue disadvantageous fertility careers, leading to high mortality in several generations (Laslett, 1980). A further possibility is that in some social, cultural, or religious groups, such as Catholics, early marriage, short birth intervals, and large families were more common than in other segments of the population. This may have resulted in a similar pattern to that found here: higher infant and child mortality in families of origin and among one's own offspring, short birth intervals, and early first reproduction. However, this alternative explanation would predict this pattern for a whole sibling group, which contrasts with the findings in the current work. Fast life history predicts, and the findings here confirm, a high-mortality pattern especially for individuals who had been exposed to sibling mortality, rather than for all siblings within a sibling set.

A further disentanglement of these mechanisms is a complicated challenge, which could not be pursued in the context of this dissertation. Nevertheless, it should be emphasized that these mechanisms do not necessarily exclude each other. Mechanisms of socialization and cultural norms could help shape subconscious preferences and fertility behavior, and play a role in linking together high-mortality environments in childhood and fertility careers in adulthood.

Mortality histories in families of both men and women contribute to the likelihood that children will die in early childhood. This finding supports the notion that fathers are important for the survival of their offspring. In most research, the role of the father has been reduced predominantly to his socioeconomic contribution as the main breadwinner of the household. For the mother, usually, a more extensive set of her characteristics that possibly affects the survival of her offspring is taken into account, which besides her education and occupation encompasses characteristics such as her reproductive history, the age at birth of her children, and religious characteristics that are important for infant survival because of their relation to breastfeeding.

The relative negligence of men in research into infant and child survival is possibly related to the perception of male and female roles within the household, with females as main caretakers and men as main breadwinners. However, it may also be an oversight, leaving out potentially large effects of the male presence on child survival and well-being. Here, for the first time, the role of the father in intergenerational consistencies in high infant and child mortality is evaluated in conjunction with the role of the mother. Both men's and women's family mortality histories predict the likelihood of their children to survive their first year and their first five years of life. Nevertheless, children's survival is more strongly affected by their mother's family history of high early-life mortality than by their father's family history of high mortality.

It was expected that parental exposure to sibling disease and mortality in childhood and the likelihood of his or her children to survive their own childhoods would be linked by parental health in adulthood. As was discussed above, individuals growing up in high-mortality households had a higher mortality risk in adulthood. Their children, in turn, may have been negatively affected by their parents' adverse health, disease, and possible absence from the household, and by their reduced ability to care for young children, as well as an untimely death. This could, at least theoretically, explain part of the link between high mortality in the family of origin and the likelihood that individuals' own children would die in childhood. In reality, however, parental death does not always affect child survival negatively. Research on the effect of parental mortality on child survival has shown that although maternal death has a detrimental effect on her young children's survival, the picture is more mixed for

paternal death. In some populations and time periods, survival of the father had strong effects on his children's survival and well-being, but in others the effects were small or nonexistent. Possibly, mothers are more crucial for small children, as they are their main caretaker. Also, mothers may have been able to access assistance and poor relief more readily than men, thus allowing for assistance in caring for their children. In line with these expectations, for Zeeland there is a negative effect of maternal death on the survival of young children, but for paternal death no such effect is found. Furthermore, in contrast to the expectations, the effect of maternal death on child death does not contribute to explaining why children of mothers from high-mortality families of origin have a higher risk of mortality.

With regard to characteristics of the family besides mortality exposure, socioeconomic characteristics contributed to the likelihood that infants and children would survive their first years of life. Infants born to farmers in particular had a relatively advantageous position and a lower likelihood of death before their first birthday. A similar result is found for the highest socioeconomic status individuals: their risk of child loss is lower than that of lower-status groups. Interestingly, that does not only apply to the socioeconomic status of the household in which the children grow up—based on the occupation of their father—but an independent effect is found for the maternal grandparents: if the mother originates from a high socioeconomic status family, the children's survival chances are better, independent of the mother's and father's own status. This positive effect of the maternal grandparents' resources on child survival may be the result of a direct economic contribution of the maternal grandparents to their daughters' households, thus reducing the grandchildren's mortality risk. Maternal background may matter more than paternal background due to the close involvement of mothers with the well-being of their daughter's offspring, while the maternal involvement with daughters-in-law is in many populations more limited. At the same time, socioeconomic status plays only a minor role in explaining variation between mothers in the numbers of early-life deaths among their children.

If socioeconomic status is important for survival in infancy and childhood, it is also likely that its effects on health can be identified in later life. These long-term effects relate to the same mechanisms that affect its short-term effects on health: in poor families, early-life conditions related to nutrition, disease exposure, and overcrowding negatively affect children's health. These health effects are likely to persist into adulthood. For adult mortality, the socioeconomic status in the family of origin as well as one's own socioeconomic status in adulthood are related to a lower mortality risk. This finding suggests that childhood conditions have a lasting effect on the health of individuals throughout their lives. Nevertheless, socioeconomic status and poverty-related mortality do not explain the relation between exposure to sibling mortality and survival in later life, or the intergenerational relation between high infant and child mortality. Although these characteristics contribute to the overall risk of mortality, their effect exists independently from the effects of sibling mortality exposure on health. Thus, socioeconomic status appears to play a relatively minor role in explaining family histories of high mortality, and infant and child mortality clustering does not result from an increasing concentration of infant and child mortality in families that are poor over time.

With regard to gender, no evidence was found that boys and girls were affected differently by their family history of mortality. Earlier studies on sex-specific mortality rates in 19th century Western Europe found some excess mortality among girls for certain age groups (Devos, 2000; Poppel et al., 2009; Janssens et al., 2010). In their families, there may have been a parental preference for one of the sexes, or behavior that led to an increased chance of mortality in early life for one of the sexes specifically. Because such preferences are shaped through cultural processes and processes of socialization, it could then be expected that gender preferences would occur in several generations of the same family. However, girls and boys appear to be affected similarly by patterns of excess mortality in their families. No evidence was found for patterns of female excess mortality. Possibly, the children that were studied here were not part of the affected age groups; that is, girls in the early teenage years. It therefore remains unclear why female adult mortality is more strongly affected by childhood sibling mortality than male adult mortality. Possibly, gender norms with regard to caretaking played a role: girls may have been expected to care for ill siblings (Van Poppel, 2000), more so than boys. Because of their increased exposure to sick siblings, girls' likelihood of being infected with disease could have been higher than among boys; alternatively, there may be sex differences in susceptibility to disease. The stronger effect of childhood sibling mortality on females' adult survival could have resulted from these sex differences.

Finally, it is important to consider the role of context in the processes that were studied here. Two populations were considered: historical Zeeland, with sky-high infant and child mortality rates in the 19th century, and historical Utah, with much more favorable levels of infant and child mortality and higher mean ages at death. For both populations, it was found that exposure to sibling mortality and disease affected survival in adulthood. Therefore, vaccination and modern health technologies, including antibiotics, only play a minor role in reducing the negative effects of exposure to mortality in early life. In Utah, where mortality was relatively low, even a single exposure to sibling mortality was related to lower survival in adulthood. For high-mortality Zeeland, only relatively strong exposures within the family of origin had a distinguishable effect on adult mortality. Possibly, as the mortality in the Zeeland

population is very high, individuals are generally more strongly exposed to disease and mortality, and therefore it is harder to pick up the effects of a single exposure in the family. Furthermore, for both Zeeland and Utah, intergenerational effects were found in child survival, showing that despite widely varying mortality contexts, the role of fast lives and reproductive choices on child survival should not be underestimated.

A broad array of possible explanations for the role of the family in explaining high mortality across the life course was considered in this dissertation. Not all alternative explanations could be included in the studies. For instance, children who lose many siblings in infancy and childhood have a smaller social network in adulthood, which limits their possibilities of accessing resources and kin assistance when they are raising their children or growing old. As relatives other than mothers and fathers may significantly contribute to child well-being and survival (Sear, Steele, McGregor, & Mace, 2002; Sear & Mace, 2008), a smaller kin network may negatively be related to their survival, or the survival of their offspring. Furthermore, unhealthy environments could play an important role. Here, access to safe, unpolluted drinking water can be problematic, or poverty may be rife, negatively affecting the survival of inhabitants. Adversity in the living environment was likely relatively stable over time. In Zeeland, industrialization occurred relatively late, and water and sewage systems were covering significant parts of the population only from the end of the 19th century. If individuals originating from such a region stayed there—or in another region with adverse living conditions—throughout their lives, this may have affected the survival of their siblings, their own life expectancy, and the chances of survival of their own offspring in infancy and childhood. In other words, spatial proximity may be the underlying, true cause of intergenerational similarities. In the studies that were conducted here, spatial aspects were not a central topic of study. Two steps were taken to involve the region of residence of individuals. First, in one of the studies, mortality levels in the municipality of birth were included, with the purpose of assessing how exposure to mortality affected survival in adulthood. Here, I assessed how exposure to mortality in regions compared to exposure to mortality in the parental household. In particular, mortality in the household played an important role in affecting survival in adulthood, much more strongly than mortality levels in the municipality of origin.

Another measure to ensure that the observed effects that were ascribed to exposure to mortality were not, in reality, caused by unobserved characteristics of families or regions in which those families lived relates to the modeling strategy. Frailty models were employed that compared siblings that were and siblings that were not exposed to mortality of a brother or sister. It was found that unobserved heterogeneity played

a role in explaining intergenerational consistencies, but at least part of the relations in increased mortality in families is related to *exposure* to mortality, rather than membership of a specific family with specific attributes shared collectively. This is a strong indication that the observed relationships are not artifacts of a process that, in reality, occurs at the regional level, but are in fact the product of processes that work within and between families in that region. Of course, exposure to mortality, disease, and adverse living conditions occurs in regions as well as families, and could at this level be even more detrimental than such exposures in the family of origin. For future research, it would be interesting to consider the exact role of the living environment *vis-à-vis* family relations in high mortality.

Furthermore, in the studies on the effects of exposure to mortality and infectious diseases in childhood on the later life course, it remains unknown whether individuals were truly exposed to the sicknesses of their siblings and to their deaths, as there is no information available on the specific household composition around the time of the event. It is unknown which siblings were infected and fell sick. At the same time, earlier research into early-life effects of exposure to disease has often relied on nationwide statistics on excess mortality due to specific diseases. The use of within-family mortality as an indicator of exposure to mortality and disease is much more precise than these statistics, which are not necessarily meaningful for certain regions or households.

The dissertation has employed a relatively novel approach, using a large database—"big data"—on a full population of a province of the Netherlands. This data has become available through the LINKS project. In the context of this doctoral project, the quality of this database has critically been assessed, both by looking into internal consistencies—that is, are life courses logically reconstructed?—and comparisons between ages at death for individuals whose life courses and families were reconstructed in comparison to the whole population of Zeeland, based on all vital event registers. Mortality peaks in the selected samples were found to reflect mortality peaks in the population rather well. Furthermore, the family reconstructions were compared to manual family reconstructions based on household registers contained in the Historical Sample of the Netherlands. It was found that LINKS is more complete with regard to kin networks in comparison to the HSN, but the HSN has more complete follow-up on life courses, and the additional strength of knowing who lived together in a household at what point in time. For the current study, family mortality histories were especially relevant, therefore necessitating working with LINKS rather than the HSN.

Based on our work on cases found in both LINKS and the HSN, cases with observations on their life course that were as complete as possible were selected for analysis. Individuals with known dates of death in Zeeland and parents who had a known date of marriage were included. Furthermore, results showed that individuals who had no known date of death were very unlikely to have died in their childhood. Thus, the analyses were based on the assumption that children without a known date of death survived their infancy or childhood. The use of sensitivity analyses—changing that assumption to an assumption of early death for these children, or leaving these children outside the selected sample—did not alter the conclusions. Third, and finally, the results have been compared across historical populations, increasing trust in the validity of the results across populations and generations.

The central aim of this dissertation was to contribute to the understanding of how health and survival are shaped within families, and how families shaped changing patterns in health survival and mortality in the course of the 19th and early 20th centuries. The research contributed to the literature on cohort and early-life effects on health and survival, life courses of individuals, and childhood mortality clustering, integrating approaches from biodemography, sociology, and historical demography. I have shown that infant and child mortality patterns are not individual but family-linked, and are even linked across generations. Families are therefore crucial in considering changes in health and mortality in historical populations, and changes in health and survival and mortality risk of adults and children should be understood from the perspective of changing conditions in families of origin. The effects of the family on the life course cannot only be linked to family-shared characteristics that negatively affect the whole set of siblings, but also to specific consequences of growing up in a family in which young-age mortality is rife.

Part of the link between sibling mortality and adverse consequences throughout the lives of individual survivors is explained by direct exposure to death and (infectious) disease in their families of origin, and part is explained by disadvantageous characteristics of the families of origin. Over the 19th and early 20th centuries, conditions of families gradually improved, and children started growing up in conditions that were unlike those in earlier days: poverty, overcrowded housing, insufficient and inadequate nutrition, and other adverse living conditions slowly declined. Both exposure to infectious diseases and structural differences between families, including their living conditions, contribute to adverse outcomes for surviving individuals. On the one hand, there is a causal process related to exposure to mortality that has adverse consequences for the physical health of exposed individuals in adulthood and late life. Their children are not as likely to survive their childhoods, and their own risk of death is higher throughout their adult years. On the other hand, poverty, socioeconomic

status, childcare, and behavioral and other family characteristics were related across generations and explain at least part of the intergenerational correlations in high mortality. This indicates that to explain patterns of high mortality in families, not just a single generation should be taken into account, but that a more comprehensive set of kin and explanations should be taken into account. Taken together, these processes suggest that the increasing life expectancy and ages at death are related to the reduction of adverse living environments and exposure to sibling disease and death during the demographic transition in the late 19th and early 20th centuries.

In this thesis, I have taken a critical look at the role of the family in shaping children's chances of survival and death, emphasizing processes that work beyond individual characteristics and exposures. I have assessed the role of family histories of high mortality, involving both mother and father histories of high mortality, emphasizing how families shape the mortality chances of the children embedded in them. I have shown that the adverse effects on individual lives of originating from a high-mortality environment does not stop after the childhood years, but can also be found as scarring effects on health and survival into adulthood and old age. Furthermore, individuals who proceed to reproduction have increased chances of offspring mortality. Part of the high early-life mortality and early ages at death in Zeeland are therefore likely to be caused by a vicious circle of disease exposure and high mortality in families, which is related to unfavorable life-course trajectories after exposure to high mortality in childhood. The family therefore plays an important role in explaining mortality, not only because causes of high mortality are transmitted within multiple generations of families, but also because it constitutes an environment in which the early-life disease and mortality exposures of children are shaped. Thus, it is crucial to consider the family in research to how health and survival have been shaped in the past as well as today.

Part II

Paper I

Families in Comparison: An Individual-Level Comparison of Life-Course and Family Reconstructions Between Population and Vital Event Registers

(revised and submitted)

Authors
Van Dijk, I.K.*
Mourits, R.*
Van den Berg, N.*
Slagboom, P.E.
Janssens, A.
Mandemakers, K.

*Shared first authorship.

DOI 10.31235/osf.io/h2w8t

An earlier version of this paper was presented at the International Population Conference, IUSSP, Cape Town, South Africa in October 2017 and at the European Population Conference, Brussels, Belgium in June 2018. The paper has benefited from comments by Jan Kok, Sören Edvinsson, and Isabelle Devos and from the comments of three anonymous reviewers and an anonymous editor.

Abstract

In demographic research, large-scale individual-level data have become increasingly available. At the same time, it remains unknown how varying sources affect the reconstruction of individual life courses and families in databases. In this paper, we conduct individual-level comparisons of family and life-course reconstructions of 495 individuals simultaneously present in two well-known Dutch datasets: LINKS Zeeland and the HSN. The first dataset is based on a province's full population vital event registration data; the other is based on a national sample of birth certificates, after which individuals were followed in population registers. We compare indicators of fertility, marriage, mortality, and measurements of occupational status of individuals found in both databases and conclude that reconstructions in both the HSN and LINKS reflect each other well. LINKS provides more complete family information on siblings and parents, whereas the HSN provides more complete life-course information, especially for individuals who migrate out of Zeeland. Life-course and family reconstructions based on linked, fragmented observations on individuals constitute a reliable alternative to such reconstructions based on continuous observations from population or parish registers.

Keywords Family reconstitutions, life course, historical demography, demographic methods, family history, microdata, population registers, civil registers, databases, family reconstruction

1. Introduction

Demographic research is increasingly conducted on large-scale datasets. Underlying these databases are sources such as population registers, parish registers, registrations of vital events, censuses, and genealogical databases. These sources have been used to reconstruct families, focusing on kinship networks, and individual life courses; for example, births, marriages, and deaths. Names, ages, birthplaces, and other personal characteristics are used to link life-course events to individuals, and individuals to each other into family networks. Over recent years, the resulting databases have provided an impulse to various research fields and have been employed for wide-ranging innovative research questions on topics, including early-life conditions (Van Dijk, Smith, & Janssens, 2018), health, genetics (Van den Berg et al., 2019), the role of grandfathers in determining socioeconomic status (Knigge, 2016), widowhood and mortality (Alter, Dribe, & Van Poppel, 2007), household structure and child mortality (Kok, Vandezande, & Mandemakers, 2011; Riswick & Engelen, 2018), health of migrants and return migration (Puschmann, Donrovich, & Matthijs, 2017), the trade-off between fertility and longevity (Gagnon et al., 2009), and numerous other topics.

As the databases rely on different source material and construction approaches—including sampling and linking techniques—it is crucial to check the quality of life-course and family reconstructions contained in newly developed databases. Ideally, family and life-course reconstructions are cross-checked, not only against population-wide external sources, but also against other sources containing individual-level information on the same people. Due to data constraints, cross-checking is usually not possible. An exception is work by Wisselgren et al. (2014), who cross-checked data sources with the purpose of exploring the success rate of linking strategies. However, the information contained in resulting individual life-course and family reconstructions has remained unexplored.

In this paper, we show a comparison of life-course and family reconstructions for the same individuals in demographic datasets using two different, independent data sources: one based on the Dutch population registers and one based on Dutch vital event registers. Our purpose is twofold: to investigate to what extent life courses and family reconstructions are represented similarly in the two databases, which employ different source material, and to determine the suitability of the data for different research questions, including questions on life spans and mortality, marriage behavior, and fertility. The results are of interest to researchers working with demographic individual-level data—continuous observations from sources such as parish registers and church books, and more fragmented observations originating from sources such as vital event registers and censuses.

2. An overview of the literature

Earlier research focusing on the quality of individual-level, large-scale demographic databases has used a variety of approaches, which consider the characteristics of the source material and the logic of the construction of the database (Delger & Kok, 1998; Gavrilov & Gavrilova, 2001; Mandemakers, 2002; Kok, Mandemakers, & Bras, 2009). Some sources under-register females or specific age groups (Zhao, 1994), do not include historically and demographically relevant events such as extramarital fertility and stillbirths, or do not include unmarried individuals or individuals without children (Zhao, 2001). Other studies have used approaches based on internal consistency of databases and comparisons to external data sources such as mortality statistics in life tables. Gavrilov and Gavrilova (2001) judged the data quality of a genealogical database of royal and noble families on completeness, accuracy, consistency, and representativeness and observed a skewed sex ratio in older generations, but relatively complete infant-birth information. Kaplanis et al. (2018) compared demographic information from genealogical family trees to external sources and found that the data reflect historical events, international trends, and life-span distributions quite well, although infant mortality is underestimated. Schellekens & Van Poppel (2016) compared population-register data to national statistics and reported that in the Historical Sample of the Netherlands (hereinafter, HSN) cohort life expectancies at age 30 may contain some overestimation for men, but not for women. Adams et al. (2002) concluded that observations on migration in vital event registrations reflected migration information in population registers (the HSN) well. Ruggles (1992) observed that migration causes underestimation of population-level demographic indicators, such as age at marriage, age at first and last birth, and mean number of children. After migration, migrants are right-censored and demographic events are no longer observed, causing an underestimation of the number of events as well as the mean age at the corresponding events.

While approaches based on external data sources and false and missed links are useful instruments to judge the quality of databases, they only provide insight into deviations between databases and demographic realities at the aggregate level, such as differences in mortality rates. Whether individual life courses and families are reconstructed accurately remains an open question. Some efforts have been made in this direction, as linkage success and percentages of correct matches across sources—such as censuses—have been used as an indicator of data quality. Ruggles et al. (2018) emphasize that in most studies, missed links (type II errors) are given too much attention, at the expense of false links (type I errors). Both errors may introduce bias into the life-course and family reconstructions. However, missed and false links not only affect whether individuals are included in demographic databases,

but also whether their children, spouses, and parents are linked to them. As a consequence, life-course transitions are more accurately incorporated in more complete databases. Due to issues relating to frequently occurring names, name changes upon marriage, and variation in spelling across sources, both false matches and failed matches may occur across data sources. Several studies compared databases to explore the success rate of various automatic linking procedures—see, for instance, Wisselgren et al. (2014) for comparisons between Swedish censuses and parish registers, and Massey (2017) for historical US data.

In addition to linking strategies, characteristics of the underlying data sources affect life-course and family reconstructions. Civil registries record individual life-course events, whereas censuses and population registers monitor households. Reconstructions using the civil registry can more easily miss a vital event such as the birth of a child. Furthermore, out-migrants are usually lost from observation after their last observed vital event. At the same time, missing a single event—such as the birth of a child or a relative's marriage—does not necessarily end observations on individual life courses. In censuses and population registers, observations are centered on households and are relatively complete. Out-migration is commonly observed, providing researchers with a date of last observation. However, once individuals leave households, their lives are not necessarily followed further, limiting observations on family networks. Thus, both differences in source material as well as strategies for following individuals across data sources are crucial for the quality of reconstructed lives and families in historical databases.

3. Data

In the Netherlands, a unique opportunity has opened up to compare the individual life-course and family reconstructions in two different types of datasets. The first is based on a sample of birth certificates (HSN) and contains continuous observations originating from the nationwide population registers. The second is based on the civil registry of Zeeland (LINKS). Individuals born in Zeeland who were included in the HSN can be identified in LINKS through an identifying combination of the municipality, year, and sequence number provided on each civil certificate.

3.1 The civil registry and LINKS

3.1.1 The civil registry

The Dutch civil registry is one of the oldest in the world, and has covered the entire country from 1812 onward. Only in France (1792) and Belgium (1796) were civil registries introduced earlier. The civil registry uses a precise municipal bureaucratic

procedure guaranteeing a high data quality. Each year, municipalities were sent three sets of books in which to record births, marriages, and deaths. For each vital event, the name and age of the individual, the parents, and/or the spouse were recorded. Vital events were registered in duplicate and stored separately, in case of flooding or fire. All pages of the registers were watermarked by a local judge and checked for completeness at the end of the year (Vulsma, 1988). As a result, the recorded vital events are likely to represent the population well. This is especially the case for marriages, which took place and were recorded at the city hall. Deaths were reported by family members in the municipality of death. For deaths outside the municipality of residence, a copy of the death certificate was sent to the home municipality. Furthermore, births had to be reported to the municipality within three working days, in principle by the father or—if the father was not known or was absent—by another authorized person, such as the midwife or the doctor who attended the birth (Vulsma, 1988), or the person in whose house the birth occurred—for single mothers, often the maternal grandfather.

The Dutch civil registry of birth, marriages, and deaths is a good source for life-course and family reconstructions. All certificates contain the date of the event, the date of the registration (birth and death certificates), the place of registration, name, and age of the person reporting the event, and the names and places of residence of the witnesses. The birth certificates contain the name of the father if he is not the person present at the registration, as well as the name and age of the mother and the name and sex of the child. The marriage certificates contain the occupation and place of residence of both spouses, the names of the bride's and groom's parents, and—if they were alive—their age and occupation. For death certificates, one of two persons reporting the death is usually a spouse or parent. The civil registers of births, marriages, and deaths become public with delays of 100, 75, and 50 years, respectively.

3.1.2 The LINKing System for historical family reconstruction (LINKS)

LINKS (LINKing System for historical family reconstruction) is based on digitized certificates from the civil registries, as indexed by the WieWasWie project, to reconstruct families. The Zeeland 2017.01 release of the database contains around 700,000 birth certificates, 200,000 marriage certificates, and 650,000 death certificates. Multigenerational families were built using linked marriage certificates, reconstructing life courses and families (see Figure 1). Of the births detailed in LINKS, 81% were linked to the marriage of their parents. In total, the dataset contains almost 2 million persons covering a maximum of seven generations. Individual life-course reconstructions resulted from linking civil birth, marriage, and death certificates: 68% of all birth certificates and 66% of all marriage certificates were linked to a death certificate (Van Dijk, unpublished). The scope of the database is large regarding intergenerational networks of family members, but observations are relatively fragmented for life courses, as these concern either linked vital events or observations linked to these events. In addition, LINKS does not contain information on addresses, co-residence of individuals, migration movements, and religion.

Indexes of civil birth, marriage, and death certificates were linked together, using combinations of at least two pairs of names of individuals, spouses, and parents, combined with time constraints based on age. Variations in the spelling of names, name changes, and the nonuniqueness of many names renders family reconstruction a complicated task. To prevent missed matches due to spelling variations, all first and last names were corrected for minor known variations in spelling. All name combinations for at least two persons, the individual and one or two of his or her parents and possibly a spouse or child, were matched. In the data release used here, certificates were only linked within the province of Zeeland, so that certificates of individuals who out-migrated from Zeeland to another province in the Netherlands or abroad were missed. This concerned a sizable part of the population; for example, those who migrated to Belgium and Rotterdam (Priester, 1998).

3.2 Population registers and the HSN

3.2.1 Population registers

Population registers were introduced in the Netherlands from 1850. The population registers were maintained by each municipality in large books, organized by streets or neighborhoods. The registers contain information on households and their individual members. For each household, the registers contain information on the family name, given names, sex, marital status, birth dates, death dates, birth places, address, professions, and religious denominations. For married couples, the head of the household was the male spouse. After his death, his widow would remain the head of the household until her death or remarriage, or until she moved into a household with an existing head (Stadsarchief Amsterdam, 2018). Relationships between the members of the household were included from the perspective of the head of the household, allowing the reconstruction of relationships between other household members. Movements into and out of the household as well as births and deaths were continuously tracked. The books containing the population registers were replaced every ten years. Every ten years, the population registers were updated based on a nationwide census. This continuous registration allows the follow-up of households for longer periods of time.

In the period of research of this paper, two important changes in the population registries were implemented. In 1920—and earlier in the large cities—the population

registration was no longer ordered by street or neighborhood, but by individual household in a card system with separate documents. From 1939 onward, the registration was no longer focused on households, but on individuals by means of personal cards. Later, in 1994, this personal card system was completely digitized. Nowadays, the system is known as the Personal Records Database (Dutch: Basis-registratie Personen, BRP) and is maintained on the national level. One year after a person's death, a summary of personal and family information becomes available for scientific and genealogical research (CBG, 2019) and for specific research purposes a request can be made to Dutch government to directly access the BRP.

3.2.2 The Historical Sample of the Netherlands (HSN)

In the HSN, the life courses of a representative group of individuals in the Netherlands are followed. The HSN enables research on detailed life courses of individuals from the 19th century for the Netherlands (Knotter & Meijer, 1995; Mandemakers, 2000, 2002). The HSN is based on a sample of birth certificates of all individuals living in the Netherlands, stratified by cohorts of ten years for the period between 1811 and 1922 and according to regional levels of population density. The sample consists of 0.75% of the births for the period 1812–72 and 0.5% of the births for the period 1873–1922. In total, the sample consists of about 85,500 individuals (IISG, 2018). Up to the present time, about 40,000 of these 85,500 persons have been followed in the population registers throughout their life course. In the HSN, these persons are referred to as "research persons" (RPs). The population-register information in the HSN was supplemented with information from the Dutch birth and marriage certificates.

In the HSN release 2010.01, entries in population registers and on personal cards of 37,137 RPs were made available (Mandemakers, 2002). For some regions, including Zeeland, the HSN already starts in 1850, when the population registers were introduced. The database includes information about the RP's household, including co-residents, occupation, and religion. Households were, in principle, only followed as long as the RP was present in that household. Siblings and other kin were eventually lost from observation when the RP moved out of the household or died, after a follow-up to the end of the ten-year population-register period. For the period after the implementation of family cards for individual households, the remaining family members were followed for up to 40 or 50 years.

4. Structural differences between the HSN and LINKS

Because of the sampling procedure and independent sources of information, structural differences exist between the databases in the life-course and family reconstructions (see Figure 1 and Table 1). First, in the HSN, life courses of RPs were traced using population registers, allowing for the reconstruction of individual life courses and the multigenerational reconstruction of RPs' families. These reconstructions are based on population-register information of a sample of RPs, and will therefore, first and foremost, reflect events occurring in the lives of these individuals. This implies that family reconstructions were only available to the extent to which family members cohabited with RPs. Questions with topics such as intergenerational and horizontal kin relations—for instance, sibling similarities in mortality—can therefore not be answered. Second, in LINKS, family reconstructions were based on indexes of vital event registration data. LINKS is not centered on a sample or a research person, but on entire families. Individuals were observed when vital events occurred in their lives, in the lives of their spouse, or in the lives of their children. Consequently, the HSN is primarily focused on life-course reconstruction and less on family reconstructions, whereas the opposite applies to LINKS, in the sense that observations on life events are used to reconstitute families. As the two databases both contain elements of life-course and family reconstructions, but vary in the purpose for which they were built—and in their construction criteria and underlying sources—they are likely to differ in the completeness of their life-course and family reconstructions. We will explore to what extent events of fertility marriage, migration, mortality, and occupational careers were observed and differed between the HSN and LINKS, with the purpose of determining the strengths and weaknesses of both databases.

Automatic record linking, as applied to LINKS, has two potential issues: missing links and false links. Within the LINKS database, record linking was done in a conservative manner, limiting the number of false links. Between individual's birth and marriage, and between marriages of parents and children, we expect that the quality and quantity of links will be high. Birth certificates had to be shown upon marriage, limiting the potential impact of misspelled names. More problems may be expected with regard to linking birth and marriage certificates to death certificates, as the names on birth certificates were not controlled by means of other documents. The same problem occurs when birth certificates of children are linked with the marriage certificate of their parents. Furthermore, in order to match death certificates to corresponding birth certificates, accurate ages at death were necessary, which may not be present for reported deaths that are not cross-checked with birth certificates. Therefore, the automatic record-linking procedure will occasionally miss matches between vital event certificates.

Some demographic information was not included in LINKS due to the characteristics of vital event registration in the Netherlands. First, in contrast to the HSN, LINKS does not encompass unmarried cohabitation or extramarital children, which may lead to an underestimation of the number of children or siblings. Second, certificates were only linked within the geographic area of a province, so persons were lost if they migrated to another province. Thus, mortality in early life was most likely measured quite accurately, but certificates of deaths and marriages occurring later in life are more likely to be unavailable. In addition, due to privacy regulations, registers of deaths occurring after 1962 have not been released yet. Finally, key indicators such as occupation and place of residence were only observed in concordance with vital events of individuals, their spouses, or their children. Migration patterns and occupational careers can be reconstructed from an individual's civil certificates, as well as from their children's civil certificates. Death certificates contain occupational information if the deceased person had an occupation at the time of death. Hence, more complete occupational and residential information was available for RPs who married or had children. Moreover, most of these vital events occur relatively early in life, so that later changes in place of residence and occupation could easily be missed. For unmarried individuals, only vital events in the family of origin and one's death certificate will be observed.

In the HSN, there were no systematic observations of events occurring to the parental family before the sampled RP was born. In addition, due to missing population registers for some periods and places, observations on RPs does not always start at birth, leading to gaps in life-course and family reconstructions. The implication is that siblings who reside elsewhere or died young may not be included in the ten-year register in which the RP first appears. As a result, the count of all known siblings reflects the count of surviving siblings—the net fertility—rather than the count of all siblings ever born—that is, the total fertility. At the same time, RPs' children were identified very accurately in the HSN because RPs were followed for their entire life course. This is illustrated by Janssens (1993), who showed for Tilburg (1849–99) that 99.8% of the children found in the birth registers were identified in the population registers. At the same time, stillbirths and children who died very soon after birth were usually not included in the birth or population registers, but only in the death registers (hereafter, "lifeless reported infants"). These characteristics limit opportunities for research on events early in the life course—such as exposure to sibling mortality or the length of birth intervals—and research on intergenerational relations in longevity, mortality, and fertility. However, as yet, the extent to which these characteristics affect demographic estimates in the data is unknown.

Table 1 The expected availability of demographic indicators in the HSN and LINKS

	HSN		LINKS	
	Availability on data sources	**Reason**	**Availability on data sources**	**Reason**
Parents				
Marriages	Incomplete	Not included if parents were not in household; Marriage date of parents often not known	Incomplete	Not available for RPs who moved out of Zeeland
RPs				
Sibship size	Incomplete	Not included if siblings died before follow-up of the RP, Or were born after RP moved out of the household	Incomplete	Not available for RPs who moved out of Zeeland
Marriages of RPs	Incomplete	Marriages incompletely registered in population registers	Incomplete	Not available for RPs who moved out of Zeeland
Fertility	Incomplete	Offspring not included if they died before registration; No stillbirths recorded	Incomplete	Not available for RPs who moved out of Zeeland
Family relations	Not always clear	Relations within household need to be logically reconstructed for the period 1850–62; family relations to the third or fourth degree may be unclear in subsequent registers	Clear	–
Occupation	Complete	Updated regularly	Incomplete	Not available for RPs who moved out of Zeeland; Only known when a vital event was registered; Measured relatively early in the life course
Later-life mortality	Complete	-	Incomplete	Not available for RPs who moved out of Zeeland
Extramarital fertility	Complete	Premarital fertility included; RPs who lived together but were not married	No information	No information on extramarital fertility
Migration	Complete	Continuous follow-up of migration in the Netherlands	Incomplete	Only known when a vital event was registered; Persons are followed through zeeland only
Children				
Child mortality	Incomplete	No information on offspring outside the RP's household	Incomplete	Not available for RPs who moved out of Zeeland

5. Data construction and approach

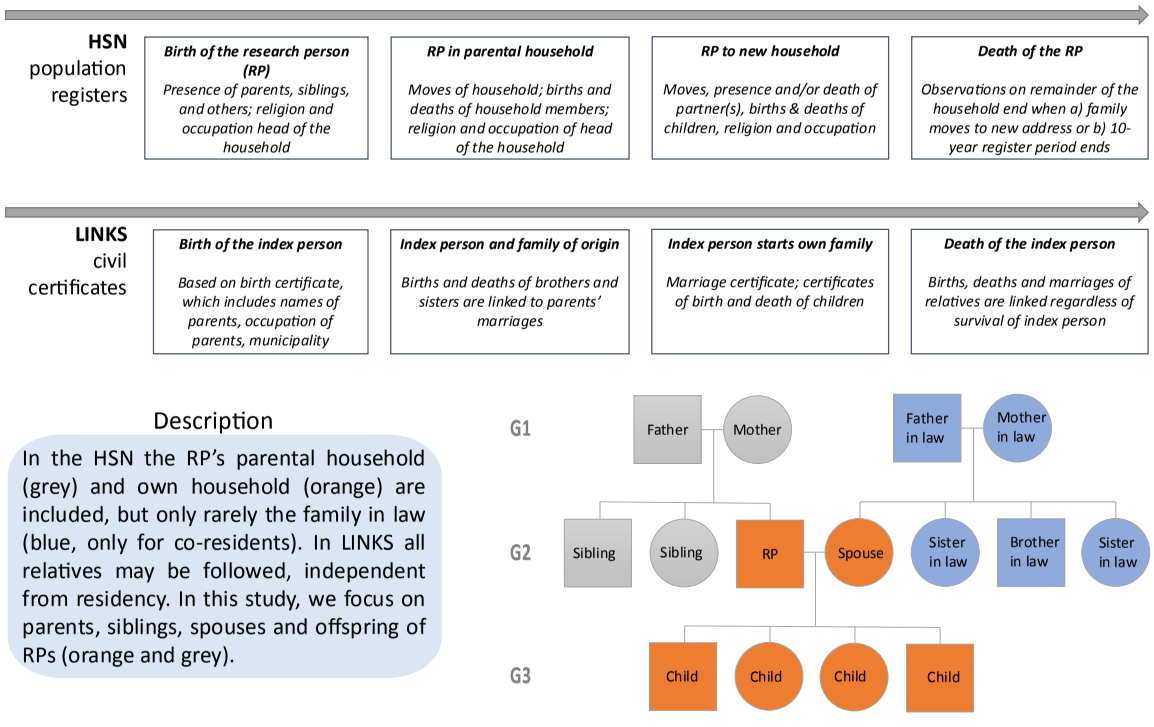
For the comparison between the HSN and LINKS, we used persons identified in both databases who were born between 1863 and 1872[1]. Drawing on data from LINKS 2017.01 (Mandemakers & Laan, 2017) and the HSN 2010.01 population-register release (HSN, 2018a,b), the 495 Zeeland-born individuals included in the HSN were traced in LINKS via unique identifiers of the birth certificates. We analyzed differences in life-course and family reconstructions of RPs in the estimation of key demographic and socioeconomic indicators. This strategy balances exploring linking success—which adds more spouses, children, and siblings to an individual—and accuracy of the reconstructed life courses and families—as missed links lead to lower estimations of numbers of spouses, children, and siblings, but higher estimations for age at first reproduction and marriage.

An overview of all available information in both datasets and expected completeness is provided in Table 1. For our analyses, we used the following indicators: sex, start and end dates of observation (HSN) or first and last observation (LINKS), birth year, and death year. We counted the number of siblings and children known, and the birth order of the RP in their family of origin. With regard to the number of siblings and children, stillbirths and infants lifeless upon civil registration were excluded, as they were unavailable in the HSN. In addition, we measured ages at first and last childbirth.

Furthermore, we noted whether RPs married or not and had children or not; calculated their age at first marriage and at death; traced whether they migrated within Zeeland, outside Zeeland (HSN), or never; and tested their socioeconomic position on consistency between both datasets using HISCLASS, a social class scheme to classify historical professions (Van Leeuwen & Maas, 2011; Mandemakers et al., 2018).

1 Earlier and later cohorts were not included here because population registers were of limited quality until 1860 and after 1915.

Figure 1 The data structure for HSN and LINKS



Both LINKS and HSN can be used for life-course (top two rows) and pedigree reconstruction. The top row shows the information based on the life-course reconstruction in the HSN. The second row shows the information based on the family reconstruction in LINKS. The pedigree shows, in general, the available persons in LINKS and HSN. The orange and gray colors represent the selection for the paper, which resembles the pedigree structure that can be derived from HSN.

Table 2 The available family ties in the HSN and LINKS for the selected 495 RPs from the 1862–71 Zeeland cohort

Relatives	Sample size	RPs with known relatives (%)
HSN		
Parents	932	475 (96)
RPs	495	
Siblings	1,447	336 (72)
Spouses	233	138 (28)*
Children	1,060	196 (40)
LINKS		
Parents	814	407 (82)
RPs	495	
Siblings	2,804	413 (83)
Spouses	188	177 (36)
Children	810	151 (31)

The 233 spouses are identified using the population registers. These numbers are used for the current work. When adding marriage certificates to the population registers, we identified 237 spouses and 277 unmarried RPs. Combined, the population registers and marriage certificates identify 324 spouses and 270 unmarried RPs. The RPs with known relatives refer to the number of RPs with, for example, known parents (N = 475). Spouses are based on the number of marriages. Hence, one RP could have had multiple spouses.

6. Results

Table 2 presents the number of RPs for whom parents, siblings, and children could be identified. In the HSN, 96% of the RPs had available parent information, for a total of 932 parents. In LINKS, 82% of the RPs had available parent information, for a total of 814 parents. In the HSN, 72% of the RPs were identified with siblings, totaling up to 1,447. In LINKS, a total of 2,804 siblings could be identified from 83% of the RPs. A total of 233 spouses were found in the HSN (28% of the RPs), while 188 spouses were identified in LINKS (36% of the RPs). Lastly, 1,060 children were identified in the HSN (40% of the RPs), whereas in LINKS 810 children were identified (31% of the RPs). The number of individuals without spouses and children appears to be high. However, many individuals in Zeeland did not reach reproductive ages, as infant and child mortality in Zeeland was very high, reaching up to 50% in some municipalities and years (Van Poppel et al., 2005; Van Dijk & Mandemakers, 2018).

Table 3 The number of available cases and the mean score on demographic indicators after selection of best cases in the HSN and LINKS separately (A) and simultaneously (B)

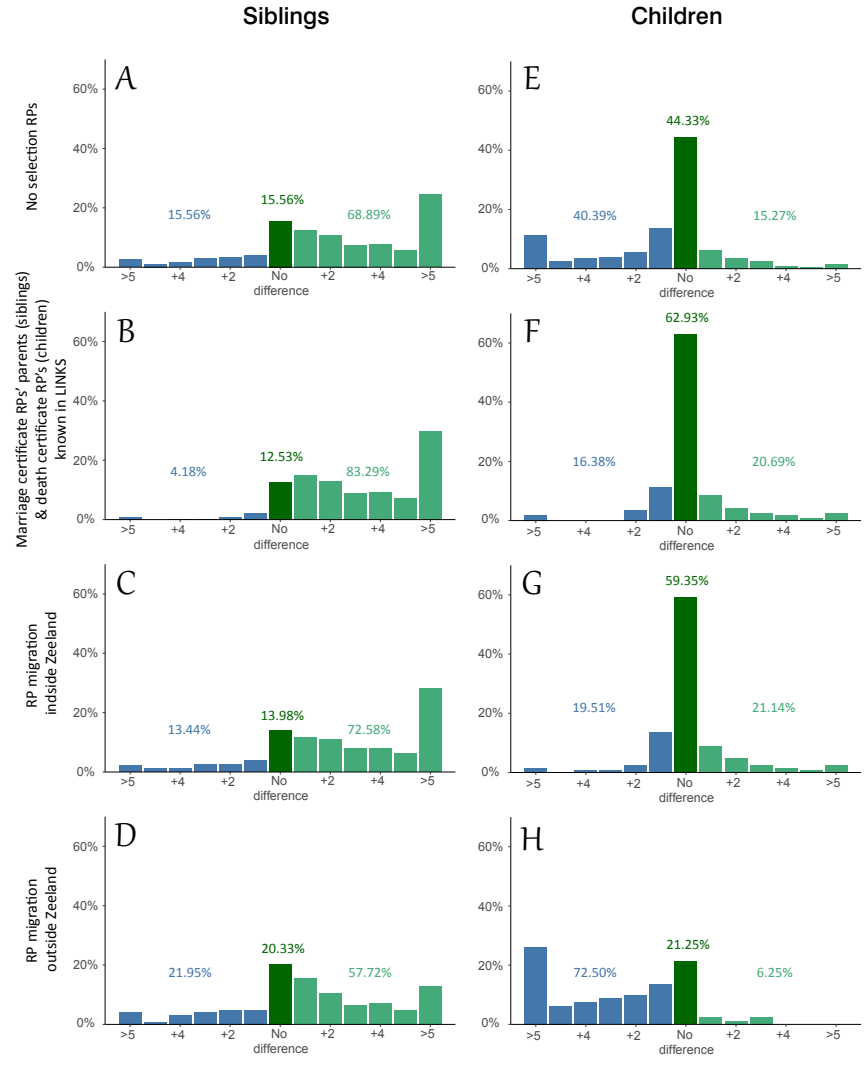
Data selection	Indicators	(A)				(B)		
		HSN	LINKS	HSN	LINKS	HSN + LINKS	HSN	LINKS
		N	*N*	*N* (mean)	*N* (mean)	*N*	*N* (mean)	*N* (mean)
Siblings								
Lived until age 30	Number of RPs with identified siblings (mean number of siblings)	221	–	221 (3.9)	–	186	186 (3.8)	–
Lived until age 30	Birth order	221	–	221 (1.8)	–	186	–	186 (6.6)
Parents known	Number of RPs with identified siblings (mean number of siblings)	–	407	–	407 (6.7)	186	186 (1.8)	–
Parents known	Birth order	–	407	–	407 (4.2)	186	–	186 (4.0)
Marriages								
Lived until age 30	Ever married*	221	146	122 (55.2)[a]	124 (84.9)	138	62 (44.9)[a]	118 (85.5)
Lived until age 30	Never married*	221	146	99 (44.8)[b]	22 (15.1)	138	776 (55.1)[b]	20 (14.5)
Ever married	Age at first marriage	137	178	137 (28.4)[c]	178 (26.3)	97	97 (27.7)[c]	97 (26.8)[c*]
Children								
Ever married	No children identified*	122	178	13 (9.5)[d]	26 (14.6)	97	8 (12.9)[d]	15 (15.5)[d*]
Children identified	Age at first childbirth	196	152	196 (27.0)	152 (26.5)	146	146 (26.8)	146 (26.6)
Children identified	Age at last childbirth	196	152	196 (37.4)	152 (36.6)	146	146 (37.4)	146 (36.8)
Children identified	Number of RPs with identified children (mean number of children)	196	152	196 (5.4)	152 (5.2)	146	146 (5.4)	146 (5.4)

Table 3 Continued

Data selection	Indicators	HSN	LINKS	HSN + LINKS	(A) HSN	(A) LINKS	(B) HSN	(B) LINKS
		N	*N*	*N*	*N* (mean)	*N* (mean)	*N* (mean)	*N* (mean)
Migration behavior								
Lived until age 18	Never moved*	236	157	149	62 (26.3)	90 (57.3)	56 (37.6)	86 (57.7)
Lived until age 18	Moved within Zeeland*	236	157	149	78 (33.1)	67 (42.7)	71 (47.7)	63 (42.3)
Lived until age 18	Moved out of Zeeland*	236	157	149	95 (40.3)	–	21 (14.1)	–
Age at death								
All	All ages at death	409	313	306	409 (40.8)	313 (34.7)	306 (33.6)	306 (33.9)
Lived until age 18	Ages at death after 18	236	157	149	236 (69.4)	157 (67.4)	149 (67.1)	149 (67.1)
Lived until age 50	Ages at death after 50	204	134	126	204 (75.1)	134 (73.6)	126 (73.6)	126 (73.6)

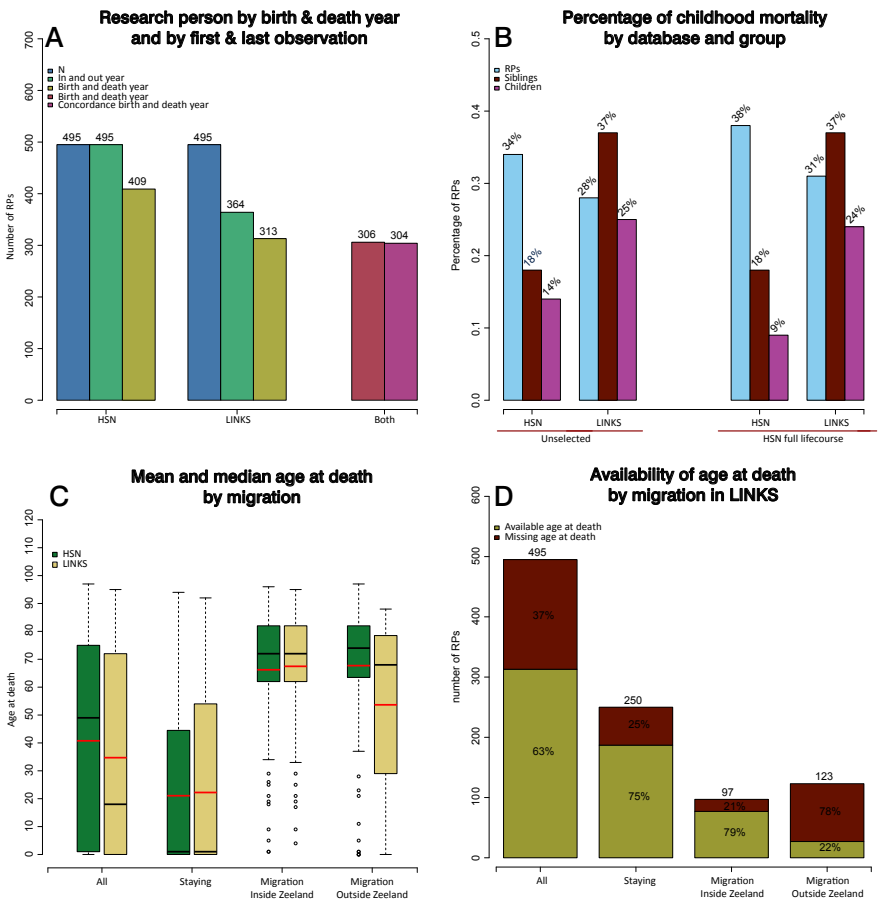
The indicators in Table 3A do not concern the same RPs, as the best cases are selected separately in the HSN and LINKS. Identical RPs are selected for both the HSN and LINKS in Table 3B. All indicators with an asterisk (*) are categorical. Hence, *N* (mean) represents the number and the proportion of RPs belonging to that subgroup. Marriages are based on the population registers in the HSN. The use of marriage certificates provides the following numbers in Table 3A: **a.** *N* = 194, mean = 87.8, **b.** *N* = 27, mean = 12.2, **c.** *N* = 137, mean = 28.4, **d.** *N* = 30, mean = 13.3. Reading example: row 7 shows that 122 persons lived beyond the age of 30 years and married at some point in their lives in the HSN and that this was the case for 124 in LINKS. These reflect, respectively, 55.2% of the 221 persons who ever married in the HSN and 84.9% of the 146 persons who ever married in LINKS. Use of the HSN marriage registers in Table 3B: **a.** *N* = 120, mean = 87.0, **b.** *N* = 18, mean = 13.0, **c.** *N* = 177, mean = 27.7, **c*** *N* = 177, mean = 26.3, **d.** *N* = 23, mean = 14.7. Reading example: row 7 shows that 138 persons lived beyond the age of 30 years in both the HSN and LINKS. Out of those 138 persons, 62 persons married at some point in their lives in the HSN and 118 in LINKS. These reflect, respectively, 44.9% and 85.5% of the in total 138 persons identified in both the HSN and LINKS. For siblings and children, *N* refers to the number of RPs with identified siblings. The mean refers to the mean number of siblings or children.

Figure 2 Matching for the number of siblings and children between the HSN and LINKS: more in LINKS (green) or HSN (blue)



Notes: The figure shows the matching for siblings and children of RPs between HSN and LINKS. Colors: brown, exact match; blue, more siblings or children identified in HSN than in LINKS; green, more siblings or children identified in LINKS than in HSN. The *x*-axis shows how many more siblings or children have been identified in either LINKS (green) or HSN (blue). In each case, the *y*-axis shows the percentage of matches corresponding to the *x*-axis. The marginal sums are illustrated with colors corresponding to the bars and sum up to 1 (100%). The legends on the left illustrate different data selections based on HSN or LINKS if explicitly stated. Numbers (*N*) per panel: A, 495; B, 407; C, 372; D, 123; E, 203; F, 116; G, 123; H, 80.

Figure 3 A comparison between HSN and LINKS for mortality data



Notes: Panel A shows the RPs by birth and death year and by first and last observations (HSN only) in absolute numbers. Panel B shows the percentage of childhood mortality (mortality <5 years) by database (HSN and LINKS) and group (unselected and full life course). Full life course indicates that HSN RPs are observed from birth. Panel C shows the mean (red color) and median (black color) age at death by migration (staying, migration inside Zeeland, and migration outside Zeeland) by database (HSN and LINKS). Panel D shows the availability of age at death by migration only for the LINKS database in absolute numbers and percentages. Migration for RPs is determined based on the HSN, since migration in LINKS is not available by definition of the source material.

Figure 4 The HISCLASS score for research persons in HSN and LINKS

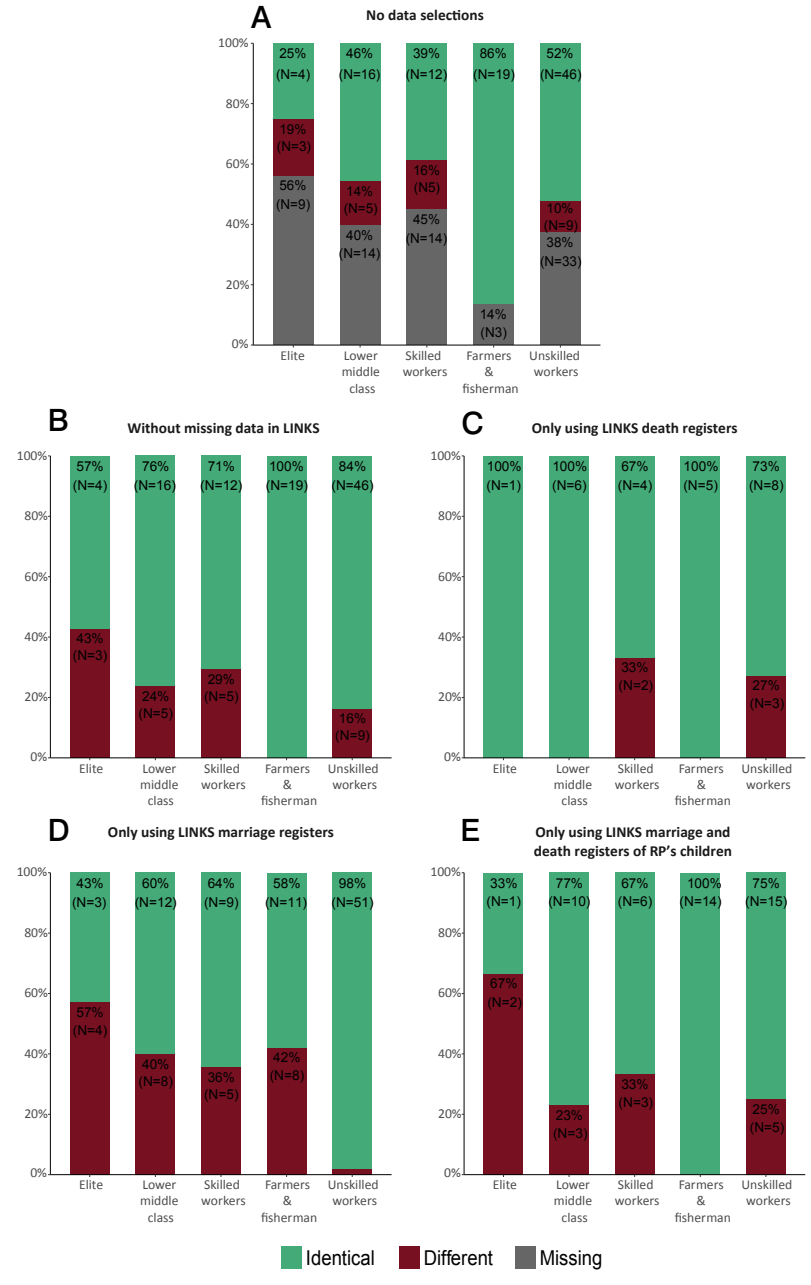


Figure 4 Continued

Notes: All panels in this figure are based on RPs with an available SES in the HSN. The highest SES is used. Overlaps and differences with LINKS are indexed. HSN was chosen as the focal point because the underlying source material for SES in the HSN is more complete. Because HSN is the focal point, LINKS data can be identical to the HSN, different from the HSN, or missing in LINKS. Seventy-three RPs are available in HSN and not in LINKS (40 males, 33 females); 32 RPs are available in LINKS and not in HSN (seven males, 25 females); and 119 RPs are available in both the HSN and LINKS (91 males, 28 females). Panel A shows the proportion of highest available SES between LINKS and HSN without any data selection in LINKS. Panel B shows the proportion of highest available SES between LINKS and HSN without the 73 missing RPs in LINKS. Panel C shows the proportion of highest available SES between LINKS and HSN with only information about death certificates used in LINKS. Panel D shows the proportion of highest available SES between LINKS and HSN with only information about marriage certificates used in LINKS. Panel D shows the proportion of highest available SES between LINKS and HSN with only information about marriage and death certificates of RPs' children used in LINKS.
Socioeconomic classes: 1, Elite, subdivided into (1a) Higher managers and (1b) Higher professionals; 2, Lower middle class, subdivided into (2a) Lower managers, (2b) Lower professionals, clerks, and sales workers, (2c) Lower clerks and sales workers, and (2d) Foremen; 3, Skilled workers, subdivided into (3a) Medium-skilled and (3b) Lower-skilled workers; 4, Farmers and fishermen; and 5, Unskilled workers, subdivided into (5a) Lower-skilled farm workers, (5b) Unskilled workers, (5c) Unskilled farm workers, and (5d) Unspecified workers.

Comparisons between the databases were conducted in two ways. First, we compared all individuals for whom relevant observations can be expected in both databases separately, with the purpose of exploring all life-course and family reconstructions (Table 3A). Because the mean scores in this table were based on different RPs, these means must be interpreted for each dataset separately. Second, we analyzed only the subsets for which we could reconstruct life courses in an identical way. Hence, we selected individuals for whom a relevant observation may be expected in both databases (Table 3B). Both tables show key demographic information for all RPs for whom it is possible to know whether they experienced the demographic event. Cases without information on the relevant selection criteria were not included. Differences between the HSN and LINKS in demographic indicators in Table 3B indicate differences in the reconstructions of life courses and families between the HSN and LINKS, whereas differences in these indicators in Table 3A may also be caused by differences between the subsets of individuals for whom information was available.

6.1 Comparisons of demographic indicators in the HSN and LINKS

Table 3A shows that the mean number of siblings and birth order were lower in the HSN (3.9 and 1.8) than in LINKS (6.7 and 4.2). These results were similar for the 186 identical cases. These differences are mainly a consequence of the research design of the HSN, in which siblings are only observed if they are part of the RP's household. Therefore, information on siblings who died young or who did not live in the household is often missing, leading to an underestimation of sibship size in the parental household of the RP.

Within LINKS, marriages were available for 84.9% of the RPs of 30 years and older, and the mean age at first marriage was 26.3. In the population-register release of the HSN, marriages were available for 55.2% of the selected RPs, with a mean age at marriage of 28.4. Table 3B shows that for the 138 common RPs, marital information was available for 85.5% in LINKS and for 44.9% in the HSN. The mean age at first marriage in LINKS was, at 26.8 years, higher than in the separate subset, whereas in the HSN it was lower, at 27.7 years. After combining the HSN population registers with the HSN marriage certificates, we observed that marriages were available for 87.0% of the RPs with a mean age at first marriage of 27.7 years (see the notes to Table 3). The higher mean age at marriage in the HSN is partly caused by right-censored observations in LINKS. Out-migration is known to cause underestimation of the number of events as well as the age at which demographic events occur (Ruggles, 2002). Nevertheless, the number of known marriages is higher in LINKS than the HSN after we selected only individuals who married in Zeeland. This is probably related to left truncation in the HSN, as not all RPs were followed for their entire life course, so that second marriages were counted as first marriages and hence the age at first marriage was overestimated.

Table 3A shows that the mean number of identified children in the families of the RPs was similar between both datasets: 5.4 children for RPs in the HSN and 5.2 in LINKS. However, the number of RPs with identified children was lower in LINKS ($N = 152$) than in the HSN ($N = 196$). Furthermore, the mean ages at first and last birth in LINKS (26.5 and 36.6 years) were lower than in the HSN (27.0 and 37.4 years). The percentage of married couples without identified children is 14.6% in LINKS and 9.5% in the HSN. Table 3B shows that for the 146 RPs who are included in both datasets, the mean age at first childbirth was 26.6 in LINKS and 26.8 in the HSN. This selection of common cases also showed the same mean number of children (5.4), although the mean age at last birth is lower in LINKS than in the HSN and the percentage of married couples without identified children is 15.5% in LINKS and 12.9% in the HSN. The smaller differences after we select cases that are available in both datasets indicates that the higher number of childless families and lower mean ages at first and last birth in LINKS are probably caused by right-censored observations due to out-migration. In

addition, the automatic linking procedure might have failed up to pick up certificates of later-born children and entire families altogether, as differences in mean age at last birth and mean number of children remain after selecting identical RPs.

The HSN and LINKS include different information on migration behavior, as out-migration from Zeeland was not observed in LINKS. Table 3A shows that, according to the HSN, 95 (40.3%) of the RPs who were alive at age 18 migrated out of the province at some point in their lives. In LINKS, 157 death certificates are available for the RPs who lived at least until age 18, suggesting that these RPs either never left Zeeland or returned to Zeeland at a later age. The HSN indicated that 140 RPs (59.6%) never lived outside Zeeland, of which 62 (26.3%) never moved at all, and 78 (33.1%) only moved within Zeeland. Vital events outside the place of birth of the RP, indicating migration between municipalities within Zeeland, were identified for 67 RPs. This pattern was similar when we compared identical individuals (Table 3B). In LINKS, we observed that 63 RPs (42.3%) who were observed after age 18 died in another municipality than their municipality of birth, whereas both vital events occurred in the municipality of birth for 86 RPs (57.7%). According to the HSN, 56 RPs (37.6%) remained in their municipality of birth, 71 (47.7%) moved within Zeeland, and 21 (14.1%) lived outside Zeeland at some point in their lives. As about one out of seven adults who were born and died in Zeeland lived outside Zeeland at some time, assumptions about cross-provincial migration behavior or the lack thereof should not be based on the presence of vital events in LINKS.

The last few rows in Table 3A present the number and mean age of death for all RPs for whom an age at death was known and the mean ages at death for individuals reaching 18 and 50 years. Because persons out-migrating from Zeeland are known in the HSN but not in LINKS, we expected that in the HSN more ages at death would be known and that the mean age at death would be higher. Indeed, fewer ages at death were known for RPs in LINKS than in the HSN, resulting in a lower mean age at death in LINKS (34.7 years) than in the HSN (40.8 years). The difference between the databases in the mean ages at death was smaller at higher ages. For those surviving until age 18, the mean age at death was 67.4 in LINKS and 69.4 in the HSN; after age 50, the mean ages at death are 73.6 and 75.1. An important reason for the declining difference over time is the lower likelihood that middle-aged individuals will out-migrate (Kok, 1997). The mitigated differences between the HSN and LINKS after identical cases were selected support our assumption that selective availability of information for out-migrated individuals plays an important role.

6.2 Comparing life-course and family reconstructions of RPs between the HSN and LINKS

Here, we take a closer look at deviations in individual life-course and family reconstructions. Figure 2 shows whether estimations of outcomes in the HSN and LINKS deviate upward, downward, or are identical. Because information may be more complete for some subsets of individuals, four groups are included: individuals with (a) no selections, (b) known marriage certificates for parents of siblings and known death certificates for RPs, (c) known migration inside Zeeland, and (d) known migration outside Zeeland.

Without any selections on the data (panel 2A), the number of siblings was higher for 69% of the RPs in LINKS, whereas 16% of the RPs in the HSN contained more siblings. The differences between the number of siblings in the HSN and LINKS were even more pronounced if a marriage certificate of the parents was known in LINKS (panel 2B). In LINKS, fewer siblings were found in 4% of all cases compared to the HSN, the same number of siblings was found in 13% of all cases, and more siblings were found in 83% of all cases. Migration within Zeeland did not affect the results (panel 2C), whereas for RPs who migrated out of Zeeland, the number of siblings in LINKS was lower than in the HSN in 22% of all cases, identical in 20% of the cases, and higher in the remaining 58% (panel 2D).

The number of children of the RPs was more similar between both datasets than the number of children from the parents (siblings of the RP). Figures 2E–H show the difference in number of children between the HSN and LINKS, which was calculated for RPs who had children identified in either or both datasets. Without selections on the data, the HSN provided the most accurate results. For 40% of all RPs, more children were found in the HSN than in LINKS; for 44% of all RPs, the same number of children was found in both datasets; and in the remaining 15%, more children were found in LINKS than in the HSN. The difference between family reconstructions in the HSN and LINKS may have been caused by interprovincial migration, as births outside Zeeland are not included in LINKS. To indicate the quality of the linking process, RPs unaffected by migration were selected by requiring the availability of a Zeeland death certificate, thus excluding migration effects. For these RPs, the same number of children was found in the HSN and LINKS in 63% and 59% of all cases, respectively (panels 2F and 2G). Where the number differed between the HSN and LINKS, there was no clear distinction in performance between the databases: the HSN performed better in some cases, whereas LINKS performed better in the others. If RPs moved out of Zeeland, a larger number of children was found in the HSN in 73% of all RPs, the same result in both sets in 21% of all cases, and a smaller number in the other 6%. Hence, the differences in family reconstructions between the HSN and LINKS are

caused by migration rather than linking quality. In general, LINKS gives a better estimate of the number of siblings, whereas the HSN and LINKS give a comparable number of children unless the RP moved out of Zeeland.

Figure 3 shows comparisons in the available mortality information between both datasets in four panels. Panel 3A shows that ages at death were known for 409 RPs (83%) in the HSN, whereas 313 RPs had an available age at death in LINKS (63%). The age at death overlapped in 304 cases (99%), for whom a death certificate was available in both databases. Panel 3B presents childhood mortality for three different groups: the RPs, their siblings, and their children. In the HSN, childhood mortality for RPs was estimated to be 7% higher than in the HSN, reflecting the good coverage of RP information in the HSN. Childhood mortality was estimated to be twice as high for siblings and almost three times as high for children in LINKS compared to the HSN. Furthermore, for RPs, the childhood mortality estimates were more accurate when selecting RPs who were followed for their entire life course. Adult mortality estimates were influenced by migration outside Zeeland. Panel 3C shows that for individuals who stayed in their municipality of birth or who moved within Zeeland, the mean and median ages at death were similar between both datasets. However, death certificates were not linked for 25% of the individuals who were marked as stayers in the HSN and 21% of the individuals who were identified as internal migrants. Some of these individuals might have survived the observation period, as death certificates are not available after 1962, whereas for other cases death certificates were not linked due to spelling and age variations on the death certificates. For RPs who left Zeeland according to the HSN, the mean and median age at death is much lower in LINKS than in the HSN. The date of death is known for only 22% of all RPs for whom we know, based on the HSN, that they lived outside Zeeland at some point in their life course (panel 3D). These return migrants have a significantly lower age at death than interprovincial migrants who died outside Zeeland. Thus, LINKS contains a selective group of return migrants and might have failed to link available death certificates due to spelling and age variations.

Figure 4 shows the differences in socioeconomic position between the HSN and LINKS based on HISCLASS (Van Leeuwen & Maas, 2011). We present social class on the abbreviated HISCLASS scale with five categories (Kok & Mandemakers, 2009). Figure 4 also shows whether RPs in the HSN with an available HISCLASS 5 score had none, the same, or a different socioeconomic status in the LINKS dataset. Figure 1 in the appendix shows the same information, but is based on the full 13-category HISCLASS. Panel 4A shows that, in total, 73 RPs had known socioeconomic information in the HSN but not in LINKS. The share of missing values varied between 38% and 45% for unskilled workers, skilled workers, and the lower middle class, was slightly higher for the elite with 56%, and only occurred for 14% of the farmers, indicating that farmers had a better chance of being included in the LINKS dataset than individuals with other socioeconomic positions. Figure 4B shows that HISCLASS scores were identical for 80% of the RPs for whom occupational information was known in both datasets. All farmers in the HSN were also farmers in LINKS. However, differences in social position were found for 22% of the others. Most discrepancies occurred for the elite (43%)—more than for the lower middle class (24%) and skilled workers (29%). Fewer differences were found for the unskilled workers (16%). Underestimation of socioeconomic status generally occurred when information on occupational status was not known after marriage (Delger & Kok, 1998). Thus, lower scores in LINKS might be a consequence of geographic mobility, which was higher for individuals with a better socioeconomic position (Ekamper, Van Poppel,& Mandemakers, 2011) and which, for the LINKS dataset, also resulted in having no occupational information on higher ages.

Figures 4C–E show comparisons of the occupational score in the HSN with three parts of the LINKS dataset: the RP's death certificate, his or her marriage certificate, and the marriage and death certificates of the RP's children. The choice for a certain certificate determined the sample size as well as the moment of observation. Occupations were only recorded on death certificates if the deceased held an occupation at the time of death. As a result, occupational information on death certificates is limited and only available for 29 cases, but the HISCLASS scores were very similar between the datasets. No different socioeconomic positions were found for the elite, the lower middle class, and the farmers, whereas 29% of the unskilled workers and 33% of the skilled laborers held a different socioeconomic position on their death certificates than in the population registers. Marriage certificates were available for 112 RPs, of whom 52 were identified as unskilled workers in the HSN. In 98% of the cases, these were also identified as unskilled laborers on their marriage certificate. However, marriage certificates are less concordant with the HSN for socially mobile individuals. Between 36% and 42% of the farmers, skilled laborers, and lower middle class had a different occupational position on their marriage certificate than in the HSN. This difference was larger for the elite (57%). The 59 RPs with marriage and death certificates of children in the LINKS dataset (panel 4E) have a better balance between sample size and matching quality in socioeconomic position than the comparison made in panel 4D. Similarly, farmers show no differences at all between the HSN and LINKS datasets. For the other groups, socioeconomic positions differ from 23% to 33% of unskilled workers, skilled workers, and the lower middle class. For the elite, the estimates seem worse, since two out of the three observations are different. Because in the LINKS dataset the number of observations were dependent on the number of children, the more children an RP had, the better HSN and LINKS reflect each other.

7. Discussion and conclusion

We expected that we would find a lower mean age at death in LINKS than in the HSN, as it has been shown that migrants are often healthier than the native population, a phenomenon known as the healthy migrant effect. Indeed, we found a lower mean life span for the RPs in LINKS (34.7 years) than in the HSN (40.8 years), which was attributable to the almost 100 extra observations of life spans that were available in the HSN in comparison to LINKS. These observations mainly concerned out-migrated adults, increasing the length of the mean life span in the HSN. Moreover, we found that individuals who were observed outside the province of Zeeland during their life course, but who returned to Zeeland, died at earlier ages than individuals who never migrated or who migrated within the province of Zeeland. This suggests that return migration occurred for health considerations, contributing to the problem of underestimation of ages at death in LINKS.

In the literature, a number of earlier studies have reported findings in line with this "salmon bias" which states that the relative health advantage of migrants in comparison to the native population may at least partially be caused by return-migration movements of unhealthy migrants. Earlier work has found that healthier individuals tend to migrate more and further in contemporary as well as historical populations (Alter & Oris, 2005; Lassetter & Callister, 2009), and, for internal migrants in modern-day China (Lu & Qin, 2014) and the US (Turra & Elo, 2008) that unhealthier migrants tend to return migrate to their countries of origin. At the same time, a historical study on Rotterdam did not find evidence for either a healthy migrant effect or a salmon bias (Puschmann et al., 2017). Evidence for the current data is in line with both mechanisms. Possibly, in Rotterdam, healthy migrant effects were counterbalanced by a heavy urban penalty affecting migrant's health, which is absent in the small towns of Zeeland. We found that if a death date was included in both the HSN and LINKS, they were identical in 304 out of 306 instances, indicating the validity and comparability of the life-course reconstructions in both databases.

Ruggles (1992) showed for English family reconstitutions that—even in the absence of healthy migrant effects—cessation of observation on individuals due to out-migration causes underestimation of the ages at which demographic events occur. As more individuals were lost from observation due to migration in LINKS than in the HSN, due to the provincial scope of LINKS and the national scope of the HSN, we expected that the mean age at which life-course transitions occur would be lower in LINKS than in the HSN. Indeed, we found that not only age at death, but also age at first marriage, first childbirth, and last childbirth were higher in the HSN than in LINKS. In addition, in LINKS, more men and women had no identified children than in the HSN.

With regard to children, family reconstructions appear to be more accurate in the HSN, as the number of identified children in the HSN is slightly higher than in LINKS. But after adjustment for migration, the number of children identified in the databases was usually the same. This finding illustrates that, indeed, the life courses of migrants are covered to a more limited extent in LINKS than in the HSN. In line with our expectations, that does not apply to the size of the sibling set of the RPs. Due to early mortality and migration, not all siblings were found in the population registers in which the RPs appeared. The linking algorithm used to construct the LINKS dataset results in very precisely reconstructed families, so that not only the number of children, but also the number of siblings is relatively accurately measured. Thus, although the life-course reconstructions appeared to be very accurate in the HSN, the parental family reconstructions were less complete than in LINKS. Analyses using total numbers of siblings and analyses of intergenerational relations in demographic behavior, including total fertility and age at marriage, should be based on LINKS rather than the HSN.

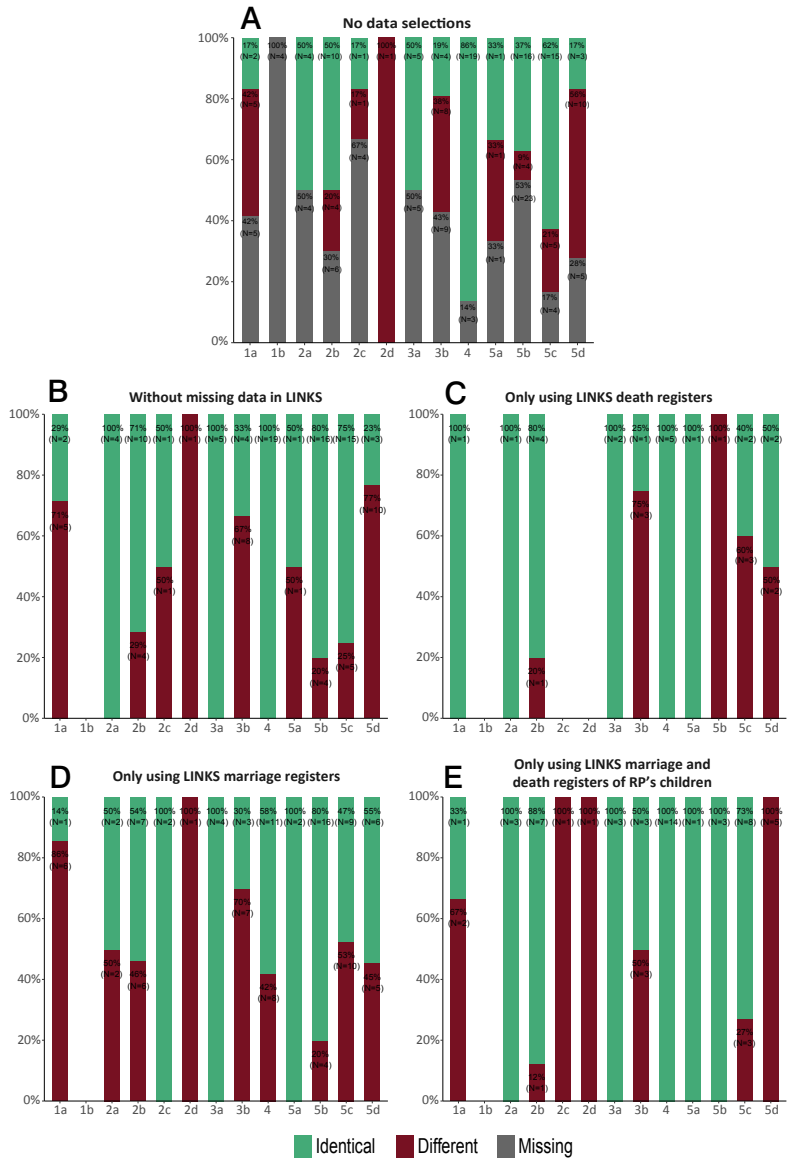
Although life courses are reconstructed relatively completely in both the HSN and LINKS, both datasets have their own strengths and weaknesses. It should be kept in mind that the purpose of the HSN was to reconstruct life courses of RPs in great detail, but not necessarily their families of origin. In line with that purpose, information on children of RPs is relatively complete, although the life courses of children were only followed until they left the RP's household, or until cessation of observations on the RP. Hence, the population-register data are relatively well equipped to reconstruct individual life courses, but less well suited to reconstruct families. This applies to censuses as well, especially if they are not continuously updated, unlike the Dutch population registers. For LINKS, it appears that record linkage of vital event observations can deliver high-quality family reconstructions. The strength of vital event register data mainly lies in family reconstructions, because individual life-course reconstructions can only be made for selective groups in the dataset. Information on household composition, unmarried cohabitation, or extramarital fertility was not available. Therefore, a full understanding of the nature of the source material is required to study inter-generational effects with vital event data. As such, the HSN and LINKS are well suited to answer a wide range of questions, but only if one utilizes the strengths of either dataset.

Information from different datasets can be combined to gain new and more complete insights into demographic behavior. The extended-family reconstructions found in LINKS can contribute to more detailed family reconstructions in the HSN; for instance, with regard to lifeless reported infant siblings and children or more detailed observations on socioeconomic status. In current versions of the HSN, marriage certificates—

which are also included in LINKS—are already used to enrich information on relationship formation found in population registers. Second, differences between the two databases may itself be of interest for family historians and historical demographers. Deviating information on siblings and children within households in the HSN and regardless of household in LINKS may provide researchers with clues on non-co-resident kin, a phenomenon on which neither database alone provides information. Similarly, supplemental observations on socioeconomic status in the HSN may enrich our understanding of the development of the status of individuals over time. As the current analyses have shown, it should be taken into account that information for certain individuals may more readily be matched between databases, such as index persons from LINKS who remained in their province of origin.

This paper has illustrated how matched databases may be used to cross-check family and life-course reconstructions and to gain insight into sources of bias. In the future, the LINKS project will be continued for other provinces of the Netherlands, implying that it may be possible to reconstruct families not only within provinces, but also between provinces, following individuals who migrate elsewhere. This will challenge current methodologies for linking individuals, as the number of individuals with similar names increases, leading to a larger number of potential false matches. Individuals who are found both in the HSN and in LINKS may be used to cross-check linking strategies for such data. Similarly, LINKS currently does not contain premarital births, an issue that may be solved by improving the linking strategy, as civil certificates do exist. The HSN may be used to improve methods for the automatic linking of premarital births to mothers and fathers, thus contributing to the further development of the LINKS database. We conclude that life-course and family reconstructions based on linked, fragmented observations on individuals constitute a reliable alternative to such reconstructions based on continuous observations from population or parish registers. Hence, through the further integration of existing sources, innovative research questions and approaches may be generated.

Figure A.1 The HISCLASS score for research persons in HSN and LINKS



Notes: All panels in this figure are based on RPs with an available SES in HSN. Overlap with LINKS and differences with LINKS are indexed by distinguishing between RPs who have the same HISCLASS score in both the HSN and LINKS, the same score on the five-category HISCLASS but a different score on the 13-category HISCLASS, and a different score on both the five- and 13-category HISCLASS scales. HSN was chosen as the focal point because, by definition of the source material, HSN is more complete.

Seventy-three RPs are available in HSN and not in LINKS (40 males, 33 females); 32 RPs are available in LINKS and not in HSN (seven males, 25 females); and 119 RPs are available in both the HSN and LINKS (91 males, 28 females). Panel A shows the proportion of highest available SES between LINKS and HSN without any data selection in LINKS. Panel B shows the proportion of highest available SES between LINKS and HSN without the 73 missing RPs in LINKS. Panel C shows the proportion of highest available SES between LINKS and HSN with only information about death certificates used in LINKS. Panel D shows the proportion of highest available SES between LINKS and HSN with only information about marriage certificates used in LINKS. Panel D shows the proportion of highest available SES between LINKS and HSN with only information about marriage and death certificates of RP's children used in LINKS.

Paper II

Early-Life Mortality Clustering in Families: A Literature Review

Population Studies (2019), *72(1)*, 79–99

Author
Van Dijk, I.K.

DOI: 10.1080/00324728.2018.1448434

This paper has benefited from comments by Isabelle Devos, Sören Edvinsson, Ken Smith, and Alice Reid, and from two anonymous reviewers and an anonymous editor.

Abstract

Research on early-life mortality in contemporary and historical populations has shown that infant and child mortality tend to cluster in a limited number of high-mortality families, a phenomenon known as “mortality clustering.” This paper is the first to review the literature on the role of the family in early-life mortality. Contemporary results, methodological and theoretical shortfalls, recent developments, and opportunities for future research are all discussed in this review. Four methodological approaches are distinguished: those based on sibling deaths, mother heterogeneity, thresholds, and excess deaths in populations. It has become clear from research to date that the death of an older child harms the survival chances of younger children in that family, and that fertility behavior, earlier stillbirths, remarriages, and socioeconomic status all explain mortality clustering to some extent.

Keywords Biodemography, child mortality, demographic methods, less developed countries, early-life mortality, family demography, historical demography, infant mortality, literature review, mortality clustering

1. Introduction

This paper offers a review of the causes of a phenomenon known as mortality clustering, defined in this study as the occurrence of excess infant and child mortality in a select group of families. Almost 25 years ago, Monica Das Gupta (1990, 1997) recognized that, despite high mortality rates in the Indian region that she was studying, infant and child mortality were concentrated in a subset of the families under study. Over recent decades, it has been shown that early-life mortality does indeed cluster in families, in both high- and low-mortality populations (Van Bodegom, Eriksson, Houwing-Duistermaat, & Westendorp, 2012; Van Poppel, Bijwaard, Ekamper, & Mandemakers, 2012), in historical populations (see, e.g., Edvinsson et al., 2005), and in present-day populations in the developing world (see, e.g., Zenger 1993; Omariba et al., 2008). Mortality clustering, or death clustering, has important implications for research and policy. For research methods, it implies that in determining the causes of early-life mortality, we should account for the correlated observations of children from the same family. For theories on early-life mortality, death clustering implies that we need to shift the focus from the child level to the level of the families in which children are embedded. Finally, for policy, it means that an infant or child death in a family may function as a red flag for further bereavement.

Over recent decades, many researchers have contributed to our understanding of the causes of mortality clustering in families. At the same time, these efforts have been fragmented—partially taking place in the field of (historical) demography and partially in development studies and anthropology, using different methodological approaches and therefore yielding insights into mortality clustering from varying analytical angles. Whereas some researchers have focused on the scarring effects of child deaths on the health of remaining siblings, others have analyzed heterogeneity between families or mothers, analyzed differences between high- and low-risk families, or assessed concentration in families in the population. In this review paper, I attempt to place the approaches and results of these research efforts within a coherent framework. A systematic literature search was conducted using relevant search engines, including PubMed, Web of Science, and Google Scholar. A search for literature on infant and child mortality or death clustering provided 681 titles to consider. I identified all papers discussing mortality clustering and removed those that did not consider infant or child mortality, overviews of the literature, contributions without data analyses (including research notes), presentations unaccompanied by papers, and earlier versions of selected works. This left 170 papers that were listed for further consideration. Full papers were carefully assessed to determine whether these contributions addressed differences in infant and child mortality between families or mothers: 53 papers met these criteria, while 117 did not. A further six

papers and one book chapter were added from other sources. In total, 60 papers and chapters are included.

In this paper, I reflect on innovative research efforts and results, identify challenges, and explore avenues for future research that have opened up in part because new, large-scale datasets have become available, for both contemporary developing-world populations and historical populations from the developed world. While there has been extensive research on spatial mortality clustering as well as mortality clustering in families, the focus of the current paper is the latter, and so research concerning spatial clustering is included only if the primary topic is early-life mortality clustering in families. To my knowledge, this is the first attempt to review the literature on infant and child death clustering. The review is based on literature on mortality clustering in the first 60 months of life (i.e., both infant and under-five child mortality) in families; I refer to this collectively as "early-life mortality clustering."

Research on death clustering does not commonly deliver insights into the medical causes of death of individuals: instead, it aims to explore why some families are more vulnerable to disease and death among their children than others (for an exceptional case where medical causes are taken into account, see Lundevaller & Edvinsson, 2012, on Rh disease). In historical and developing-world populations, communicable diseases (such as diarrhea and measles) constitute a major cause of death (Van Bodegom et al., 2012). Families have social, economic, and behavioral characteristics, which are risk factors for exposure and vulnerability to diseases (Mosley & Chen, 1984; Sastry, 1997). Death clustering seems to be caused by a complex interplay of all kinds of causes: not only do socioeconomic characteristics seem important, but genes, the size of the family, maternal care practices, and the health of the mother also seem to be related to the risk of early-life mortality. But the risk factors are not exclusively the domain of the mother and other family members: socioeconomic, genetic, behavioral, and environmental roots of individual susceptibility to death can be found at the child level and the community or population level too (Madise & Diamond, 1995; Sastry, 1997). These three analytical levels (child, family, and community) have been the focus in different strands of research within the mortality-clustering literature, and all three are taken into account in this review.

The paper proceeds as follows. In the next section, the concept of mortality clustering is explored in depth. Four approaches to measuring and modeling death clustering are evaluated on their merits. Then I discuss current evidence, focusing in turn on the level of the child, the family, and the community. Finally, pressing challenges for future research are identified. Although extensive research has been conducted in some subfields and results have been reaffirmed by multiple scholars, other subfields have received more fragmentary attention, and yet others remain completely unexplored. First, I turn to the concept and measurement of early-life mortality clustering.

2. Early-life mortality clustering: meaning, measurement, and modeling

Mortality clustering is a crucial concept for understanding early death, as research has repeatedly shown that families play a pivotal role in the survival chances of infants and small children. In most studies on mortality clustering, the authors present the distribution of deaths across families in the population under study, using various measures. Often, the share of deaths in high-risk families is shown as a proportion of the total number of deaths in the population; in other cases, the proportion of mothers who experience a certain share of the deaths is reported. For instance, 12.6% of mothers accounted for 62.2% of child deaths in rural Punjab, India in the 1980s (Das Gupta 1990), while Edvinsson et al. (2005) found that in 19th century Skellefteå, northeast Sweden, approximately 50% of deaths were found in 10% of the families. Typically, such numbers are used to show the extent to which clustering plays a role in mortality in the populations being studied.

Although sometimes presented as an indicator for mortality clustering, these numbers are not necessarily meaningful. In measuring whether populations experience inequality in the number of early deaths between families, it is essential to take both the size of the family and the binomial distribution of the probability of mortality into account. First, large families can be expected to lose more children than small families. If 62% of the babies were born into 12% of the families in rural Punjab at the end of the 20th century, it would not be remarkable if these large families experienced 62% of the total infant mortality. In that case, their mortality rate would be the same as that of the population as a whole. Second, in addition to family size, we need to account for binomial distributions to assess whether mortality clusters in families in a population. If early deaths follow a binomial probability distribution, the implication is that not all families will experience the same realized mortality rate: there must be 'lucky' large families who escape early death among their children and 'unlucky' small families who have to bury multiple children. Even if one in five children dies before their fifth birthday, chance predicts that more than 10% of large families with ten children will not lose a single child. Thus, in assessing whether deaths cluster in the population, we should take both chance variations and the size of the family into account, and this constitutes a major methodological challenge.

A further challenge is related to the assumption of many statistical techniques commonly applied in infant and child mortality research that observations—in this case, observations on the mortality of siblings—are independent from each other. Considering that mortality clusters in high-risk families, observations on children originating from the same family are not independent (Beise & Voland, 2002; Omariba et al., 2007). Thus, we have to account for these linked survival chances as well as chance variations and family size in our statistical models.

Many methodological solutions have been proposed for these issues, resulting in varying definitions of mortality clustering at three analytical levels: the children, their families, and their communities. Four different approaches are commonly used in the literature on mortality clustering in families. These approaches are summarized in Table 1. A first approach, used at the child level, is to use sibling death as an indicator of mortality clustering in sequence models. The second and third approaches are applied at the family level. In these cases, some authors use random intercept models assessing heterogeneity in the likelihood of death of a child at the level of the family as an indication of mortality clustering. Other authors distinguish high-risk from low-risk families by a set threshold; for instance, at least two deaths in the sibling group (Das Gupta, 1990). The fourth and final approach uses the difference between the expected and observed numbers of families with a certain number of deaths at the community or population level as an indication of mortality clustering.

These variations in methods and definitions of mortality clustering have resulted in a rather heterogeneous field of research, split up by the analytical level of interest, in which results are not easy to compare and are seldom considered in an integrated fashion. In this paper, I attempt to reunite findings from these varying angles, exploring methods and findings, and identifying gaps in the literature. First, the four approaches are discussed in more detail.

The first approach models the impact of the death of one sibling on their other siblings, interpreting the relationship as a causal effect, using sequence models or dynamic models (Table 1, approach 1). Examples include the work by Curtis et al. (1993), Zenger (1993), Arulampalam and Bhalotra (2006, 2008), and Omariba et al. (2008). Analytically, the focus lies on the effect of the death of a sibling on survival of the other siblings after accounting for the correlated death risk of the sibling set. Neither parity nor the binomial chance distribution are usually taken into account, meaning that the focus is not on excess mortality, but on early-life death itself, as a result of a previous sibling's death. Furthermore, authors often use Markov models to assess scarring effects. These models assume a strict sequence of events: the death of an older sibling influences the death of a younger sibling. This assumption is not usually met in models that incorporate mortality after the first birthday and, consequently, there has been more attention on scarring effects on infants (<12 months) than on older children (12–60 months); for an exception, see Ikamari (2000). Finally, neither family nor community effects are incorporated in the models.

Table 1 An overview of the methodological approaches to mortality clustering

Approach	Definition of mortality clustering	Example methodology	Takes binomial chance distribution into account?	Insight into family variation?	Insight into child variation?	Insight into community variation?	Example studies
1	Sibling death	Sequence analysis (dynamic models)	No	No	Yes	No	Zenger 1993 Arulampalam & Bhalotra 2008 Omariba et al. 2008
2	Mother heterogeneity	Variance of random intercept, usually in survival analysis (frailty models)	No	Yes	To some extent	Depending on approach	Sastry 1997 Janssens & Pelzer 2012 Scalone et al. 2017
3	Families experiencing mortality beyond set threshold	Logistic regression	Depending on approach	Yes	No	No	Das Gupta 1990, 1997 Edvinsson et al. 2005 Vandezande & Matthijs 2013
4	Difference between observed and expected number of deaths	Simulation models; binomial models	Yes	Yes	Depending on approach	No	Ronsmans 1995 Holmberg & Brostrom 2012 Lundevaller & Edvinsson 2012

In the second approach, random intercept models are applied to focus on mother heterogeneity in the likelihood of death and survival of children, modeling family-level variables in conjunction with child-level variables (Table 1, approach 2). Often, these models include a description of the (explained) variance at the family level and sometimes the community level. The results commonly show the extent to which variation exists between families before and after controlling for family- and child-level characteristics, and the variance of the mother- or family-level effects is interpreted as an indication of mortality clustering (for a discussion of the interpretation of model parameters, see Scalone et al., 2017). The models do not compare the observed distributions of deaths with the expected distributions, although solutions involving simulation techniques have been suggested (Holmberg & Broström, 2012). The effects of unobserved family characteristics are assumed to be independent from other predictors in the model. Furthermore, the models assume that all children are affected equally by their family's characteristics and, therefore, offer limited insights into differences within families. Examples include work by Guo (1993), Sastry (1997), Bolstad and Manda (2001), and Janssens and Pelzer (2012).

A third approach also focuses on the family level, applying a threshold approach to distinguish between high- and low-risk families and using logistic regression analysis to model explanatory characteristics at the family level (Table 1, approach 3). While earlier authors used criteria such as sibling death (Das Gupta, 1990) or multiple child loss (Das Gupta, 1997; Kuate-Defo & Diallo, 2002), more recent authors have used the size of the family and population mortality rates to determine which families experience mortality beyond what can be expected. For instance, Edvinsson et al. (2005) considered families to be high risk if their infant death rate was twice the population-level infant mortality rate or higher. While this approach to mortality clustering is intuitive and takes fertility and mortality patterns in the population into account, it should also be noted that there is an operationalization effect. In large families, it is more likely that at least one infant or child death occurs and, therefore, the likelihood of a large family being a low-mortality family is lower. This effect has been illustrated graphically by Vandezande (2012), who showed that small families were more likely to be classed as low-mortality families than large families. Furthermore, although this approach puts family characteristics in the spotlight, there is less opportunity to assess whether child or community characteristics play a role in explaining mortality clustering.

In the fourth approach, the expected number of families experiencing a certain number of deaths is compared with the observed number, to determine the extent to which excess deaths occur in the population (Table 1, approach 4). Models of excess deaths usually assume that the probability of child death is equal for all families within a group of a given parity. Furthermore, it is not easy to integrate the effect of individual characteristics of children or the mutual influence of siblings on each other's survival chances. At the same time, these models take the binomial chance distribution and parity into account most explicitly and allow for comparisons between populations exhibiting different fertility patterns.

The choice of a particular analytical model influences both the level at which we look at mortality clustering and the definition of mortality clustering—and, because of model assumptions, it is also closely related to the potential explanatory characteristics and mechanisms that can be addressed. While sequence models offer insights mainly into factors at the child level (such as birth order and interactions within the sibling set, including survival of the preceding sibling), random intercept models and threshold models allow us to pinpoint factors that are mainly relevant for differences between families, as do models that are based on excess deaths. In the following sections, I discuss the four approaches as distinguished above in terms of the results on mortality clustering that have been produced. This discussion is organized according to the three different analytical levels: the child, the family, and the community. In the third section, I discuss research that predominantly gives insight into child-level predictors of early-life mortality. In the fourth section, I move on to discuss studies providing insights into family-level predictors, including threshold and heterogeneity approaches. In the fifth section, I discuss studies that yield insights into the sources of community-level variation. An overview of all publications on mortality clustering, including a summary of their approach and main results, can be found in Table 2. In a final section, I discuss new questions and directions for research, including some innovative methodological approaches to mortality clustering.

Table 2 An overview of publications on early-life mortality clustering

Author(s) and year	Population[a] and period[b]	Families	Children	Age group[c]	Approach[d]	Main results
Alam, 1995	DSS, Teknaf, Bangladesh (1983–84)	NA	3,729	0–35 months	1	Birth spacing and survival are more strongly associated if the sibling survives
Alam & David, 1998	DNFS, Matlab, Bangladesh (1977–90)	NA	32,650	0–35 months	3	Adjacent siblings' mortality risk is affected by death of older sibling; neonatal mortality is higher, toddler mortality lower
Alter et al., 2001	Sart, Belgium (1812–1900)	NA	918; 839; 469; 113 respectively	1–14; 15–29; 30–54; and 55+ years	1, 2	Correlation between siblings in mortality weakens after age 15 and disappears after age 30; some evidence for acquired immunities
Arulampalam & Bhalotra, 2006	DHS, Uttar Pradesh, West Bengal and Kerala (–1998)	2,340–7,297	5,950–29,937	Infants	1	Mortality clustering in families is reduced by availability of contraceptives
Arulampalam & Bhalotra, 2008	DHS, 15 Indian states (–1999)	2,340–9,370	NA	Infants	1, 2	Scarring effects and evidence for sex preferences found
Beise & Voland, 2002	Krummhörn, Germany (1720–1874)	NA	3,530	Children	2	Maternal but not paternal grandmothers improve likelihood of survival; timing of death indicates that causes are exogenous
Bhalotra & Van Soest, 2008	DHS, India (–1999)	7,286	29,747	Neonates	1	Scarring effects and some proof for replacement behavior found
Bolstad & Manda, 2001	DHS, Malawi (–1992)	2,911	4,838	Infants	2, 1	No scarring effect found if family risk is taken into account; risk factors—early succeeding conception and short breastfeeding duration
Curtis et al., 1993	DHS, Brazil (–1986)	2,308	4,752	Postneonates	2, 1	Birth intervals are strongly related to survival
Curtis & Steele, 1996	DHS, Bolivia, Kenya, Peru, and Tanzania (–1989)	4,754–5,739	11,586–13,134	Neonates	2, 1	Biological mechanisms are likely because of similar strength of family associations in neonatal mortality in four different populations
Das Gupta, 1990	Punjab, India (1980–90)	About 1,800	1,520	Children	1, 3	Risk factors: short birth intervals, short breastfeeding duration, and birth weight; socioeconomic status and childcare abilities matter
Das Gupta, 1997	Punjab, India (–1984)	674	NA	Children	3, 4	Short birth intervals appear to be the effect of mortality; evidence for clustering found especially in low-status and low-education groups
Edvinsson et al., 2005	DDB, Sundsvall and Skellefteå, Sweden (1803–1900)	20,005	133,448	Infants	3	Risk factors: social status, number of marriages and earlier stillbirths
Guo, 1993	Six communities in Guatemala (–1976)	851	3,120	Children	2, 1	Relatively small effect of households on mortality found beyond household economic status and mothers' education
Gyimah, 2009	DHS, Ghana (–1998 & –2003)	3,540	4,938	Children	2	Higher risk found for older children in polygynous households
Holmberg & Broström, 2012	DDB, Skellefteå, Sweden (1831–90)	8,062	37,074	Infants	4, 2	Smaller effect of scarring found once death clustering is taken into account
Hussain et al., 2001	DHS, India (–1993) and Pakistan (–1991)	5,447 and 3,993 respectively	17,531; 14,050	Children	2	Consanguineous marriages play a significant role in infant mortality
Ikamari, 2000	DHS, Kenya (–1989)	NA	16,426	Children	1	Sibling death risks are correlated, especially in infancy
Janssens et al., 2010	Twente, the Netherlands (1875–99)	163	733	0–24 years	1	Higher mortality found for girls in late childhood than for boys
Janssens & Pelzer, 2012	HSN, two cities in the Netherlands (1900–30)	353	1,556	Children	2	Occupation of the mother is unrelated; older mothers beneficial
Kippen, 2011	Tasmania (1852–57)	NA	NA	All ages	3	Deaths within families often occur closely spaced in epidemic years

Table 2 Continued

Author(s) and year	Population[a] and period[b]	Families	Children	Age group[c]	Approach[d]	Main results
Kippen & Walters, 2012	Sart, Belgium (1812–1900)	NA	2,123	Children	1	Evidence found for sibling competition over resources and for death clustering
Kuate-Defo & Diallo, 2002	Africa-wide DHS data (–1978/1998)	NA	NA	Children	3	Preceding child death, contraceptive use, and birth spacing affect risk
Lalou, 1997	PRDH, Quebec, Canada (1621–1730)	NA	NA	Infants	4	Wet nursing and difficult deliveries due to maternal health given as explanations
Last, 1992	Farmstead in Nigeria (–1992)	30	131	All ages	Qualitative	Replacement of departed child may lead to subsequent elevated mortality risk due to timing of next birth and behavioral factors
Lindkvist & Broström, 2006	DDB, Skellefteå, Sweden (1831–90)	59,757	133,448	Infants	2	Intergenerational transmission of mortality clustering found, especially for girls
Lundevaller & Edvinsson, 2012	DDB, Skellefteå, Sweden (1860–1900)	4,943	23,067	Perinates	4	Rh disease explains more than a third of perinatal mortality
Lynch & Greenhouse, 1994	DDB, Sundsvall, Sweden (19th century)	5,754	20,626	Children	1	Scarring effect found, especially for adjacent siblings. Increased risk of death already present for oldest children in family before death of siblings
Madise & Diamond, 1995	Malawi (–1988)	3,043	6,258	0–15 years	1	Effect found of socioeconomic factors and short birth intervals, controlled for mortality clustering in the household
Manda, 1998	DHS, Malawi (–1992)	2,650	3,927	Infants	2	Biological and socioeconomic factors contribute to death clustering in families; no significant community variation in infant deaths
McMurray, 1997	DHS, Burundi, Uganda, and Zimbabwe (1987–89)	2,529–3,433	11,078–16,075	Children	3	Includes several regions and finds that clustering is lower in urban and malaria-free regions, and higher for mothers living in poor conditions
Miller et al., 1992	DNFS, Matlab, Bangladesh (1975–80) and Cebu, the Philippines (1983–86)	NA	1,755 and 3,029, respectively	0–24 months	3	Children born after short birth interval have higher risk of death in the first two years of life, even after accounting for gestation—this is not explained by shorter birth intervals in families with a high mortality risk for children
Myntti, 1993	Village in Yemen (1988–89)	16	NA	All ages	Qualitative	Socioeconomic factors, lack of social support, and attitudes are relevant
Nault et al., 1990	PRDH, Quebec, Canada (1608–1729)	?	17,010	Infants	4, 3	Birth intervals are related to mortality due to breastfeeding and maternal depletion
Nonyane et al., 2013	Sylhet, Bangladesh (2001–05)	13,457	24,485	Neonates	4, 2	Random household variation is almost completely due to between-mother effects and is largely explained by child and mother characteristics
Omariba et al., 2007	DHS, Kenya (–1998)	5,716	23,348	Children	2	Biodemographic factors are more important in infancy than in childhood
Omariba et al., 2008	DHS, Kenya (–1998)	5,717	23,351	Infants	4, 1, 2	Unobserved heterogeneity and scarring play a role in infant mortality
Pakot, 2015	Transylvania, Romania (1850–1939)	1,883	8,841	Infants	1	Longer birth intervals are related to a lower mortality risk, especially for neonates
Pavard et al., 2005	PRDH, Quebec, Canada (1625–1759)	NA	58,365	Postneonates—15 years	2	Mother's death negatively affects survival, especially for older girls
Pelzer & Janssens, 2014	HSN, four cities in the Netherlands (1880–1920)	656	2,682	Infants	1	Mortality is more strongly determined by regional differences than religion
Quaranta et al., 2017	Five historical European populations	381–44,429	1,445–207,071	Infants	3, 2	Intergenerational transmission of mortality clustering occurs in five European populations

Table 2 Continued

Author(s) and year	Population[a] and period[b]	Families	Children	Age group[c]	Approach[d]	Main results
Ranjan et al., 2018	NFHS, four Indian states (–2006)	11,992	38,392	Infant	1, 2	Find effects of both scarring and mother heterogeneity
Rao et al., 1997	NFHS, Goa, India (–1993)	1,331	NA	Children	3	Maternal education may be important because of hygiene behaviors
Reid, 2001	Derbyshire, England (1917–22)	NA	29,537	Stillbirths and neonates	1	Stillbirths are concentrated in certain women; greater influence of external factors for neonatal mortality than for stillbirths
Reid, 2002	Derbyshire, England (1917–22)	NA	24,743	Postneonates	1	Previous child deaths affect all-cause mortality, and mortality from wasting, diarrhea, and respiratory diseases
Ronsmans, 1995	Niakhar, Senegal (1983–89)	1,664	12,752	Children	4	Presence of known within-family heterogeneity does not explain the variance between families, suggesting an interfamily component
Saha & Van Soest, 2011	HDSS, Matlab, Bangladesh (1982–2005)	25,088	64,344	Infants	1	Scarring works through birth intervals, which become shorter after death of preceding infant
Saha et al., 2014	HDSS, Matlab, Bangladesh (1987–2005)	NA	107,367	Neonates	2	Education of the mother is protective and low socioeconomic status of the father detrimental for child mortality, especially for communicable diseases
Sastry, 1997	DHS, northeast Brazil (–1986)	1,051	2,946	Children	2, 1	Community variation in mortality is more relevant than family variation
Scalone et al., 2017	Granarolo, Italy (1900–1939)	1,214	3,968	Perinates and neonates	2	Variation between sharecroppers and 'landless laborers and nonrural workers' in mortality clustering; occupation irrelevant for perinatal mortality
Sear et al., 2002	West Kiang, Gambia (1950–74)	NA	2,294	Children	2	Living maternal grandmother reduces child mortality among toddlers
Shah & Dwivedi, 2011	Sewa, Gujarat, India	33	NA	Neonates	3 and qualitative	Family history of high mortality is mainly related to prematurity and low birth weight causes of death
Van Bodegom, 2012	Northeast Ghana (2002–10)	1,703	9,288	Children	2	More variance occurs at household than village level; influence of both fathers and mothers found; socioeconomic status and water sources play an influential role
Van Poppel et al., 2012	LINKS, three provinces of the Netherlands (1812–1903)	90,000	485,303	Stillbirths and children	1, 2	More death clustering found in later periods; change seen over time in effect of stillbirth on death of index child
Vandezande, 2012	COR*, Antwerp, Belgium (1846–1905)	322 and 376, respectively	1,222 and 1,826, respectively	Stillbirths and infants	2, 3	Death risks are correlated between generations in maternal and paternal line; age at death is relevant—late neonatal births are most predictive
Vandezande, 2013	COR*, Antwerp, Belgium (1846–1905)	406	1,634	Postneonates		Postneonatal deaths are transmitted across generations in the paternal line
Vandezande & Matthijs, 2013	COR*, Antwerp, Belgium (1846–1905)	376	1,826	Infants	3	Evidence found for intergenerational mechanisms; these are not due to reproductive behavior, social class effects, or persistent regional effects
Willführ & Gagnon, 2012	Krummhorn (1720–1874) and PRDH, Quebec, Canada (1670–1720)	6,445 and 13,932, respectively	29,935 and 115,013, respectively	Infants	4	Stronger concentration of infant deaths in Quebec, possibly explained by better maternal quality in populations with low marriage rate
Zaba & David, 1996	Census of Kenya (–1979)	194,198	691,546	Infants	4	Parity and age explains part of the clustering of deaths; more variation at higher ages and parities
Zenger, 1993	Matlab, Bangladesh (1968–73)	2,832	7,304	Neonates	1	Neonatal deaths are most strongly associated for adjacent siblings; birth spacing effect on mortality is stronger if sibling survived

Notes: [a]COR* sample from Antwerp, based on the first three letters of the family name; DDB, Demographic Database; DHS, Demographic and Health Survey; DNFS, Determinants of Natural Fertility Survey; DSS, Demographic Surveillance System; HDSS, Health and Demographic Surveillance System; HSN, Historical Sample of the Netherlands; NFHS, National Family Health Survey; LINKS, LINKing System for historical family reconstitution; PRDH, Programme de recherché en démographie historique.[b]Where only the end of the observation period is provided (e.g., –1986), observations are cross-sectional, based on retrospective survey or census data. [c]Infants refers to the age group 0–1 and children to the age group 0–5; other age ranges are specified. [d]For an overview of the approaches, methods, and their characteristics, see Table 1. The approaches are listed in order of predominance.
NA, Not applicable.

3. The role of scarring: the influence of the death of an older sibling

While mortality clustering refers to high mortality in certain families, many current studies have focused, paradoxically, on the survival of individual children. The focus has been on the effect of the death of a sibling, usually an older sibling, on the survival of an index child (see Table 2, studies marked as approach 1). The relationship between survival of siblings has been interpreted as a scarring effect, which can be attributed to several mechanisms. First, exposure to infectious disease and death may result in damage to the bodies or immune systems of surviving siblings (Bengtsson & Lindström, 2000; Alter et al., 2001; Barker et al., 2002; Finch & Crimmins, 2004). Second, bereaved parents may be depressed, which can have detrimental effects on the remaining children because of suboptimal caretaking. Third, child deaths are related to short birth intervals through two mechanisms: after the death of an infant, breastfeeding is interrupted, which leads to faster restored fecundity (Nault et al., 1990); and, also, parents may decide to replace the lost child quickly (Arulampalam & Bhalotra, 2006, 2008; Bhalotra & van Soest, 2008). In turn, short birth intervals lead to higher maternal and child mortality, as they may result in maternal depletion, which increases the risk of pregnancy complications, preterm delivery, and low birth weight (Curtis et al., 1993; Omariba et al., 2008).

Using this approach, it has been shown that neonatal deaths (those between birth and the 28th day after birth) and postneonatal deaths (those later in infancy) are the most strongly related between adjacent pairs of siblings (Zenger, 1993; Curtis & Steele, 1996). However, in families that will eventually lose many children, this relationship is present even early on: during the years in which the family is still small, the oldest children already have increased likelihoods of death (Lynch & Greenhouse, 1994). One of the explanations for scarring effects between pairs of siblings is short birth intervals (Saha & van Soest, 2011). As just discussed, closely spaced births are related to maternal and early-life child mortality. However, research has shown that short birth intervals are detrimental if the preceding sibling survives, but that these effects are weaker if the preceding sibling has died (Zenger, 1993). In addition, death clustering is more pronounced among higher parity women (Zaba & David, 1996). These results indicate that the effect of birth intervals may be partially explained by resource competition (Kippen & Walters, 2012) or transmission of infectious disease between closely spaced siblings or in overcrowded households (Das Gupta, 1990; Alam, 1995; Kippen, 2011). Finally, some authors have found that the effect of scarring appears to be dependent on the sex of the preceding sibling, which could indicate that son preferences play a role in the correlation between sibling deaths (Arulampalam & Bhalotra, 2008).

Later research in this vein not only took the effect of sibling survival into account, but used sequence models that were able to control for family-level heterogeneity, thus accounting for the fact that observations within families are not independent. These studies have shown that scarring and unobserved heterogeneity both play a role in explaining early-life mortality, emphasizing that there are differences between families in mortality levels and also that the death of a sibling has a causal effect on survival of the next sibling (Arulampalam & Bhalotra, 2006, 2008; Omariba et al., 2008). However, once clustering in the family is taken into account, scarring seems to play a much more modest role (Bolstad & Manda, 2001; Saha & van Soest, 2011; Holmberg & Broström, 2012). Saha and van Soest (2011) have suggested that, when keeping heterogeneity constant (i.e., controlling for death clustering), the death of a sibling scars remaining siblings, increasing the risk of death by 29% in a contemporary population in Bangladesh, mainly through a decrease in the interval to the next birth. This study is a welcome contribution among the methodologically more advanced contributions to scarring research, as it has identified a pathway by which scarring influences the family. In itself, scarring is a black box, as it does not clarify the mechanisms by which a first death causes further deaths and does not reveal why the death of a child is so detrimental to their brothers and sisters. Furthermore, it tells us little more than that a first death is related to higher odds of further deaths and should thus function as a red flag for local governments and health-care workers, among others. Still, the causes of the initial death remain unknown and we learn little about the families in which deaths occur.

An additional problem of research into the role of scarring is that the econometric methods that are commonly used to determine the impact of scarring in the family have strict model assumptions. First, models that take both unobserved heterogeneity between families and a scarring effect into account use a Markov approach. Markov models assume that only adjacent children are influenced by the death of a sibling (Omariba et al., 2008), but that may not be entirely true. The mechanisms that cause a relationship between the mortality of adjacent brothers and sisters may influence nonadjacent children as well as adjacent children. For instance, all children are likely to suffer from the consequences of parental depression or maternal depletion, not just younger or adjacent children. In addition, it has been shown that effects of sibling death on nonadjacent siblings do exist, although the effects are largest for adjacent children (Alam, 1995). Omariba et al. (2008) included random effects in their models to capture the influences of nonadjacent children; however, this solution assumes that the nonadjacent effects influence all other children in a similar fashion. Second, for the assumption of sequence to hold, older siblings must die before younger siblings. Scarring research must focus on infant mortality, for this assumption will not always be met if research is extended to child mortality. At the same time, death

among younger siblings may influence the survival of both older and younger siblings. Finally, as discussed above, death clustering implies that there is a stronger concentration of early-life deaths in families than would be expected. As the scarring models often include only the effect of death of the adjacent sibling, it is unclear to what extent there is an unusual concentration of deaths in these families. For an exception and a discussion of higher-order Markov models to solve this issue, see Omariba et al. (2008).

A caveat should be made here, however. Not only is the mechanism behind scarring effects often unclear and the method not completely sufficient, but there have also been favorable developments in infant and child mortality rates in less developed countries over recent years. In addition, families have become smaller on average (UNICEF et al., 2014). Therefore, scarring effects on siblings will have become less relevant for health and survival. This suggests that part of the decline in mortality in modern-day less developed countries is caused by a self-reinforcing mechanism: because fewer children die, the health of fewer siblings is damaged, causing further reductions in child mortality. Why variation exists between families in the total experience of early-life mortality, then, becomes an even more important question.

To conclude this section, research into the effect of the death of a sibling on individual survival has generated important insights into the antecedents of excess child loss in the past and in the developing world. At the same time, the frequently cited explanation —scarring—remains a "black box" explanation, as the mechanisms through which brothers and sisters are molded by the death of their sibling remain unclear in most research. Furthermore, research has neglected the opportunity to assess how the effects of sibling deaths differ between families of varying backgrounds. Existing research efforts can relatively easily be extended to do this, for instance, by the application of simulation models or by including family-level heterogeneity in models and research reports.

4. Characteristics of families and mortality clustering

A second group of studies have incorporated family characteristics more explicitly in the definition and measurement of mortality clustering and in the characteristics that are linked to early-life mortality. In early research, death clustering often referred to families experiencing multiple infant or child deaths. Other research has used heterogeneity between families in their children's chances of dying in their early years and determined how much variation in early-life mortality between mothers is explained by the inclusion of explanatory characteristics. This approach allows for the inclusion of variation at both the family and community levels. Finally, researchers have used excess deaths in families at the population level to explore family-level causes of excess mortality, or used simulation models to assess whether excess deaths are explained by the family characteristics that are known determinants of early-life mortality. The main results from studies using these approaches can be found in Table 2 under approaches 2–4.

Explanations for differences in mortality risk between families have been sought in fertility behavior, socioeconomic status, childcare practices, and maternal health and biological characteristics. With regard to fertility behavior, multiple pregnancies, especially if closely spaced, may result in maternal depletion. Furthermore, in larger families, children compete for the limited resources available to the family. Zaba and David (1996) have shown that the variability between mothers in the risk of their children dying—an indication of mortality clustering—is highest among mothers at the highest parities. Several mechanisms could explain this relationship. First, mothers who experience more child deaths tend to proceed to higher parities. Second, life-history theory predicts that individuals exposed to high mortality will reproduce early, have shorter birth intervals, and invest less in their children, which is essentially predicting mortality clustering among mothers exposed to high-mortality environments (Störmer & Lummaa, 2014).

Other explanations have been sought in the realms of socioeconomic status and childcare practices (Das Gupta, 1990, 1997). In India, death clustering has been partially explained by the caste membership of mothers, which is related to poverty and access to health care (Ranjan et al., 2018). Lalou (1997) found that infant mortality was higher and birth intervals shorter in the higher classes, which may be explained by their practice of using wet nurses. In addition, in the lower classes, inadequate feeding patterns may have led to poor health among mothers, resulting in repeated problems during delivery, repeated stillbirths, and high maternal mortality (Lalou, 1997). Janssens and Pelzer (2012) found that mothers who had worked in factories, who at the time were assumed to turn into "bad mothers," did not do worse than other women. Das Gupta (1997) acknowledged that there was more variety in parenting behavior among the women with relatively low levels of education in her sample than among those with higher levels of education, resulting in higher mortality among the children of women with lower education. This implies that maternal care plays an important role in child mortality, but that there is also an overlap with socioeconomic status. Although there is little evidence for a consistent relationship between socioeconomic status and childhood mortality for much of the 19th century (Janssens & Pelzer, 2012), Scalone et al. (2017) found that, in 19th century Italy, sharecroppers were less likely to experience mortality clustering than landless laborers and nonrural

workers. This indicates that co-resident kin—common among sharecroppers—decrease the likelihood of mortality clustering. However, it is unclear to what extent these families co-resided with their kin; moreover, economic deprivation and seasonal migration may be alternative explanations for the higher mortality levels of the landless laborers and nonrural workers. Thus, future research should attempt to distinguish carefully between the role of proximate kin and the role of socioeconomic conditions.

With regard to social and biological characteristics, Pavard et al. (2005) found that maternal death decreased the chance of survival of the children affected, especially girls, which suggests that the presence of kin and maternal care both play an important role in survival. Similarly, Edvinsson et al. (2005) found that both mattered in determining risk status: remarriage after the death of a spouse (interpreted as a largely social characteristic related to instability of the family) and the experience of stillbirth (considered by the authors to most likely have been caused by biological mechanisms including health status and genetic factors) were both related to the likelihood of being a high-risk mother. Willführ and Gagnon (2012) used a Lorenz curve and Gini coefficients for the expected and observed distributions of deaths over the population, and showed that there was less clustering in Krummhörn (Germany) than in Quebec (Canada). The extent to which deaths cluster differed between recomposed and nuclear families, showing that in Quebec maternal care

abilities may have played a role in mortality clustering in the families of remarried widowers (Willführ & Gagnon, 2012). Furthermore, consanguineous marriages, which are common in some parts of the world, including Pakistan and India, are related to excess early-life mortality (Hussain et al., 2001). Earlier research has suggested that frailty—and hence the risk of death—is caused by genetic disposition (Vaupel et al., 1979; Vaupel, 1988; Yashin et al., 1995). However, in the absence of genetic information, it is difficult to attribute remaining family variance to genetic causes. Remaining familial heterogeneity may be considered the upper limit of the influence of genes shared in families.

Furthermore, it has been shown that there is an intergenerational component to death clustering (Lindkvist & Broström, 2006; Vandezande & Matthijs, 2013; Quaranta et al., 2017). Parents who lost many siblings in infancy are more likely to have a high-mortality family themselves. Thus, mortality clustering is transferred between generations. The causes of high mortality in the first generation—be they social, biological, or economic—appear to be transmitted to the next generation (Vandezande, 2012). Alternatively, exposure to high mortality in the family of origin itself may be a cause of faster reproduction, lower investments in children, and higher early-life mortality in the second generation, as life-history theory predicts. It remains unclear, however, which characteristics explain this intergenerational transmission of mortality clustering. It is interesting to note that intergenerational transmission occurs through both the paternal and maternal lines (Vandezande, 2013). The influence of fathers on family-level mortality beyond socioeconomic status has been shown in only one earlier paper. Van Bodegom et al. (2012) used a polygamous community in Ghana to show that variation in mortality could be found at the paternal and maternal levels separately. Mortality clustering research commonly focuses only on mother-level variation and explanations, but these papers have shown that the father plays a larger role than is commonly assumed. At the same time, earlier work has shown that there are large differences between women in the same polygynous unions (Last, 1992). In comparison with monogamous unions, the offspring experience survival disadvantages in later childhood (Gyimah, 2009). Furthermore, other kin may affect survival as well, as earlier research has shown that maternal grandmothers improve the survival chances of their grandchildren (Beise & Voland, 2002; Sear et al., 2002).

Family-level studies often include individual-level explanations for early-life mortality and these may interact with family-level characteristics. In earlier research, it has been shown that heirs often fare better than other children (Voland, 2000) and that parents may show bias toward one sex at the expense of the other (Pavard et al., 2005; Arulampalam & Bhalotra, 2008; Janssens et al., 2010). As discussed earlier, families facing economic hardship may be more likely to lose their children and infants. However, the burden may be divided unequally among the family members. For instance, during food shortages, boys' nutritional needs may be better met than those of girls. Similarly, the dependency ratio of the family results in resource dilution: resources run thin in large households, especially if the children are still young. Who suffers from this lack of resources depends on the way in which resources are divided among household members. In addition to family size, other household characteristics, including household earning capacities and parental competence, are also likely to change over the life course of the family. The effects of these household characteristics on children depend on their sex and birth order, as these are related to their responsibilities within the household. Thus, household characteristics impact on the likelihood of death among infants and children in a complex fashion and, moreover, may change over time.

Interaction models focusing on both child and family characteristics may help us to achieve a more thorough understanding of these processes, by explaining family heterogeneity in mortality risks. However, models using random variation between families assume that all children within the household are influenced in a similar fashion. By incorporating interaction terms between family characteristics (including their risk status) and individual characteristics, these processes may be explored in future research.

To summarize, a focus on mother heterogeneity and familial risk status has several advantages: it clarifies why families differ in the likelihood that mortality occurs among their infants and children, and it enables competing explanations to be tested simultaneously (Edvinsson et al., 2005). Furthermore, community effects can be modeled in conjunction with family effects and researchers may assess how much random variation between families is explained by their models. At the same time, when the focus is on variation between families, it remains unclear to what extent family differences in mortality levels reflect chance fluctuations within communities. While social, economic, and biological mechanisms have all been shown to play a role in mortality clustering, it is still not clear: (1) whether the same explanations apply for clustering in the past as for clustering in modern-day less developed countries; (2) which explanations matter most; and (3) which children are most at risk in high-risk families.

5. Mortality clustering in populations

Mortality clustering as a phenomenon shows that mortality is concentrated in families, affecting the early-life survival chances of their children, and manifests itself at the population level as an excess concentration of deaths in some families. Although the methodological angles discussed in the previous two sections focus on the child and family levels, respectively, most papers use population-level measures to show the extent to which mortality clustering manifests itself in the populations under examination. Several researchers have taken a step beyond that and managed to explore to what extent death clustering in the population can be explained by the characteristics of children and families. Results from these papers can be found in Table 2, under approach 4. To determine the relevance of the explanations given and whether we should continue to search for new explanations, we need to assess whether there is still excess clustering in families (exceeding the clustering as predicted by chance) after taking our explanations into account. In other words, we should return to the population level. Holmberg and Broström (2012) have proposed using simulation models to find the extent to which mortality clustering in the population is explained. Using this method, they first account for family size and use the binomial chance distribution to find the extent to which deaths cluster in a population. Second, models with random effects are estimated to analyze the causes of mortality clustering. Third, these explanatory variables are incorporated in simulation models to find the expected number of deaths, which is compared to the observed number of deaths. Thus, the extent to which there is still excess mortality after including known explanations for mortality clustering is assessed.

While most studies into early-life mortality clustering assume that it can be found in any population, death clustering seems to play a larger role in some communities than in others (Van Poppel et al., 2012; Vandezande, 2012). For northeast Brazil, Sastry (1997) has shown that taking community-level variance in mortality into account explains almost all variance between families in the number of children who succumb, suggesting that mortality clustering may be overestimated in some contemporary research. Strikingly, Saha and van Soest (2011) used a random cluster analysis to show that in Bangladesh, the main source of variation in early-life mortality is found in families, not in communities. Similar results were found by Manda (1998). This means that, here, families play a more important role in death clustering than their communities.

The causes of community variation have rarely been the subject of research. In some rare examples addressing differences between populations, it has been proposed that some populations are culturally quite homogenous and socioeconomically comparatively equal, which results in relatively little inequality in mortality between families (Curtis et al., 1993; Guo, 1993). For instance, Guo (1993) found only small differences between families in Guatemala after accounting for varying income levels and suggested that, in high-mortality populations with relatively low socioeconomic inequality, unexplained variance is low because mortality resulting from poverty and community effects is still very high. In modernizing societies, on the other hand, varying levels of access to health care and new health interventions could increase inequalities between families, as better-off families are better equipped to take advantage of these new services (Saha & van Soest, 2011). However, other studies have found unexplained mortality clustering in relatively high-mortality populations, such as the Netherlands in the 19th century (Van Poppel et al., 2012), indicating that increasing inequalities after the start of the epidemiological transition are not the sole cause of mortality clustering. Spatial factors including access to safe sources of water (Van Bodegom et al., 2012) and rural–urban differences (Vandezande, 2012) have been shown to play a role in mortality clustering in high-mortality contexts. In addition, cultural factors such as attitudes toward breastfeeding may play a role (Pelzer & Janssens, 2014). In future research, both population- and family-level characteristics need to be taken into account to clarify whether significant familial mortality clustering exists, or whether populations play a more important role than previously thought.

Another advantage of a stronger focus on the role of community characteristics in mortality clustering would be the insight delivered into the contextual determinants of changing mortality patterns, such as the causes of the epidemiological transition. Generally, historical demographers and international development researchers have

failed to make use of the opportunity to explain the epidemiological transition better by focusing on the changing characteristics over time of families who lose their infants and children. In other words, the experiences of families who have carried the burden of mortality—or escaped early-life mortality—sharpen our understanding of mortality in the past and in the modern-day developing world, especially if we take longitudinal changes in the causes of clustering into account. Furthermore, the role of context in high familial mortality helps to explain why mortality clusters in some families but not in others. Why is it that some families are especially vulnerable to adverse circumstances and how does that change over time? To explore that, we need to research the interaction between family and community characteristics. Thus, we need to find out which families experience excess deaths and why. This strategy may help to explain why death clustering is more common in some populations than in others and to identify the determinants of death clustering at the family level.

6. The way forward and concluding remarks

After 25 years of research on infant and child mortality clustering, important insights into the causes of early-life mortality clustering have been generated. The research has shown that in populations around the world and over time, there are differences between families in the degree of risk of mortality among their children. It has become clear that, taking the varying risk between families into account, the death of an older child harms the survival chances of younger children in that family and that fertility behavior, earlier stillbirths, remarriages, and socioeconomic status all explain mortality clustering to some extent. However, many questions on the nature of the phenomenon remain and several promising avenues in mortality clustering research have been neglected in research to date. In that sense, death clustering remains a black box: to open it, we need to concentrate on the pivotal role played by the families in which deaths cluster.

First, social, economic, and biological mechanisms have each been shown to play a role in mortality clustering; and birth intervals, birth order, age of the mother, and other demographic characteristics also matter, as does an effect of scarring from sibling death. However, it remains unclear whether the same explanations apply for clustering in the past as for clustering in modern-day less developed countries, as systematic comparisons are lacking. In high-mortality societies, spatial differences appear to play a more important role, whereas in transitioning and low-mortality societies, inequalities between families are more important. These community-level characteristics are an underexplored but promising avenue for future research. Modernization of economies eventually benefits most individuals, but especially during the initial phases of modernization, the benefits may be limited to some families; others may see no benefits and modernization may even have a negative influence for some, increasing inequalities in early-life mortality. The interaction between shifting conditions over time and the extent to which deaths cluster in families is an important avenue for future research, one that has remained unexplored so far. Furthermore, because the levels of child and infant mortality differ between populations, as does the extent to which deaths cluster in families, we should develop strategies to take the varying levels of mortality between populations into account in our models. Currently, part of the inequality in mortality shown by our models may be caused by the accidental attribution of community-level variance to the family level.

Second, which explanations matter most and for whom—which children, of what age—remains as yet unknown. More innovative questions may be asked if researchers focus on the interactions between the analytical levels that were addressed earlier in this paper. The child's characteristics may interact with those of their family, which may be more beneficial to some children than to others. Similarly, the community environment in which the family lives may be more challenging for some families than for others and, at the same time, have a more detrimental effect on some children within that family than on others, such as on later-born children or girls. These interactions between analytical levels have not yet been taken into account in a systematic way. Finally, it is not yet known whether the causes of mortality are similar for high- and low-mortality families. Do the same characteristics put infants and children at risk in both high- and low-risk families, or are there specific explanations for mortality in high-risk families? These questions are essential for making progress in the field of early-life mortality clustering in families.

Third, at the center of most death-clustering research, we find the nuclear family; family relationships beyond parents and children in a nuclear household are seldom considered in such research. However, it has been shown that intergenerational transfers of mortality clustering may apply if the causes of mortality clustering are transferred between generations (Vandezande, 2012). Furthermore, conjugal and extended families may be more important in populations in transition than previously acknowledged in demographic and development research. This means that adults (other than biological parents) who co-reside with families should not be excluded from our theories and models. In addition to vertical (intergenerational) similarities in mortality clustering, horizontal similarity to siblings may enhance understanding of the phenomenon: siblings may share learned behavior, socioeconomic status, environments, and biological difficulties, resulting in similar patterns of mortality among their children. Finally, although most research calculates rates of mortality

clustering based on mothers, it has been shown that fathers may play an important role too and, therefore, should be considered in more detail. The practical implications of these suggestions for innovative research questions concern the tailoring of policy interventions. If mortality clustering is socially contagious or inherited, policy interventions aimed at combating early-life mortality should focus on the family members of people who experience mortality clustering among their children, rather than on all families or only on the mothers and siblings of parents in whose families mortality clusters.

The limits for answering questions such as these have often been set by the availability of suitable data. In sources concerning developing-world populations, such as the Demographic and Health Survey (DHS) and digitized censuses, individuals are surveyed about their reproductive histories and household characteristics. These databases are often rich in detail, with intricate insights into the socioeconomic status and knowledge of parents and the conditions surrounding births. Furthermore, similar census and survey questions are asked in many countries, allowing for international comparisons. Most research on death clustering focuses on one or two regions only, so including more regions in research, using similar data sources and approaches in particular, would enhance our understanding of the phenomenon (for an example, see McMurray, 1997). Several drawbacks of these data should be mentioned as well. The retrospective reconstruction of reproductive histories is sensitive to cultural differences, social desirability, memory, and emotional sensitivities. In addition, often only women are interviewed, while the paternal perspective is neglected. Finally, women who did not survive their reproductive period are not included in cross-sectional surveys.

In historical demography, the use of censuses and population or parish registers has been common. These datasets often give insight not only into an individual's life course, but also into those of their parents, grandparents, and offspring. Traditionally, these datasets were often small in scope, concerning only one village or region over a short period of time. Therefore, they often did not contain information about close relatives residing outside the region of concern, or (if a sample was taken) who did not live in the same household, therefore limiting the scope of research based on these data sources. In recent years, the digitization of sources, including civil certificates and population and parish registers, has led to the increasing availability of high-quality data sources on individuals. By linking information from various sources, the scope of historical demographic datasets has increased, allowing for comparisons between siblings, research on the linked lives of multiple generations, and comparisons of populations in multiple (sub)regions. Thus, new areas for research have been opened up by the increasing computerization of data.

Fourth, many questions remain with regard to the period of death. While some authors choose to focus on postneonatal deaths (e.g., Curtis et al., 1993; Reid, 2002), others focus on neonatal deaths (see, e.g., Bhalotra & van Soest, 2008; Nonyane et al., 2013). Yet others include all perinatal deaths, including stillbirths (Reid, 2001; Lundevaller & Edvinsson, 2012). It has been argued that the timing of death is related to its cause. Postneonatal death may be related more to factors such as kin support (Sear & Mace, 2008; Sear & Coall, 2011), childcare (Edvinsson et al., 2005), and other exogenous characteristics (Lalou, 1997; Beise & Voland, 2002), while perinatal death is assumed to be related more to factors experienced during pregnancy and to health and genetic problems (Lalou, 1997; Reid, 2001; Vandezande, 2012; Pakot, 2015). For instance, Reid (2001, 2002) found that a woman's history of child deaths was related to further postneonatal mortality of her infants, especially that caused by wasting, diarrhea, and respiratory diseases. At the same time, her history of stillbirths and miscarriages affected the likelihood of repeated stillbirths and miscarriages in the future. This implies that early-life mortality clustering in some women is related to endogenous or perinatal factors, whereas in other women it may be related to exogenous factors including behavior.

However, this argument does not always hold: social and cultural practices may play a large role in the first weeks of life as well (Nonyane et al., 2013). For instance, in 19th century Iceland, the incidence of neonatal tetanus was extremely high through infection of the umbilical stump (Garðarsdóttir, 2002). Similarly, Lalou (1997) showed that neonatal mortality peaked during an epidemic of smallpox. In the literature, a systematic approach to the timing of death is often lacking and infant deaths seem to be more commonly addressed in the field than early-life child mortality (see Table 2). A way forward may be found following research that addresses the influence of sibling mortality in several periods of early life separately. For instance, Alam and David (1998) have shown that infant deaths decrease the likelihood of toddler deaths in families, but toddler deaths increase the likelihood of infant deaths, illustrating that the ages at death of siblings are related. Alter et al. (2001) took mortality over the life course into account, showing that deaths of siblings were most strongly related in early childhood, weakened after age 15, and disappeared after age 30. Children from high-mortality households tended to survive longer, pointing to an acquired immunity effect. Furthermore, research into the causes and timing of death of children in high-mortality households may help us to disentangle whether causes are behavioral (such as gastrointestinal and external causes of death) or biological (such as genetic and cardiovascular diseases), or should be understood as caused by the transmission of disease between siblings if, for instance, multiple siblings succumb to the same cause of death in a short time frame (Kippen, 2011). In addition, researchers could investigate the extent to which the factors related to mortality clustering are related to

specific causes of death instead of all-cause mortality. Causes of death have only rarely been incorporated in research into the mortality of siblings (for examples, see Reid, 2001, 2002; Kippen, 2011; Saha et al., 2014).

Finally, early-life mortality clustering is an exciting field of research, with many avenues yet to be explored. Addressing the questions explored in this review of the literature on mortality clustering would provide us with important insights into the demographic transition and help us to understand current and future developments with regard to early-life mortality and health in the developing world. Although data sources and models remain highly complex and challenging for researchers, asking sophisticated new questions and applying more thorough analysis to existing data should help the determinants of early-life mortality and the role of families become much clearer in the near future. The existing models and innovative research approaches that have been discussed in this review can be used to help us answer these questions.

Paper III

Like Mother, Like Daughter: Intergenerational Transmission of Mortality Clustering in Zeeland, the Netherlands, 1833–1912

Historical Life Course Studies (2018), 7(2), 28–46.
Special Issue on Intergenerational transmissions of infant mortality using the Intermediate Data Structure (IDS). Editors: Luciana Quaranta and Hilde L. Sommerseth

Authors
Van Dijk, I.K.
Mandemakers, K.

Identifier
http://hdl.handle.net/10622/23526343-2018-0003?locatt=view:master

An earlier version of this work has been presented at the Conference of the ESHD, Leuven, Belgium, in September 2016 and at the Conference of the PAA, Chicago, US in April 2017. The paper has benefited from the comments of Luciana Quaranta, Hilde Sommerseth, and Paul Puschmann; of participants to the Special Issue in Historical Life Course Studies; and from Sören Edvinsson, Isabelle Devos, Jan Kok, Eline Slagboom, Niels van den Berg, and Rick Mourits.

Abstract

The burden of infant mortality is not shared equally by all families, but clusters in high-risk families. As yet, it remains unclear why some families experience more infant deaths than other families. Earlier research has shown that the risk of early death among infants may at least partially be transmitted from grandmothers to mothers. In this paper, we focus on the intergenerational transmission of mortality clustering in the Netherlands in the province of Zeeland between 1833 and 1912, using LINKS Zeeland, a dataset containing family reconstitutions based on civil birth, marriage, and death certificates. We assess whether intergenerational transmission of mortality clustering occurred in Zeeland, and if so, whether it can be explained on the basis of the demographic characteristics of the families in which the infants were born. In addition, we explore the opportunities for comparative research using the Intermediate Data Structure (IDS). We find that mortality clustering is indeed transmitted from grandmothers to mothers, and that the socioeconomic status of the family, the survival of mothers and fathers, and the demographic characteristics of the family affected infant survival. However, they explain the heterogeneity in infant mortality on the level of the mother only partially.

Keywords Mortality clustering, infant mortality, intermediate data structure, biodemography, early-life exposure

1. Introduction

The burden of infant mortality is not shared equally by all families: some families lose a considerably larger share of children in infancy than others, a phenomenon known as early-life mortality clustering. In earlier research, it has been shown that death clustering has an intergenerational component, as surviving children from high-mortality families experience higher mortality among their own infants than individuals who grew up in low-mortality families (Lindkvist & Broström, 2006; Vandezande, 2012). However, as yet, it is unknown how family histories of high infant mortality can be explained. The aim of this paper is threefold. First, it is to determine whether intergenerational continuities between grandmothers and mothers in infant mortality occurred in the province of Zeeland, the Netherlands, in the 19th and early 20th centuries. Second, we aim to disentangle the contribution of the socioeconomic status of the household in relation to mortality clustering among mothers *vis-à-vis* mechanisms of intergenerational transmission that work through exposure of mothers to sibling mortality and affecting their reproductive careers. Our third aim is methodological: we use the Intermediate Data Structure (IDS) for our research and reflect on opportunities for comparative research using this format.

Intergenerational transmission in increased mortality among infants has been found for several historical populations (Quaranta et al., 2017). This shows that death clustering is not limited within one generation but has roots in earlier generations. However, in earlier work, explanations concerning intergenerational correlation within infant mortality have remained largely unaddressed and, as yet, it is unknown how these family histories of high mortality come about. Several explanations for this phenomenon are possible, which can roughly be divided into indirect and direct effects. On the one hand, grandmothers and mothers may share certain characteristics—including social and behavioral characteristics and genes—which are related to high infant mortality, causing high mortality in multiple generations (Vandezande, 2012). On the other, exposure to mortality in the family of origin itself may affect the likelihood that future mothers experience infant mortality (Nettle, Coall, & Dickins, 2011; Störmer & Lummaa, 2014).

In the past, data constraints often limited opportunities for conducting research into intergenerational similarities in infant mortality and survival. Here, we use LINKS, a large-scale historical database that contains up to seven generations of individuals experiencing at least one vital event in the province of Zeeland to reconstruct life courses of grandmothers and mothers and assess mortality hazards of infants in an intergenerational fashion. We aim to assess which factors play a role in the intergenerational transmission of high infant mortality from grandmothers to mothers.

2. Theory

Almost 25 years ago, Monica Das Gupta (1990, 1997) recognized that child mortality clustered in a select group of families, despite high overall mortality rates in the Indian region that she was studying. Some families raised all their children to adulthood while others lost many children under the age of 5. Although this variation was recognized in earlier research (for an overview, see Das Gupta, 1997; Edvinsson & Janssens, 2012), Das Gupta's work put death clustering strongly on the research agenda. Differences in maternal ability were identified by Das Gupta as one of the prime causes of differences between families' infant mortality levels among their infants and young children. Over recent decades, it has been shown that early-life mortality does indeed cluster in a select group of high-risk families (Curtis, Diamond, & McDonald, 1993; Lynch & Greenhouse, 1994; Janssens, Messelink, & Need, 2010), including families in the province of Zeeland (Van Poppel, Bijwaard, Ekamper, & Mandemakers, 2012). These findings illustrate that some families face hardships that other families do not, and shift our understanding of early-life mortality from the level of the individual child to the level of the families in which they are embedded.

Earlier research has suggested that there are structural differences between families in which deaths cluster—high-risk families—and families losing relatively few or no children. For instance, family size (Zaba & David, 1996), maternal ability (Das Gupta, 1990, 1997), maternal death (Pavard, Gagnon, Desjardins, & Heyer, 2005), remarriage of the mother (Edvinsson, Brändström, Rogers, & Broström, 2005), earlier stillbirths (Edvinsson et al., 2005), and Rh disease (Lundevaller & Edvinsson, 2012) have been related to families' excess risk of early deaths (for an overview, see Van Dijk, 2018). In addition, it has been shown that an increased mortality risk of infants is related between generations in the maternal line, which suggests that family histories of high mortality play a role in mortality clustering (Vandezande, Moreels, & Matthijs, 2010; Vandezande & Matthijs, 2013). However, earlier work explanations for the intergenerational correlation in infant mortality have remained largely unaddressed. In this paper, we address direct and indirect mechanisms that may affect infant mortality in multiple generations, working through shared social characteristics between grandmothers and mothers or through exposure to mortality in the family of origin, and thus affecting life courses and infant survival.

With regard to the characteristics of the family, the socioeconomic status of the family in particular may be relevant (Janssens & Pelzer, 2012). Earlier research has shown that there existed a socioeconomic gradient in infant mortality in the Netherlands, with lower infant mortality among the middle and higher classes than the lower classes (Van Poppel, Jonker, & Mandemakers, 2005). It is hypothesized that this gradient in infant mortality increased over the course of the 19th century, possibly due to healthier living environments for the upper classes (Van Poppel et al., 2005) and was caused by a slow spread of knowledge about health and access to medical innovations from the upper classes to the lower classes (Antonovsky & Bernstein, 1977). Some evidence has been found for a relationship between socioeconomic status and death clustering, as the variability in child mortality was greater between low-educated and low social class women in India (Das Gupta, 1990, 1997) and mortality clustering was lower among sharecroppers in 19th century Italy than among landless laborers and nonrural workers (Scalone, Agati, Angeli, & Donno, 2017). The effect of socioeconomic status may not be limited to one generation, as the socioeconomic status of the grandparents may be reflected in the socioeconomic characteristics of the mother through a process of social reproduction (Zijdeman, 2009). In addition, the socioeconomic conditions of the grandparent may affect health of the mother, including her reproductive health and caretaking capabilities.

In addition to effects of socioeconomic status through social reproduction, exposure to high mortality itself may also directly affect future mothers, leading to increased mortality in several generations (see Figure 1). Several pathways may exist. The health of girls raised in high-mortality households may be damaged by childhood adversity, including lack of nutrition and exposure to infectious disease in the parental home. Adults whose health is compromised by growing up in high-mortality environments may live shorter and unhealthier lives (Barker, Eriksson, Forsen, & Osmond, 2002; Crimmins & Finch, 2006; Lindeboom, Portrait, & Van den Berg, 2010). As infants and young children are highly dependent on their parents, the health and survival of parents is essential for the survival of young-age offspring. Furthermore, childbearing capacities of women may be damaged, resulting in reproductive problems, stillbirths, and child loss in adulthood (Lalou, 1997). Thus, exposure to high mortality may be related to high mortality among one's own offspring in adulthood, explaining why intergenerational continuities in high infant mortality between grandmothers and mothers may exist.

Alternatively, evolutionary biology emphasizes that exposure to childhood adversity affects life histories, including reproductive careers and infant mortality. Individuals who are exposed to childhood hardships, such as food shortages, or stressful events in childhood, including mortality of siblings, are expected to employ fast life strategies for reproduction, to increase the chances of successful reproduction in risky environments. These exposures are therefore related to a lower age at menarche, a lower age at first marriage, a lower age at first birth, and shorter-spaced births of children (Nettle et al., 2011; Störmer & Lummaa, 2014). Furthermore, as part of this strategy, individuals are expected to invest fewer resources in individual children and

experience higher infant mortality (Chisholm, 1993; Störmer & Lummaa, 2014). Through the reproductive strategy of the mother, grandmothers' experiences of infant mortality may be linked to the survival of their grandchildren. This paper is part of a larger project that investigates the extent to which intergenerational transmission of mortality clustering occurs in several historical populations in Europe (see also Broström, Edvinsson, & Engberg, 2018; Donrovich, Puschmann, & Matthijs, 2018; Quaranta, 2018; Sommerseth, 2018). The analyses are partially conducted using a common database format and are based on a common script for the selection of cases, the construction of key variables, and the analysis of data (see Quaranta, 2018), based on a release of the database LINKS Zeeland in the Intermediate Data Structure (hereinafter, IDS). The IDS stores datasets originating from different sources and databases into a similar structure, which enhances opportunities for comparisons between datasets and therefore between regions and periods (Alter, Mandemakers, & Gutmann, 2009; Alter & Mandemakers, 2014).

The results of common analyses conducted for several historical populations in Europe will be discussed in more depth elsewhere. However, most importantly, after selecting cases on similar criteria, constructing databases in a similar way, and using identical analyses, we find that there is indeed intergenerational transmission of mortality clustering in several historical populations. These effects are stronger for grandmothers who lost two or more children than for grandmothers who lost one child in infancy. Furthermore, the effects are not explained by the demographic characteristics of the family. Because of the similar procedures for data selection, we can exclude the possibility that differences between populations are caused by differences in data construction between populations. In the conclusions of this paper, we will discuss our experience in working with the IDS in the LINKS Zeeland database.

In the current work, the focus lies on intergenerational transmission of mortality clustering and its causes in the province of Zeeland. We contribute to the literature by including the role of socioeconomic status, fast life histories, and the compromised health of mothers in intergenerational family histories of high infant mortality. Now, we turn to the data construction.

3. Data

The data analysis is based on LINKS, a database that contains vital events of the population of Zeeland, the Netherlands, between 1812 and 1912 (Mandemakers & Laan, 2017). The used data are released in the format of the Intermediate Data Structure (IDS) version 4, a common format intended for longitudinal databases concerning historical data (Alter & Mandemakers, 2014).

3.1 The historical context

Zeeland is a largely Protestant province located in the southwestern coastal region of the Netherlands, and consists mainly of islands. In the 19th century, the province was predominantly rural and agricultural, and the economy was mainly dependent on market-oriented, capital intensive farming, focused on growing wheat, flax, and madder. Although the majority of the population belonged to the unskilled agricultural-worker class, wealthy farmers constituted a small but prosperous share of the population (Bras & Neven, 2007). At the end of the 19th century, the agricultural sector went through a prolonged crisis, which was especially felt among commercial farmers (Beekink, Boonstra, Engelen, & Knippenberg, 2003). Industrialization occurred relatively late (Blom & Lamberts, 2006), as it only took off in Zeeland after 1900 (Van Poppel et al., 2012).

Until water pipes were constructed in the province of Zeeland, from 1910 onward, in most municipalities rainwater was used as the main source of drinking water. Surface water was often unfit for human consumption due to salinization. Especially during warm summers, there could be a lack of drinking water. In addition, water in canals and cities was often unsafe and its use led to a high prevalence of waterborne diseases. Malaria was endemic (Hoogerhuis, 2003), and the life prospects of infants and children were short, with a child mortality rate reaching more than 40% in some years (Van Poppel & Mandemakers, 2002; Hoogerhuis, 2003). At the time, the high infant mortality rate in Zeeland was blamed on women who did not breastfeed their children, as they were working outside the home, especially in agriculture (Hoogerhuis, 2003). In the province, the onset of the demographic transition occurred relatively late in comparison to other countries (Hofstee, 1981). Infant mortality began to decline after 1880, but remained high in an international perspective until the early decades of the 20th century (Van Poppel et al., 2005).

3.2 Data sources

The LINKS project aims at reconstructing all 19th and early 20th century families in the Netherlands, based on a digitized index of all civil certificates from this period called *WieWasWie* ("WhoWasWho"). For over 20 years, numerous volunteers have

been working to build the index, which contains not only the names of born, deceased, and married persons, but also the names of their parents, places of birth, ages, and, in part, their occupational titles. The database contains vital events of almost 2 million persons from the province of Zeeland.

Obligatory vital event registration was introduced in the Netherlands in 1812. The registers of birth, marriage, and death become public with delays of 100, 75, and 50 years, respectively. Indexes of the registers have been digitized completely for the province of Zeeland and have been released for the period 1812–1912 for births; 1812–1937 for marriages; and 1812–1962 for deaths. Unlike population registers, civil registers do not contain information about religion, addresses, and household composition. At the same time, the civil registers have a large scope, enabling the identification of kin over a large area and in several generations. Using family reconstitutions based on civil registries, we can apply an intergenerational approach to demographic research questions. In this case, we can follow vital events from families occurring in Zeeland over time for up to five generations.

In the LINKS dataset, civil certificates of individuals were linked together based on the first and last names of pairs of individuals. The linking process consisted of four steps. First, marriage certificates of children were linked to marriage certificates of parents, using the names and ages on the certificates. Second, death records were linked to birth records. Third, birth and death records were linked to the marriage certificate of the parents. Fourth, birth and death records were linked with marriage certificates of individuals. Small inconsistencies and double links—when certificates linked to several other certificates—were resolved by selecting exact links or by not accepting established matches. To avoid false matches, matching was limited within the time frame on which events logically may occur. For example, matched marriage certificates belonging to women outside their fertile years but with matching parental names were not accepted. Matches could be missed for several reasons: if vital events occurred outside the province of Zeeland, if events occurred outside the time frame for which certificates have been indexed, if names were changed, and if the variation in spelling of names was too large between civil certificates. Persons who migrated between municipalities in Zeeland could be followed over time; however, individuals who left the province of Zeeland were lost from observation. Analyses have shown that family reconstructions are reliable with regard to family size, age at first birth, and other demographic indicators (Van den Berg et al., 2018).

An exact death date is known for almost half of our sample (47.6% of births). For an additional quarter of the sample (25.0% of births) other follow-up observations exist for vital events such as marriage or childbirth in Zeeland. For 27.3% of the births in the sample, there is no follow-up observation. As our data consist of vital event certificates, the end of observation of marriages and deaths in 1937 and 1962 means that individuals who did not marry before 1937 or die before 1962 are not observed in the vital event registration after birth. In addition, individuals who out-migrated from Zeeland before the incidence of vital events are not found in the vital event indexes of Zeeland. Finally, for some individuals, a link between their follow-up vital event certificates and birth certificates has not been established. For individuals without follow-up observation, we assume that they survived until their first birthday.

3.3 Selection of cases

To prepare the data for analysis and to conduct the analysis, a common approach is used for several papers. This approach is based on the condition that all data are structured in the same way, following the Intermediate Data Structure (Alter & Mandemakers, 2014). On the basis of this structure, an extraction program, here developed by Quaranta (2018b), selects cases based on similar criteria, builds variables for analysis in an identical fashion, and specifies the basic analyses. For an extended discussion of the data selection and variable construction, see Quaranta (2018b).

The analysis focuses on the intergenerational transmission of mortality clustering, researching correlations in familial risk between grandmothers and mothers. Intergenerational transmission of mortality clustering encompasses a paradox. In extreme high-risk families, all children may die before they reach adulthood and/or are able to reproduce. To transmit high risk across generations, there must be a second generation that reproduces itself. Therefore, we conduct our analyses on a select group of families. First, there must be a second generation of mothers, which implies that (i) at least one female child is born, who must (ii) survive childhood and (iii) reproduce. In addition, (iv) grandmothers can only be selected if they were at risk of experiencing infant mortality and had a surviving, reproducing daughter, which implies that they must have produced at least two children.

For our analysis, we select singleton births in the province of Zeeland. The selected infants are the grandchildren of maternal grandmothers who married in 1812—the year in which vital event registration became obligatory in the Netherlands—or later. The second generation was born from 1812 onward, and the third generation (grandchildren) from 1833 onward, which is therefore the starting point of our observations (see Figure 1). Observations of births end in 1912. We include 203,802 infants in the analysis, who are born to 43,582 mothers and 28,118 grandmothers. This means that for a considerable number of grandmothers, several daughters are included in the analysis, which is taken into account in a robustness check (see Section 3.3). Lifeless reported infants are not included in the analysis. Lifeless

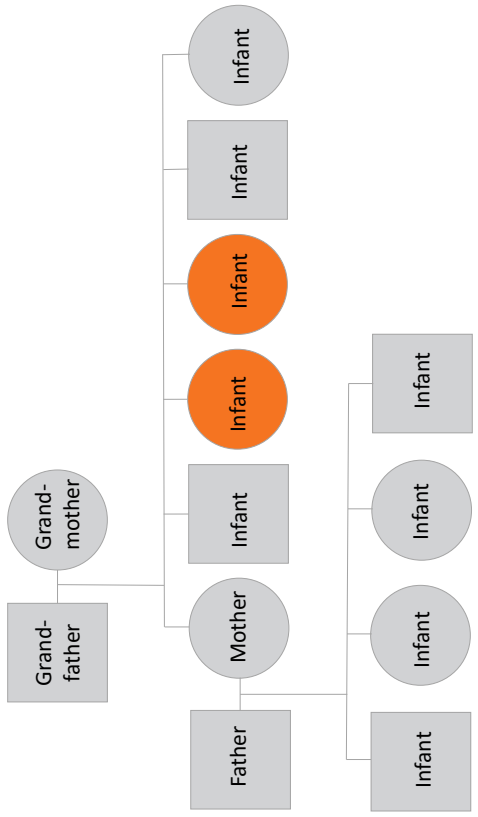


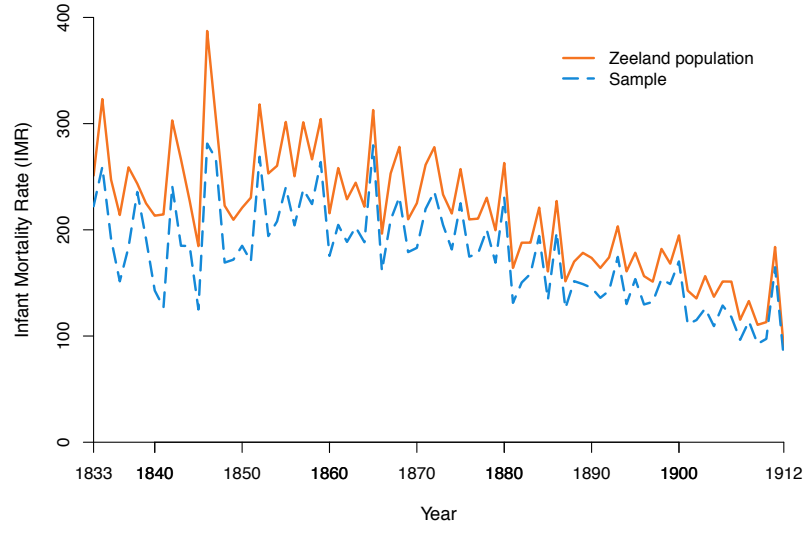
Figure 1 Multigenerational influences on infant mortality

Note: Infants who died before their first birthday are shown in orange.

registered infants concern 4.3% of the total number of births. These cases encompass both stillbirths and live births of infants who died before they were reported to the municipality, which had to be completed within three working days of the birth (Vulsma, 1988). Estimates suggest that 71% of lifeless reported infants concern stillbirths and 29% concern live births of infants who died before they were reported to the municipality (Van Poppel et al., 2012). As it is unclear which children were born alive and for how long they lived after birth, we exclude lifeless registered children from the analysis, which will lead to a slight underestimation of infant deaths in the neonatal period and to conservative estimates of the intergenerational relation in infant mortality risk.

Figure 2 shows the population infant mortality rate (IMR), based on the indexes of all birth and death certificates in the province of Zeeland and the IMR of the sample on which this study is based. Although the IMR in the sample is lower than in the population, changes at the population level are followed rather well. Differences are related to the specific characteristics of the studied sample, due to our selection of infants who can be linked to their mothers and their maternal grandmothers. Children born in 1833 to known grandmothers are born to grandmothers who married in 1812

Figure 2 IMR Zeeland population and sample LINKS, 1833–1912



Source: LINKS dataset Zeeland_2017_01 (IDS version).

or later; thus, both their grandmothers and mothers were relatively young, increasing their own likelihood and their offspring's likelihood of survival. In addition, in the early period of our sample, relatively strong deviations between the population and sample IMR can be seen; however, here, the number of infants in the sample is still very low, as these are the first children of the first children of the grandmothers. The low number of cases in these years results in relatively strong fluctuations of the sample IMR. Finally, we select infants originating from families who had been in Zeeland for at least two generations—as their grandmothers and mothers had married and reproduced in Zeeland—suggesting relatively strong residential stability. These children may have had the added benefit of access to kin networks, which may have contributed positively to child survival through assistance in child-rearing.

3.4 Variable construction

In Table 1, we present descriptive statistics for the included variables. Our analyses concern survival of the third generation, the infants, with the analysis centered on survival between birth and the first birthday. Of the infants selected in the sample, 16.5% died before their first birthday. The primary explanatory variable is the number of infant deaths of the maternal grandmother, which is categorized as follows: zero deaths among her infants (28.0%), one death (27.0%), and two or more deaths (45.0%). We control for the number of births of the grandmother, as grandmothers with more births are more likely to have experienced an infant death. As grandmothers must produce a reproducing daughter and be at risk of experiencing at least one infant death, the minimum number of observed births for the grandmother is two. The number of births has been categorized into four groups: two births (3.5%), three births (5.0%), four to six births (23.6%), and seven or more births (67.8%). The birth order of the infant is included, as there usually exists a J-shaped relation between birth order and risk of infant death: earlier and later births have a higher risk of mortality. This variable is similarly categorized, with the addition of a category for the first birth (20.8%). Further categories include the second birth (17.5%), the third birth (14.1%), the fourth to sixth births (27.3%), and the seventh birth or higher (19.8%).

Furthermore, we control for the age of the mother, categorized into three groups —15–24, 25–34, and 35–50 years old—as earlier research has shown that the risk of mortality differs between young mothers, older mothers, and average-aged mothers. To measure whether reproduction had started very early, in line with our hypotheses on fast life histories, we added a measure concerning whether mothers had their first child before the age of 20. In Zeeland, very early reproduction is rare: 1.3% of the infants have a mother who had her first child before the age of 20. In addition, we control for the sex of the child, as there are substantial differences between boys and girls in survival through their first year of life. The infant mortality rate fluctuates over

Table 1 The descriptive statistics

		Percentage	Mean	Min.	Max.
Infant deaths		16.6			
Grandmother child deaths			1.77	0	15
	Zero	28.0			
	One	27.0			
	Two or more	45.0			
Grandmother births			8.55	2	27
	Two	3.5			
	Three	5.0			
	Four to six	23.6			
	Seven or more	67.8			
Females		48.7			
Birth order of the child			4.17	1	22
	First	20.8			
	Second	17.1			
	Third	14.0			
	Fourth to sixth	27.7			
	Seventh or higher	20.4			
Age of the mother			30.5	15.4	50
	15–24	20.1			
	25–34	56.2			
	35–50	23.8			
Mother has first child young (<20)		1.3			
SES of the father			52	41	99
	Low	38.8			
	Low to average	28.1			
	Average	19.6			
	Average to high	9.9			
	High	3.7			
Father is a farmer		17.1			
SES of the grandfather			50	41	99
	Low	44.7			
	Low to average	29.1			
	Average	17.3			
	Average to high	6.2			
	High	2.6			
Grandfather is a farmer		19.6			
Death of the mother		.89			
Death of the father		.77			
Number of infants		203,802			
Number of mothers		43,582			
Number of grandmothers		28,118			

Source: LINKS dataset Zeeland_2017_01 (IDS version).

the course of the 19th century, slowly decreasing from 1880 onward (Wolleswinkel-Van den Bosch, 1998) (see Figure 2). We include the birth year of the child, which is centered on the mean birth year, 1884.

The socioeconomic status of the grandfather and the father are measured using the HISCAM scale (Lambert, Zijdeman, Van Leeuwen, Maas, & Prandy, 2013). All codes were provided by the HSN standardized, HISCO-coded and classified occupational titles, release 2013.01 (Mandemakers et al., 2013). The scale is divided into five categories because of possible nonlinear relationships between occupational status and infant mortality. The categories are made in such a way that the group sizes are approximately equal, with the exception of the group for the highest statuses, and include the groups low (HISCAM score 41–46), low to average (47–52), average (53–59), average to high (60–69), and high (70–99) socioeconomic status. Socioeconomic status is missing for 849 fathers and 1,432 grandfathers and is imputed with the mean. Socioeconomic status was measured using the highest known occupational title of the father and the grandfather. In a robustness check, we have included the lowest socioeconomic status of the grandfather and the father instead of the highest socioeconomic status (see Section 3.2). In addition, we include a measurement for the farmer status of the father and the grandfather, which is based on the HISCLASS scheme code 8 (Van Leeuwen & Maas, 2011). Finally, we include time-varying variables indicating the death of the father or the mother through the first year of life of the child. Infants and young children are highly dependent on family members for survival in their early life, and the death of a parent can therefore detrimentally affect the chances of survival of newborns and infants. We use the date of death of the parent to assess whether a parent died in the first year of life. If no date of death is known, we use the last date of observation of the parent. If the last observation is the birth of the child, we assume that the parent survives the first year of life of the infant[2].

4. Results

The analyses are divided into two parts. The first part is based on the code developed by Quaranta (2018b) and is conducted in Stata 14. Three survival models are estimated: (1) Cox regression models, (2) Weibull models with clustering on the mother, and (3) Weibull models with clustering on the grandmother. For a detailed discussion on methodology, see Quaranta (2018b). The models are controlled for basic demographic variables, including the age of the mother, the number of births of the grandmother, the birth order of the child, and sex. The second part of the analysis is conducted in R 3.32, using the package *coxme*. We estimate frailty models with clustering at the mother level and stratify our analyses by the year of birth of the child. First, we discuss whether we find mortality clustering in families in the Zeeland population.

4.1 Clustering of infant mortality

In measuring whether populations experience inequality in the number of early deaths between families, it is essential to take both family size and the binomial distribution of the probability of mortality into account. First, large families can be expected to lose more infants than small families. Furthermore, if early deaths follow a binomial probability distribution, the implication is that not all families will have the same realized mortality rate: there must be lucky large families who escape early death among their infants and unlucky small families who have to bury multiple infants. Even if one in five infants dies before their first birthday, chance predicts that more than 10% of large families with ten children would not lose a single infant. Thus, in assessing whether deaths cluster in the population, we should take chance variations and the size of the family into account, and therefore start analysis at the level of the population.

Figure 3 shows the expected and observed proportion of mothers experiencing a certain number of deaths. Here, we present figures for mothers with a total of four and eight births. The expected number of deaths has been calculated using the infant mortality rates in the population; observed numbers are taken from the data. The dashed lines represent the expected number and the filled lines the observed number of mothers with a certain number of deaths. The figure shows that the number of mothers experiencing relatively moderate numbers of deaths is below what would be expected, while there are more mothers experiencing no deaths at all, or relatively high numbers of deaths in their families. These results are in line with earlier, similar explorations (see, e.g., Zaba & David, 1996), and illustrate that in the province of Zeeland, deaths cluster in certain families with a relatively high mortality risk, while other families experience fewer deaths than would be expected or escape deaths among their infants altogether. In the next step, we turn to the data analysis.

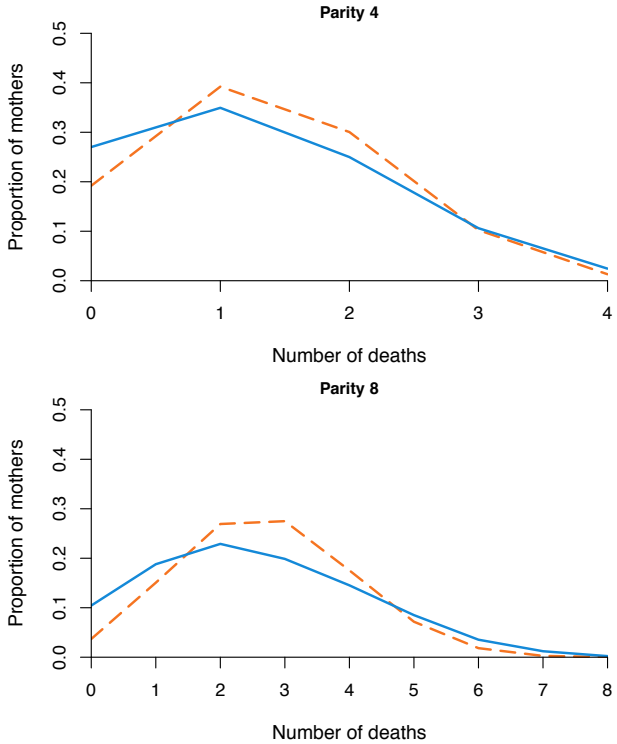
4.2 The main results

Figure 4 shows the cumulative hazard, based on a Cox proportional hazards model, for infants by grandmothers' number of infant deaths: zero, one, and two or more deaths. If the grandmother lost one infant, their grandchildren have a greater risk of death in infancy; if the grandmother lost two or more infants, this risk accumulates further. Estimated hazard ratios can be found in Table 2. All ratios are relatively stable

2 For some infants, observations on mothers (5.4%, $N = 11{,}052$) or fathers (4.5%, $N = 9{,}248$) end before the first birthday or the death of the infant. Here, we assume that the parents did not die before the first birthday or death of the infant.

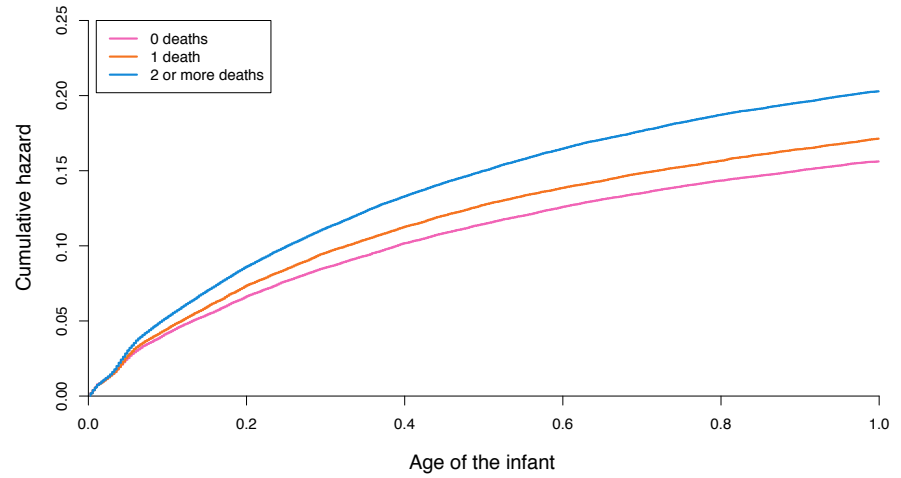
across all three models estimated: Cox models without shared frailty and Weibull models with mother´s and grandmother´s shared frailty. The risk of infant death is 12% higher if the grandmother lost one infant. In the case in which the grandmother lost two or more infants, this risk increases to a 34% higher risk of infant mortality. Thus, our main characteristic of interest, the number of infant deaths of the grandmother, is shown to be related to a higher likelihood of death among infants. The models control for the demographic characteristics of the family, including the number of births of the grandmother, the sex of the child, the birth order, and the age of the mother. In addition, we control for changes over time by including a measure for the birth year (centered).

Figure 3 The expected and observed numbers of families with a given number of deaths, by parity



Source: LINKS dataset Zeeland_2017_01.
Note: Dashed orange line, expected distribution; filled blue line, observed distribution.

Figure 4 The cumulative hazard by number of infant deaths of the grandmother



Source: LINKS dataset Zeeland_2017_01 (IDS version)

Furthermore, the models show that the sex of the child, his or her birth order, and the age of the mother contribute to the risk of infant mortality: boys, later-born children, and children from young mothers have a higher risk of dying in infancy. Without controlling for other characteristics of the family or child, the relative risk of infant mortality increases by 14% and 40% for infants with grandmothers who experienced one and two or more infant deaths versus grandmothers with zero infant deaths, respectively (results not shown). In addition, although a small share of frailty variance is explained, most variance between mothers concerning the likelihood that their infants will die remains unexplained. Thus, although the demographic characteristics of the family do contribute to the risk of infant mortality, they only partially explain the association between the number of infant deaths of the grandmother and mortality among her grandchildren.

Differences in intergenerational transmission of mortality risk may exist within families. Certain infants within families may be more strongly affected than others, as a result of gender preferences or biological differences. In most populations, females have a survival advantage over males during their life course. Excess mortality of girls in comparison to boys is therefore often interpreted as the result of gender preferences (Curtis et al., 1993; Bongaarts & Guilmoto, 2015). We have tested whether there is a significant interaction between the sex of the infant and the effects of the inter-

generational transmission of mortality risks. If intergenerational continuities in high infant mortality affect girls in particular, the effect can possibly be explained by preferences for boys over girls in multiple generations, leading to high mortality among girls due to infanticide or neglect. For the province of Zeeland, we do not find evidence that there is a sex-specific process that might involve gender preferences (results available on request), contrasting with some earlier work on survival of adolescent girls in Belgium (Devos, 2000) and one- to nine-year-old girls in the Netherlands (Van Poppel, Schellekens, & Walhout, 2009), where girls from elite families in particular experienced excess mortality.

4.3 The role of the survival of the parents and the socioeconomic status of the father and grandfather

In Table 3, results are presented for additional analyses using Cox regression models with mother shared frailty. In the first model, we analyze the effect of infant deaths of the grandmother controlling for her number of births and the gender of the infant. In the second model, we control for other demographic characteristics: the birth order of the child and the age of the mother at the birth of the child. In addition to the earlier models (Table 2), we add the age of the mother at first birth. In the third model, we add the socioeconomic status of the father and the maternal grandfather, the farmer status of the father and grandfather, and the death of the mother and the father in the first year of life of the infant. All models are stratified by the year of birth of the child, to account for changes in the likelihood of infant mortality for both grandmothers and infants over time, to ensure that these do not drive the effects of the number of deaths of the grandmother on infant survival.

Our main results are in line with our earlier findings. A grandmother's infant deaths are significantly related to the survival chances of her grandchildren. In the second model, where we add control variables, the effect of infant deaths of the grandmother decreases slightly, but remains significant and sizeable (HR 1.12 and 1.34 for one and two or more infant deaths, respectively). The likelihood that an infant will die increases by birth order, while infants from average-aged mothers (25–35, reference category) are better off than those of younger mothers (15–24). Infants from relatively old mothers (35–50), however, do not significantly differ from those from average-aged mothers in their mortality risk. Since we have controlled for the age group of the mother at the birth of the index child and the birth order of the child, we find that the age of the mother at her first birth affects the likelihood that her infants will die: children born to a first-time mother who is below the age of 20 are 19% more likely to die during their first year of life. This result is in line with theories of fast life histories and our expectations, but the strength of the effect is remarkable.

Table 2 The effects of the grandmother's infant deaths on infant survival, Zeeland, 1833–1912

	Cox model		Weibull model, mother shared frailty		Weibull model, grandmother shared frailty	
	Hazard rate	*p*-value	Hazard rate	*p*-value	Hazard rate	*p*-value
Infant deaths of the grandmother						
Zero	*ref.*		*ref.*		*ref.*	
One death	**1.12**	0.00	**1.12**	0.00	**1.13**	0.00
Two or more deaths	**1.34**	0.00	**1.34**	0.00	**1.35**	0.00
Births of the grandmother						
Two births	*ref.*		*ref.*		*ref.*	
Three births	0.93	0.06	0.93	0.12	0.93	0.10
Four to six births	**0.89**	0.00	**0.90**	0.00	**0.89**	0.00
Seven or more births	**0.84**	0.00	**0.84**	0.00	**0.84**	0.00
Sex						
Female	*ref.*		*ref.*		*ref.*	
Male	**1.25**	0.00	**1.26**	0.00	**1.26**	0.00
Birth order						
First	*ref.*		*ref.*		*ref.*	
Second	**1.11**	0.00	**1.12**	0.00	**1.12**	0.00
Third	**1.17**	0.00	**1.17**	0.00	**1.17**	0.00
Fourth to sixth	**1.30**	0.00	**1.30**	0.00	**1.30**	0.00
Seventh or higher	**1.67**	0.00	**1.60**	0.00	**1.62**	0.00
Child birth year, centered	**0.99**	0.00	**0.99**	0.00	**0.99**	0.00
Age of the mother						
25–34	*ref.*		*ref.*		*ref.*	
15–24	**1.09**	0.00	**1.09**	0.00	**1.09**	0.00
35–50	0.99	0.37	1.02	0.20	1.01	0.43
Intercept			0.17		0.17	
Frailty variance			0.43		0.32	
Number of infants	203,802		203,802		203,802	
Number of infant deaths	33,763		33,763		33,763	

Source: LINKS dataset Zeeland_2017_01 (IDS version).

In the third model, we add the socioeconomic status of the parents and grandparents. The effect of the number of grandmother infant deaths remains significant. We find that infants from the highest socioeconomic group have a better survival prospect than their peers from lower socioeconomic strata. In addition, infants whose mothers originate from highest-status households have better survival chances. Thus the effect of socioeconomic status can be tracked not only to the family of the infant itself, but also to the socioeconomic background of the mother. In addition, we find that infants whose fathers are farmers are better off than infants whose fathers have nonfarming occupations. This effect cannot be found for farming occupations of grandparents.

Furthermore, in Model 3, we include the survival of parents, as young children are highly dependent in their first year of life. We find that the death of the mother during the first year of life is related to higher chances of infant death (HR: 1.49). On the other hand, we find that paternal death is not related to a higher chance of early-life mortality of infants; rather, the death of the father is related to lower mortality chances of infants (HR: 0.77). This result is in line with earlier research showing that maternal death is more detrimental for infant survival than paternal death (Beekink, Van Poppel, & Liefbroer, 1999; Breschi & Manfredini, 2002) and that the death of parents can be compensated for, for instance, by the presence of other kin (Kok, Vandezande, & Mandemakers, 2011), remarriage (Andersson, Hogberg, & Åkerman, 1996), and out-migration (Breschi & Manfredini, 2002). Finally, very little frailty variance between mothers is explained by including the socioeconomic status of fathers and maternal grandfathers and maternal and paternal survival in the model. In addition, the main effect of the number of infant deaths of grandmothers on the survival of grandchildren is only slightly reduced and remains significant. Taken together, these results suggest that these characteristics play a minor role in multigenerational effects in infant mortality.

4.4 Sensitivity checks

We conduct sensitivity checks for several assumptions with regard to data selection and variable construction. The results are available on request. The robustness checks include selections of infants whose grandmother survives until her 50th birthday, infants whose grandmother is observed until her 50th birthday or death, infants whose grandmothers' husband survives until her 50th birthday, and the exclusion of infants with unknown dates of death or no observations after their birth from the analysis. No substantial differences were observed between our baseline models and the models with alternative case selections.

Table 3 The effect of the grandmother's infant deaths on infant survival, extended models, Zeeland, 1833–1912: Cox models with mother shared frailty

	Model 1		Model 2		Model 3	
	Hazard rate	***p*-value**	**Hazard rate**	***p*-value**	**Hazard rate**	***p*-value**
Infant deaths of the grandmother						
Zero infant deaths	*ref.*	*ref.*	*ref.*	*ref.*	*ref.*	*ref.*
One infant death	**1.12**	0.00	**1.12**	0.00	**1.11**	0.00
Two or more infant deaths	**1.34**	0.00	**1.33**	0.00	**1.31**	0.00
Births of the grandmother						
Two births	*ref.*	*ref.*	*ref.*	*ref.*	*ref.*	*ref.*
Three births	0.92	0.06	0.93	0.09	0.92	0.08
Four to six births	**0.88**	0.00	**0.89**	0.00	**0.89**	0.00
Seven or more births	**0.83**	0.00	**0.84**	0.00	**0.84**	0.00
Sex						
Female	*ref.*		*ref.*			
Male	**1.25**	0.00	**1.25**	0.00	**1.26**	0.00
Birth order						
First	*ref.*	*ref.*			*ref.*	*ref.*
Second	**1.11**	0.00			**1.12**	0.00
Third	**1.16**	0.00			**1.16**	0.00
Fourth to sixth	**1.27**	0.00			**1.27**	0.00
Seventh or higher	**1.54**	0.00			**1.55**	0.00
Age of the mother						
25–34	*ref.*	*ref.*			*ref.*	*ref.*
15–24	**1.09**	0.00			**1.09**	0.00
35–50	1.01	0.40			1.01	0.57
Mother early first birth	**1.19**	0.00			1.19	0.00
SES of the father						
Low					*ref.*	*ref.*
Low to average					**1.01**	0.00
Average					0.96	0.26
Average to high					0.95	0.10
High					**0.82**	0.00
Father is a farmer					**0.80**	0.00

Table 3 Continued

	Model 1		Model 2		Model 3	
	Hazard rate	*p*-value	Hazard rate	*p*-value	Hazard rate	*p*-value
SES of the grandfather						
Low					*ref.*	*ref.*
Low to average					0.99	0.01
Average					0.98	0.41
Average to high					0.93	0.02
High					**0.85**	0.00
Grandfather is a farmer					0.97	0.09
Survival of the father					**0.77**	0.00
Survival of the mother					**1.49**	0.00
Frailty variance	0.36		0.35		0.34	
Number of infants	203,802		203,802		203,802	
Number of infant deaths	33,763		33,763		33,763	

Source: LINKS dataset Zeeland_2017_01 (IDS version).
Note: The models are stratified by the year of birth of the child.

Furthermore, we divide the analysis into three different time periods: 1833–74, 1875–94, and 1895–1912. Children born within each time range are analyzed separately. We find no substantial differences between these analyses and our extended models. In addition, we include the lowest socioeconomic status known in the data instead of the highest observed socioeconomic status. Here, we find some differences with the main results. With the lowest socioeconomic status included, we find that each paternal socioeconomic status group has lower infant mortality than the last. For the socioeconomic status of the grandparent or the farmer status of the father or grandfather, we do not find differences in the results. Our substantive conclusions do not change, however: a higher socioeconomic status of the father reduces the likelihood that infants will die; but this effect does not explain the relation between infant deaths of the grandmother and the risk of death of the infant. In addition, it does not explain the frailty variance, or the differences between mothers in their chances of infant mortality.

Finally, instead of using a HISCAM scale of socioeconomic status, we include socioeconomic status using HISCLASS to investigate whether effects of socioeconomic status are class specific beyond the difference between farmers and nonfarmers. Here, we find that infants of higher professionals, lower clerical and sales personnel, and farmers and fishermen are relatively well off in comparison to elite groups, whereas children of lower-skilled farm workers have increased odds of dying in infancy. As for the social class of the maternal grandfather, we find that grandchildren of lower-skilled workers and unskilled and lower-skilled farm workers have a significantly higher risk of dying in infancy. However, these effects do not explain the relation between grandmothers' infant deaths and the risk of death of grandchildren. It appears that—comparable to the main results that we presented—the main dichotomy exists between farmers and nonfarmers, and higher classes versus lower classes. Thus, our substantive conclusions remain unaltered.

5. Discussion and conclusion

In this paper, we have analyzed the intergenerational transmission of infant mortality risk between maternal grandmothers and mothers, using a large dataset from Zeeland, the Netherlands. The aim of this paper was threefold: first, to determine whether there existed intergenerational transmission of mortality risk in Zeeland; second, whether this intergenerational transmission can be explained by the socioeconomic characteristics of the family, health of parents, and fast life histories; and third, to assess the advantages and potential disadvantages of conducting comparative research using the Intermediate Data Structure (IDS).

5.1 Intergenerational transmission of mortality risk

In our analysis, we have focused on intergenerational continuities between grandmothers and mothers in mortality among their infants. In line with earlier research, we have found that the risk of infant deaths of grandmothers and mothers is correlated. This was shown in previous research on 19th century Antwerp (Vandezande, 2012) and Skellefteå (Lindkvist & Broström, 2006). Analyses were conducted using Cox and Weibull survival models with shared frailty at the level of the mother or grandmother. All analyses led to similar conclusions, as did the sensitivity analyses: the risk of infant mortality is correlated between generations. Especially for grandmothers experiencing two or more deaths among their infants, the risk of infant mortality of their grandchildren is higher. We found that these results hold after controlling for the number of births of the grandmother, the birth order of the child, and several demographic characteristics of their families. Thus, we showed that there is a familial component in early-life mortality in 19th and early 20th century Zeeland. With regard to the mechanism explaining the relation between mothers and daughters in their experiences with infant mortality, we used Cox regression analyses to model the intergenerational correlation in infant mortality between grandmothers and mothers, incorporating several explanatory characteristics. We incorporated the socioeconomic

status of the maternal grandfathers and fathers, including whether or not they were farmers, the survival of parents during the first year of life of infants, and the early onset of reproductive careers.

In the current work, we did not find evidence that survival of parents plays an important role in explaining intergenerational transmission of high infant mortality. We expected that survival of mothers could play an important mediating role, as mothers exposed to high mortality in early life may be unhealthier in adulthood (Crimmins & Finch, 2006) and maternal survival is important for the survival of infants and children. Although the death of mothers has a strong negative impact on the likelihood that their infants will survive their first year of life, we found no evidence that parental survival explains the intergenerational relations between grandmothers and mothers in infant mortality.

We did find some preliminary evidence that fast life histories may play a role, as the demographic characteristics of the family—including the birth order of the child and the age of the mother at the time of her first birth and at the time of birth of the infant—were found to partially explain the relation between mortality in the family of the mother and among her own children. At the same time, the effects appear to be relatively minor: after including demographic characteristics, the effect of infant mortality of the grandmother remains strong and significant, and only a small share of the frailty variance between mothers is explained. Yet, lower investment in children, having more children, and marrying earlier may all contribute to high-mortality patterns among women originating from high-mortality environments (Nettle et al., 2011; Störmer & Lummaa, 2014). In addition, sociocultural factors could play a role: women from disadvantaged environments, characterized by early fertility and high infant mortality, may be socialized to pursue disadvantageous fertility careers, leading to high mortality in several generations (see, e.g., Laslett, 1980). In future research, more attention should be paid to the effects of mortality in the sibling set on mortality in the second generation, which may be mediated by the marriage age and birth intervals of the mother.

Socioeconomic status affects the mortality of infants—with lower infant mortality among farmers and in the highest-status groups. Not only was the socioeconomic status of the mother found to affect her infant's survival, but the socioeconomic status of the grandfather mattered as well. However, socioeconomic status played only a small role in explaining mortality clustering among mothers. This result contrasts with work on late 20th century northeast Brazil, which has illustrated that in some regions socioeconomic status plays an important role in explaining concentration of early-life mortality in households (Sastry, 1997). For textile cities in the Netherlands in the 19th century, however, no evidence was found that mothers from the working classes experienced higher mortality among their infants than mothers from other classes (Janssens & Pelzer, 2012). In addition, earlier historical work showed that the higher classes did not necessarily do better than the lower classes with regard to infant mortality, which may be explained by the use of wet nurses (Lalou, 1997). This contrasts with our findings, as we find a lower risk of infant mortality among the highest classes in Zeeland.

Preferences for boys over girls may lead to higher mortality among girls in several generations, at least partially due to neglect of girls. In the current research, such a preference would be indicated by intergenerational patterns of high infant mortality at the expense of the survival of female infants. However, in contrast with earlier work (Devos, 2000; Van Poppel et al., 2009), we did not find evidence for family patterns in sex-specific infant mortality, as boys and girls were similarly affected by high mortality among their grandmothers' infants.

In conclusion, we have emphasized the role of demographic characteristics in explaining infant mortality clustering in families, showing that there exists an intergenerational component to mortality clustering. However, questions about the causes of these intergenerational continuities remain. In future research, a further attempt should be made to disentangle the causes of these intergenerational continuities. While we have shown that the socioeconomic status and demographic characteristics of the family cannot explain these similarities, other possible explanations include genetic factors and intergenerational stability in living conditions. Families who have lived in the same place for multiple generations may be subject to similar living conditions over time, resulting in intergenerational continuities in mortality among infants. On the other hand, if they migrate out of their region of origin, they may be subject to different mortality regimes, resulting in intergenerational change in mortality among their children. Furthermore, in the current analysis we were unable to include infants born in families who had out-migrated from Zeeland. Possibly, out-migrants' infant mortality differs from stayers' mortality patterns, leading to possible under- or overestimations of intergenerational transmission of mortality clustering. Finally, future research should attempt to incorporate the causes and timing of death, to assess whether intergenerational continuities extend to these characteristics as well.

5.2 The use of IDS for a common analysis

This paper is part of a project in which several research groups have used a common script for selection of cases, construction of variables, and analysis, based on a common data format known as the Intermediate Data Structure (IDS). The IDS format provides a way to store datasets based on different sources into a similar structure,

with the purpose of facilitating comparative research in historical demography. However, not all issues of comparative research can be solved by storing data in a similar format and constructing variables in similar ways. Each historical demographical database has been constructed from different sources, with different purposes, and with varying requirements for the inclusion of persons. Although the IDS structure ascertains that database structures are identical across datasets, which guarantees that the program for extraction of variables and analysis will run with every dataset, different characteristics of underlying databases may lead to problems in case selection, variable construction, and analysis. The design of a common program for analysis is a complicated process, as approaches to data selection and variable construction that are relevant in one database may be less relevant in another, and necessary requirements specific to another database may be overlooked, possibly leading to problems downstream.

In the current work, the analyses were based on three-generation information, including grandmothers, mothers, and children. For a precise selection of mothers and grandmothers with complete life-course information, ideally, data with continuous observations from the date of birth or in-migration until the date of death or out-migration, such as data from population and parish registers, are used. With such data, decisions on sample selection can be based on the length of observation of individuals and their relatives—in this case, concerning children, their mothers, and their grandmothers. However, for other types of datasets, including family reconstitutions based on vital event registration such as LINKS, no continuous observations are available and we cannot observe individuals and their relatives in a specific time window, but we rely on more fragmented observations. As case selection based on the observation window is not possible for such data, it is more difficult to select cases for data analysis.

A second problem arises when datasets contain persons who are related to others and satisfy the conditions for selection, but at the same time do not contain meaningful observations. For instance, in the LINKS database, the names of parents are given on birth, marriage, and death certificates. Thus, the database identifies these parents, and contains a record for them, but the case only contains relational information and a name. Therefore, to select a case in LINKS, it is insufficient to observe a linked grandmother: we have to know more about her than just her name. In several other commonly used historical demographical datasets, similar problems could occur; for instance, because of a sample design to which incomplete life-course data of significant others has been added, or because in multigenerational pedigrees, ancestors are identified but not followed over time.

Thus, in using the IDS format and common scripts for comparative research, it is not always easy to decide which cases should be selected for analysis. To make the right selection for an analysis such as the one conducted in this paper, one has to ascertain that relevant characteristics are known for selected persons. To ease this problem, it would be useful if IDS-structured databases were to include a characteristic distinguishing relatively complete cases—key actors—from cases that do not contain meaningful observations. This indication may be based on the source of information about this person—the way in which the existence of this person has been verified—and the available data for this person in the database. Second, scripts could be developed that add this information to the IDS dataset, the so-called extended IDS (Quaranta, 2016). Third, extraction software for comparative research on IDS-structured databases should be developed with care to ensure that the specific characteristics of each database are accounted for.

Whereas the use of a common script for data construction is not without problems, the advantages of using a common approach are also manifold. Within the current project, intergenerational transmission of infant mortality between grandmothers and mothers has been approached in a comparable way between several historical populations. This has allowed new insights into the extent to which family histories of high mortality are relevant for the survival of infants, providing valuable new insights into the determinants of infant survival in the past.

Paper IV

Repeating a Lethal Experience: Mother's and Father's Exposure to Sibling Mortality and Offspring Mortality

(submitted)

Author
Van Dijk, I.K.

Earlier versions of this paper were presented at the European Population Conference, Brussels, Belgium in June 2018, at the Posthumus Conference, Utrecht, the Netherlands in May 2018, and at the seminar for the Radboud Group for Historical Demography in April 2018. The paper has benefited from the comments of Rebecca Sear, Sören Edvinsson, Ken Smith, Isabelle Devos, Kees Mandemakers, Eline Slagboom, Marjan Beekman, Mar Rodriguez-Girondo, Niels van den Berg, and Rick Mourits.

Abstract

This paper shows that risk of both infant and child mortality is transmitted across generations. Offspring of mothers and fathers exposed to sibling mortality face an increased risk of early child death (0–5), with stronger maternal effects. The effects are additive but do not interact. It is investigated how early-life mortality is linked together across generations, paying specific attention to two groups of mechanisms: mechanisms that connect direct exposure to mortality in the sibling group and offspring mortality, and family-shared characteristics that are related to child mortality. Analyses are based on full-population family reconstitution data from the UPDB (Utah, US, 1874–1929) and LINKS (Zeeland, the Netherlands, 1833–1917). Two approaches are used, first, comparing all families using clustered Cox survival models and, second, controlling for unobserved heterogeneity between families using fixed-effects models. The intergenerationally high risk of child death can partially be attributed to demographic and socioeconomic characteristics of the family, especially young age at first reproduction and a short interval between births among parents exposed to sibling mortality. Both exposure to sibling mortality and family-shared factors play a role in linking child mortality between generations. It is concluded that adverse reproductive trajectories and childcare behavior are likely to link together high child mortality across generations.

Keywords Early-life effects, infant mortality, child mortality, biodemography, death clustering, intergenerational transmission, life-course research

1. Introduction

Families play a pivotal role in survival of young children, and in some families the risk of death is considerably greater than in others. Child mortality in historical and developing-world populations clusters in high-risk families, which experience more deaths than could be expected given their size and the mortality rates in their communities (Das Gupta, 1990; Edvinsson & Janssens, 2012; van Dijk, 2018). Family histories of early-life mortality play a role in this phenomenon, as it has been shown that infant mortality of grandmothers and mothers are related between generations in several historical populations in Europe (Lindkvist & Broström, 2006; Vandezande, 2012; Quaranta et al., 2017; Broström et al. 2018; Donrovich et al., 2018; Quaranta, 2018; Sommerseth, 2018; Van Dijk & Mandemakers, 2018). It is likely that characteristics between families that are shared between generations play an important role here (Vandezande et al. 2010), including influences from the family of origin on family structure (Liefbroer & Elzinga, 2012), age at marriage (Van Poppel et al., 2008), family size (Murphy, 2013), and socioeconomic status (Knigge, 2016). However, the exact mechanisms remain unclear.

Disentanglement of the mechanisms underpinning this intergenerational stability contributes to our understanding of child mortality in historical and developing-world populations and the change in levels of child mortality over time as well as within populations. This paper innovates in comparison to the earlier literature by using an interdisciplinary approach and individual-level longitudinal data on several consecutive generations of family members to explore the mechanisms that link together high child mortality across generations, in the paternal as well as the maternal line. The results, discussed in more detail below, show that infant and child mortality are not stand-alone phenomena, but that family histories of mortality play an important role in child survival, through effects that work via direct exposure as well as through disadvantageous conditions in child and adulthood.

Broadly, there are two groups of mechanisms that play a role in multigenerational continuity in high infant and child mortality: family-shared characteristics and exposure to sibling death in childhood. Intergenerational continuities in offspring mortality may originate from either or both of them. First, family-shared characteristics that are related to infant and child mortality levels—including socioeconomic status, demographic characteristics such as fertility, and social and behavioral characteristics—may be (socially) inheritable (Anderton et al., 1987; Lyngstad & Prskawetz, 2010; Wolfinger, 2011; Carvalho, 2012), as well as genes. If these characteristics are transmitted from parents to their children, it is likely that the children will experience similar levels of infant and child mortality upon reproduction.

Second, parental exposure to sibling mortality may affect young-age mortality among their offspring through affecting exposed individuals' behavior and preferences. For instance, work from evolutionary biology on life-history theory suggests that family exposure to sibling mortality affects surviving sibling's age at first childbirth, marriage, and length of birth intervals (Störmer & Lummaa, 2014); demographic studies found long-lasting negative effects of exposure to sibling mortality on health and survival, which may be related to stress responses (Yu et al., 2017) and scarring of health (Alter et al., 2001). In turn, earlier age at marriage, short birth intervals, or adverse health in adulthood may negatively affect the survival of infants and young children, linking together parental exposure to high mortality in childhood and offspring mortality. While the first group of explanations predicts an intergenerational relation between all deaths occurring in the sibling group and mortality in the second generation, the second view on intergenerational relations in infant and child mortality suggests that it is parental exposure to sibling mortality that is particularly relevant in predicting mortality among offspring.

In order to unpack the mechanisms linking high mortality in multiple generations, it is essential to include fathers in addition to mothers. Commonly, in most research on the role of parents in infant and child mortality, most attention is paid to the role and characteristics of the mother, including characteristics such as her age at childbirth (Hobcraft, McDonald, & Rutstein, 1985), education (Das Gupta, 1990), occupation (Janssens & Pelzer, 2012), and previous child-caring experience (Donrovich et al., 2018). The inclusion of the father has often been limited to socioeconomic aspects such as his occupation and income (Schellekens & Van Poppel, 2012). A notable exception is work by Van Bodegom et al. (2012), who have used polygynous unions to show that fathers as well as mothers play a key role in shaping early-life conditions of children and significantly affect their chances of survival. In many historical populations, on which the current work is based, primary caretaking tasks were the domain of the mother, whereas most males were primarily breadwinners. The inclusion of mortality in the family of origin of the father in addition to that of the mother may therefore help disentangle the causes of intergenerational high mortality. Furthermore, it is assessed whether mother and father effects are additive, and matter for offspring survival independently from a given mortality level in the sibling group of the other parent; whether they interact with each other, exacerbating existing inequalities; or whether paternal exposure is irrelevant given the maternal mortality experience, or vice versa.

The major innovation of this paper is disentanglement of the mechanisms that link together early-life mortality across generations in both the maternal and the paternal line through family-shared effects and effects of exposure to high-mortality environments in childhood. Furthermore, unobserved heterogeneity is controlled for by conducting within-family comparisons of the effect of parental exposure to sibling death on offspring mortality, comparing exposed siblings to nonexposed siblings, and assessing the effects of exposure on their offspring's survival through infancy and early childhood (birth to age 5). In addition, the role of all sibling mortality in the sibling groups of fathers and mothers is investigated by comparing offspring mortality between parents originating from families experiencing various levels of child mortality. The father's and mother's childhood exposure to sibling mortality and their age of exposure to sibling mortality are explicitly addressed, to pinpoint crucial windows of exposure. The results presented in this paper increase our understanding of crucial intergenerational links in mortality that have a major impact on health and survival of children and the factors that give rise to successful reproduction and raising of children in historical populations.

2. The theoretical background

This paper offers an exploration of the causes of intergenerational continuities in high mortality in families. Explanations for intergenerational continuity in early-life mortality between grandparents and parents may be sought in individual effects of mortality exposure in childhood on the one hand and intergenerational continuities within families in social, demographic, and biodemographic characteristics that are related to high infant and child mortality on the other. In the first group, the emphasis lies on the effects of parental exposure to sibling mortality on health, survival, and aspects of their reproductive careers in adulthood, which in turn affect the health and survival of their own children. Second, through processes of socialization and social reproduction, behavioral characteristics and preferences, including factors affecting childcare behavior, and socioeconomic status, including poverty, are transmitted between generations. These characteristics are related to the likelihood of infant and child mortality and survival, and therefore, intergenerational continuities in these characteristics may lead to increased levels of child mortality in several generations. Examples include preferences for breastfeeding versus artificial feeding and other aspects of child-feeding behavior, which are related between generations through processes of socialization within families and other social groups, including religious groups. An overview of the theories can be found in Box 1.

Three mechanisms may link together child mortality in multiple generations through causal effects related to mortality exposure, identifying a direct relationship between mortality exposure and effects on the life course, including the health and reproductive health of the exposed individual and the reproductive career. Because of these changes, eventually, the survival chances of the offspring of affected individuals are affected negatively.

First, the exposure of individuals to hostile conditions—such as disease epidemics and sibling mortality—may result in cumulative damage to the health of individuals and reduced life expectancy (Barker et al., 2002; Lindeboom, Portrait, & Van den Berg, 2010; Doblhammer, van den Berg, & Lumey, 2013). Childhood experiences with sibling deaths may inflict damage on individuals through exposure to infectious disease (Van Dijk, Janssens, & Smith, 2018), or may indicate that food and caretaking habits in the parental family were insufficient and therefore resulted in damage to individuals' bodies (Norton et al., 2017), and are related to increased morbidity and mortality in adulthood (Yu et al., 2017). Surviving individuals from high-mortality families who have children as adults may struggle with health issues, caretaking responsibilities, labor, and other duties, and are more likely to die when their children are young, resulting in adversity for their own offspring. It should be noted that due to out-selection of relatively frail individuals from the sibling group, it is also possible that surviving siblings will be relatively more robust and have higher survival rates in adulthood (Alter et al., 2001).

Second, for women, damaged reproductive health related to adversity in childhood may lead to repeated stillbirths and increased perinatal mortality among women's offspring (Lalou, 1997). Indeed, healthier mothers produce healthier offspring (Reid, 2001), and thus the nutritional status and living standards in the family of origin of the mother, as well as exposure to disease and death, may affect the survival chances of children.

Third, life-history theory suggests that there are differences between individuals in the timing of life-course transitions, including the start of reproduction and the length of birth intervals, which are affected by the extrinsic mortality risk. Individuals originating from high-mortality environments are more likely to pursue a fast life track that is related to a younger age at the start of reproduction, and shorter birth intervals, with the purpose of increasing the number of children even in difficult conditions, which may lead to an early death for individuals (Störmer & Lummaa, 2014). The earlier onset of reproduction and shorter intervals lead to closely spaced large sibling sets, resulting in fewer resources available per individual child and overcrowding in the household, which increases the likelihood of infant and child mortality among the offspring of affected individuals. Fast life histories may therefore link relatively high infant and child mortality in several consecutive generations. Although the sociobiological and demographic literature provides numerous studies confirming the existence of fast life mechanisms working between high mortality and other early-life adversity and a diverse set of outcomes including age at first menarche, childbirth, and birth interval (Chisholm et al., 2005; Störmer & Lummaa, 2014), many ambiguities still exist with regard to the pathways of influence. The literature does not agree on the timing and nature of the

Box 1 An overview of the hypotheses on exposure to mortality in the second and third generations

Mechanism	Link between first and second generations	Mechanism	Mother (M) or father (F) linked?	Timing related?	Other indicators
Cumulative damage	Damaged health of parents results in decreased caretaking and other capacities	Response to mortality exposure	M as main caretaker		
Damaged reproductive health	F as main provider	No	Early death of parents		
Life-history theory	Damaged reproductive health results in problems with pregnancies and childbearing	Response to mortality exposure	M	Yes: puerperal period	
Socioeconomic inequalities	Fast reproduction; short birth intervals and lower investment in children lead to higher mortality	Response to mortality exposure, possibly indirectly through grandparents	M as main caretaker	No	Parental age at first birth; length of birth intervals
Behavioral and caretaking capacities or habits	Low SES related to child mortality and transmitted between generations	Socially inheritable characteristics	F as main provider	No	
Genetic deficits	Caretaking shortfalls and habits cause high mortality in multiple generations	Socially inheritable characteristics	M as main caretaker	No	

cues that individuals use to assess whether they are living in a high-mortality environment. In the literature, it is unclear whether the link between mortality in the childhood environment and later fertility behavior in their own family is created through the parents, or whether individuals take their cues directly from their childhood environment. Earlier work has suggested that individuals assess risk levels from family-level mortality rather than from individual exposure to child mortality (Störmer & Lummaa, 2014), suggesting that the influence of sibling mortality on offspring mortality does not necessarily work through direct exposure but may work through the parents of affected sibships.

In the case of direct individual exposure within families, several time windows may be relevant. It is possible that a critical period exists, so that exposure before or after that period is less relevant or, alternatively, it may be that regardless of the exact timing, all exposure in childhood leads to cumulative negative experiences. Here, exposure to sibling mortality is measured for several potentially crucial time frames—infancy and early childhood (0–3), middle childhood (3–6), and late childhood (6–12)— as well as all deaths during childhood (0–12). The *in utero* period (from nine months before birth until birth) is tested separately.

Not all intergenerational similarities may indicate family-shared or causal effects, as grandparents and parents may live and reproduce in communities with similar characteristics, and therefore have similar mortality outcomes in their families. In communities with very high incidences of infant mortality and low out-migration, it is less surprising that consecutive generations lose considerable numbers of children; in addition, in developing-world and historical populations there are considerable differences between communities in the level of infant and child mortality. Earlier research has suggested that after taking community-level effects into account, almost all variance between families in the number of child deaths disappears (Sastry, 1997), possibly because in high-mortality environments almost all variance is explained by income inequalities (Guo, 1993), although the evidence is mixed (Saha & van Soest, 2011; Van Poppel et al., 2012). Thus, the community could play an important role in the establishment of intergenerational relations in child mortality, that are not driven by family characteristics but, rather, by the tendency of multigenerational families to live close together. To ensure that our findings are not driven by correlations due to shared communities, within-family comparisons are used, which make it possible to take a closer look at the role of exposure in the intergenerational relation in child mortality.

In addition to effects related to exposure to sibling mortality, shared characteristics in families may also affect the intergenerational link in offspring mortality. Here, it is not the event of mortality or exposure to mortality that affects offspring survival, but characteristics that are shared within families and that affect multiple generations in a similar fashion. Learned caretaking behavior, socioeconomic characteristics, and demographic characteristics may be shared within multiple generations of the same family, as well as genes. These characteristics are not related to direct exposure to sibling mortality, so that all deaths of parental siblings—independent from timing and including those occurring before this parent's birth—may be related to increased likelihood of death for the parents´ own offspring. Furthermore, it is likely that social characteristics related to mortality are especially important in the post-neonatal period. It is generally assumed that mortality due to caretaking behaviors and other external characteristics occurs in later phases of infancy than mortality related to internal characteristics and childbearing, which are supposed to be especially important in the neonatal period and for stillbirths (Van Dijk, 2018).

In historical Europe, intergenerational relations in child mortality are likely to be sex specific due to gender differences in dominant roles in the household and differences in the reproductive and immune systems. The male role in caretaking in historical populations is smaller than for women, as it was predominantly considered a task for mothers and other females in the household. Indeed, the death of fathers barely affects the survival of young children (Sear & Mace, 2008), while the absence of mothers does have a detrimental influence on survival in early life (Kok, Vandezande, & Mandemakers, 2011). Other authors have suggested that fathers do play a supportive role in the caretaking of children (Lovejoy, 1981), although the role of the supporting adult may also be played by paternal and maternal grandmothers (Hawkes et al., 1998). At the same time, Vandezande (2013) found that child mortality clustering was related across generations in the paternal line, and concluded that in historical populations fathers played a more important role in caring for their children than previously assumed. However, in his work, the roles of the paternal and maternal families of origin were not considered simultaneously, and thus it remains unclear which parental background plays a more important role in intergenerational relations in high mortality among children and whether mortality histories from each side contribute independently to the offspring's survival.

In addition to caretaking characteristics, socioeconomic characteristics are passed on between generations in both the maternal and the paternal line (Vandezande & Matthijs, 2013). Poverty and socioeconomic deprivation may constitute an alternative explanation for intergenerational relations in child mortality clustering. In historical Europe there were socioeconomic (Bengtsson & Van Poppel, 2011) and occupational

gradients in infant and child mortality (Lalou, 1997) and child mortality clustering (Janssens & Pelzer, 2012; Scalone et al., 2017), which increased during the 19th century (Bengtsson & Van Poppel, 2011) and still exist today. In a process of social reproduction, socioeconomic status and occupational characteristics are transmitted between generations (Zijdeman, 2009). This may lead to intergenerational consistencies in mortality levels as childcare behaviors, household living arrangements, and other social characteristics related to infant and child mortality are shared within social and occupational groups (Das Gupta, 1997; Edvinsson & Janssens, 2012; Scalone et al., 2017). In particular, the father is expected to matter here, as in our period of research—the late 19th century—in most households men were the main providers. Thus, poverty and deprivation in the father's family of origin as well as in his own family could be an important cause of intergenerational mortality continuities in the paternal line.

In this paper, it is assessed whether an intergenerational relation exists in mortality among the father's and mother's siblings and mortality among their children before age of 5. The paper innovates on the earlier literature through its focus on mechanisms linking high mortality across generations, inclusion of the role of fathers, exposure to mortality, and effects on children beyond infancy (between the first and the fifth birthday). An individual-level longitudinal approach is used, based on several consecutive generations of family members. A distinction is made between mechanisms involved in intergenerational transmission of high mortality by assessing the exposure and the timing of exposure using family fixed-effects models. To further establish the role of the family in deaths after infancy and to distinguish between factors related to childbirth and pregnancy and externally affected causes of young-age mortality, effects are estimated for all children below age 5 as well as for infants.

3. The data

The analyses employ two unique, large-scale historical demographic datasets arising from two widely differing environments: the UPDB (Utah, US) and LINKS (Zeeland, the Netherlands). Both data sources contain large-scale demographic information about historical populations with multigenerational pedigrees, which allow us to follow individuals and their siblings, parents, and children over their life course. For Utah, birth cohorts from 1874–1929 are included, following children's survival until 1934, and for the Netherlands birth cohorts from 1833–1912, following children's survival until 1917.

Utah, a western state in the United States, was first settled by immigrants from North America and Europe after 1847. Initially, many migrants were associated with The Church of Jesus Christ of Latter-day Saints (i.e., the Mormon church), but over time the share of non-Mormons and inactive Mormons grew. The largely Protestant province of Zeeland is situated in the southwestern corner of the Netherlands and consists of islands that were relatively isolated at the time. Both populations were high-fertility populations during the period of this study, with a mean number of children around eight (see Table 1). Despite harsh conditions in the early years of settlement, the population in Utah was relatively healthy in comparison to that of Zeeland. Potable water was available through runoff from the mountains. In Zeeland, a lack of potable water existed due to salinity. In Zeeland, adult life expectancy was low and infant and child mortality high, reaching 50% in some years and communities (Hoogerhuis, 2003; Van Poppel & Mandemakers, 2002).

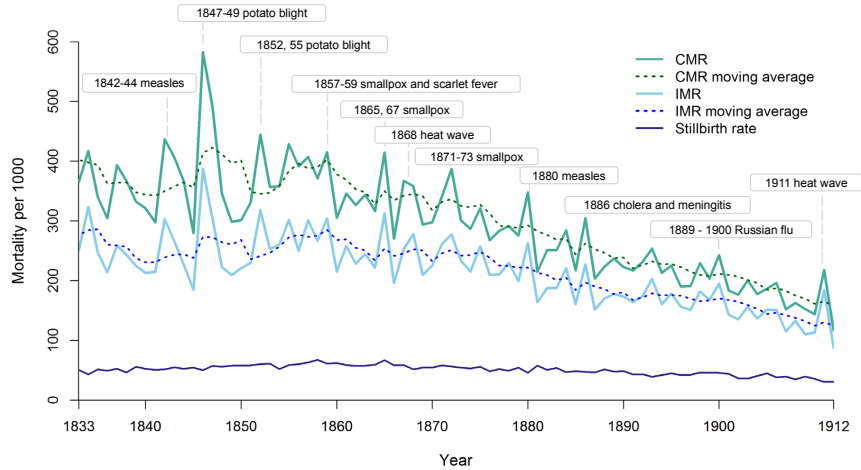
For both populations, high-quality family reconstitution databases exist: the Utah Population Database (UPDB) and LINKS (LINKing System for historical family reconstruction) Zeeland. Both databases contain large numbers of individuals linked into multigenerational pedigrees with demographic and mortality information. The UPDB has been used to assess the effect of social, biological, and genetic factors on longevity and life spans (Kerber et al., 2001; Garibotti et al., 2006; Smith et al., 2009; Van den Berg et al., 2019). LINKS has been established in demographic research on the Netherlands (see, e.g., Van Poppel, Jonker, & Mandemakers, 2005; Van Poppel et al., 2008; Bras, Van Bavel, & Mandemakers, 2013; Knigge, 2016) and has been subject to thorough checks of its quality and reliability (Van den Berg et al., 2018).

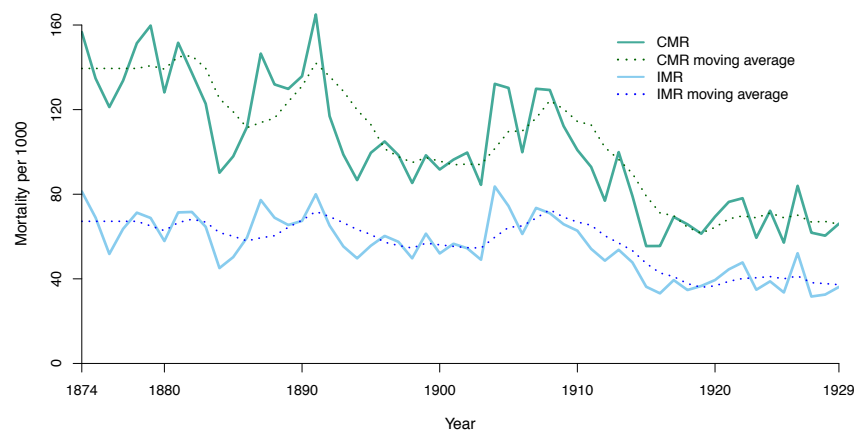
The UPDB contains demographic and genealogical information linked to medical records and other data sources. The UPDB was initially based on Utah founders and their descendants, employing genealogical records from the archives of the Utah Family History Library. These records contain demographic and mortality information on the pioneers of Utah, their parents, and their children, and have been linked into intergenerational pedigrees with information from other official administrative and medical sources, including birth and death certificates and US Census records. The full database includes over 11 million individuals. Both individuals with and without an affiliation with The Church of Jesus Christ of Latter-day Saints (LDS, or Mormons) are included in the UPDB. For our analyses, individuals are selected who were born in Utah between 1874 and 1929, for whom information is available about the families of origin of their parents and about their own survival. Furthermore, eligible individuals are followed until death, out-migration, loss to follow-up, or survival to age 5. There are 385,870 eligible individuals in the UPDB (see Table 1), who are called egos and individuals interchangeably.

The LINKS data contain demographic, genealogical, and mortality information based on linked indexes of digitized civil certificates. These originate from obligatory vital event registration, which was introduced in the Netherlands in 1812. The certificates include information on the names of born, married, and deceased persons, the municipality of the event, the individuals' place of birth, age, occupational titles, and the names of parents. Through the linkage of certificates pertaining to the same individuals using first and last names of ego and his or her parents, the life courses of spouses and children have been reconstructed. Earlier research has suggested that the family reconstructions have been quite successful and that families in the database are representative for the place and time, with the exception of issues related to the out-migration of healthy migrants (Van den Berg et al., 2018). Database construction has been completed for Zeeland and is ongoing for other provinces of the Netherlands. The full database contains information on 1,930,157 individuals from Zeeland who experienced a vital event in that province between 1812 and 1912 for births, 1812 and 1927 for marriages, and 1812 and 1957 for deaths. For the analysis, 215,364 persons are selected, who were born between 1833 and 1912, whose parents and grandparents are known, and whose parents have at least one sibling. For a full overview of selection criteria and the number of remaining cases, see Table 1 in the Appendix.

For the state of Utah, individuals are followed until their last observation date. For Zeeland, it is assumed that individuals without follow-up survived early childhood until age 5. For a small minority of the cases, a death certificate may not have been indexed or linked although the child died before the age of five. In the analyses, these cases will weaken the associations between our variables of interest. Furthermore, children who were lifeless upon registration are not included in the analyses. These cases encompass both stillbirths and births of infants who died before they were reported to the municipality, which had to be completed within three working days of the birth (Vulsma, 1988). Estimates suggest that 71% of lifeless reported infants concern stillbirths and 29% concern live births of infants who died before municipal registration of birth (Van Poppel et al., 2012). As it is unclear which children were born alive and for how long they lived after birth, lifeless registered children are excluded from the analysis, which will lead to a slight underestimation of infant deaths in the neonatal period and to conservative estimates of the intergenerational relation in child mortality risk.

Figure 1 The child mortality rate (CMR), the infant mortality rate (IMR), and the stillbirth rate: (a) for the province of Zeeland, 1833–1912; (b) for the State of Utah, 1874–1929





Sources: LINKS Zeeland 2017_01 (Mandemakers & Laan, 2017), UPDB.
Note: Stillbirths are not separately registered for Utah. Moving averages represent the rolling average mortality rates over five-year periods. For peak mortality years, the following literature has been consulted: Hofstee (1974), De Meere (1982), Hoogerhuis (2003), Ekamper, Duin, Van Poppel, and Mandemakers (2010), and Krantenbank Zeeland [Newspaper Archive Zeeland] for the relevant years.

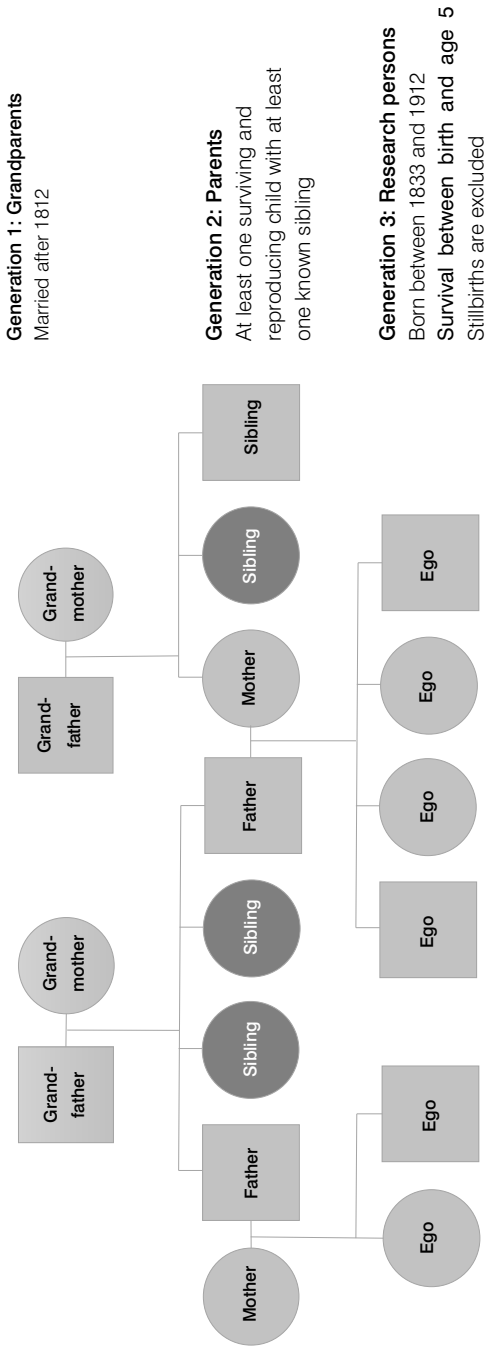


Figure 2 The design of the study

Table 1 The descriptive statistics

	Zeeland			Utah		
	Min.	Max.	Mean	Min.	Max.	Mean
Child variables						
Death before fifth birthday	0	1	0.24	0	1	0.07
Age at last observation	0	5	4.01	0	5	4.75
Short birth interval (<18 months)	0	1	0.32	0	1	0.09
Sex: female	0	1	0.49	0	1	0.49
Birth order	1	24	4.43	1	20	4.23
Twin birth	0	1	0.03	0	1	0.02
Total number of children	**215,364**			**385,879**		
Mother variables						
Family size	2	25	8.25	2	19	8.3
Sibling deaths, not exposed	0	16	0.94	0	15	0.52
Sibling deaths, *in utero*	0	4	0.06	0	4	0.03
Sibling deaths, 0–3 year old	0	6	0.55	0	6	0.11
Sibling deaths, 3–6 year old	0	7	0.51	0	5	0.18
Sibling deaths, 6–12 year old	0	7	0.36	0	7	0.32
Has first child young (<21)	0	1	0.14	0	1	0.50
Age at birth of child	14.4	45.0	30.5	12.5	45.0	28.9
Age at birth 14–20	0	1	0.03	0	1	0.11
Age at birth 21–29	0	1	0.47	0	1	0.49
Age at birth 30–34	0	1	0.26	0	1	0.20
Age at birth 35–49	0	1	0.24	0	1	0.20
Death before end of observation of child	0	1	0.03	0	1	0.02
SES of family of origin	0	9.9	5.02	0	10	2.75
Family of origin in farming	0	1	0.21	0	1	0.42
Inactive Mormon	–	–	–	0	1	0.14
Active Mormon	–	–	–	0	1	0.76
Total number of mothers	**42,361**			**84,424**		

Table 1 Continued

	Zeeland			Utah		
	Min.	Max.	Mean	Min.	Max.	Mean
Father variables						
Family size	2	25	8.22	2	19	8.2
Sibling deaths, not exposed	0	14	0.91	0	12	0.51
Sibling deaths, *in utero*	0	4	0.05	0	3	0.03
Sibling deaths, 0–3 year old	0	6	0.54	0	4	0.22
Sibling deaths, 3–6 year old	0	7	0.50	0	7	0.18
Sibling deaths, 6–12 year old	0	8	0.38	0	7	0.33
Has first child young (<21)	0	1	0.09	0	1	0.19
Death before end of observation of child	0	1	0.02	0	1	0.02
SES of family of origin	0	9.9	5.04	0	1	2.78
Family of origin in farming	0	1	20.9	0	1	0.42
SES adulthood	0	9.9	5.16	0	10	2.98
Is a farmer	0	1	0.19	0	1	0.39
Inactive Mormon	–	–	–	0	1	0.11
Active Mormon	–	–	–	0	1	0.82
Total number of fathers	**41,475**			**81,670**		

The design of the study can be found in Figure 2. All parents' sibling deaths before the age of 12 are measured, which is done separately for the sibling set of the mother and the father. The windows of exposure are accounted for by measuring separately exposure to sibling deaths between the mother´s and father´s birth and their third birthday, and between the parent´s third to sixth birthdays (early childhood) and their sixth and 12th birthdays (late childhood). In addition, deaths that occur when the mother and father are *in utero* are included. Stillbirths and deaths before registration are excluded, for reasons discussed above. In the analyses, the number of siblings of the father and mother who are alive at the beginning of the four age intervals is controlled for, as the likelihood that they will experience sibling mortality in childhood is related to the number of siblings who are alive.

The descriptive statistics can be found in Table 1 and a graphical representation of the occurrence of parental sibling deaths by the age of the parent can be found in Figure 3. Due to high childhood mortality in the province of Zeeland (see Figure 1), it is relatively common for individuals to experience sibling mortality during childhood, especially between ages one and six (see Figure 3, right).

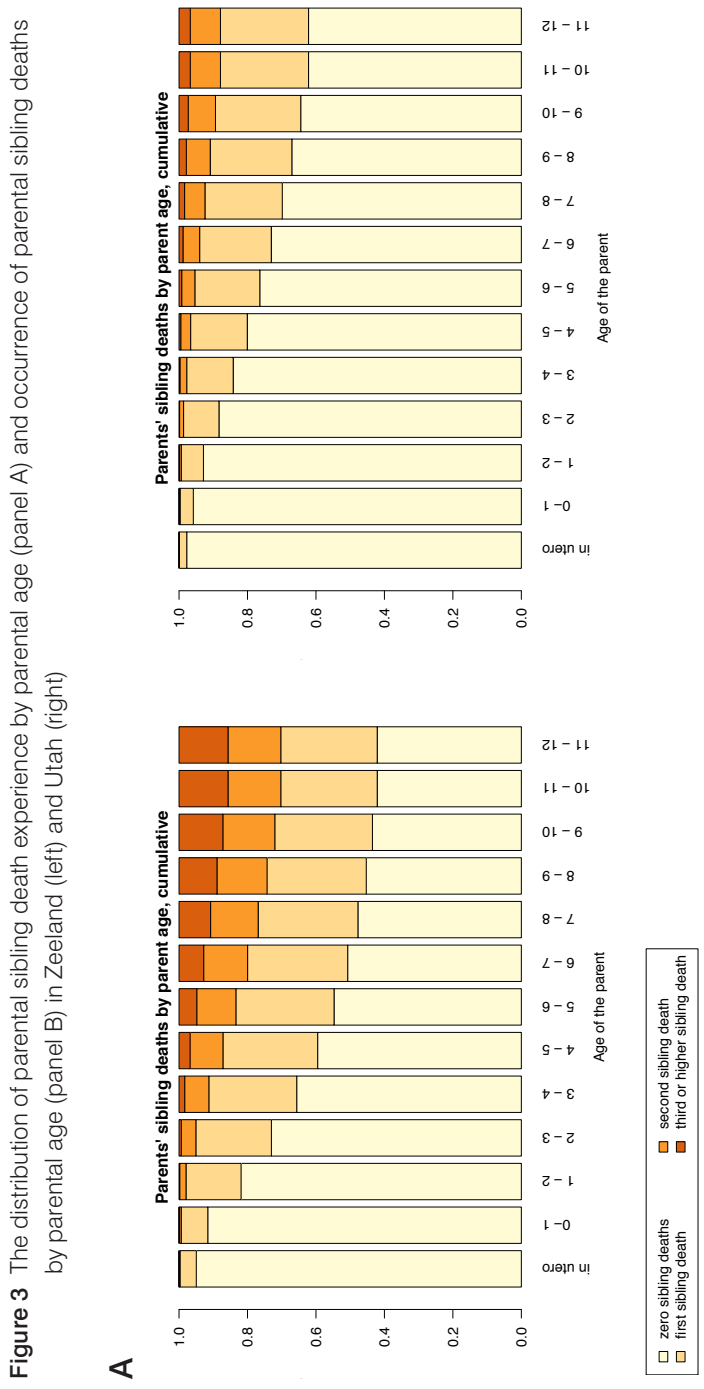
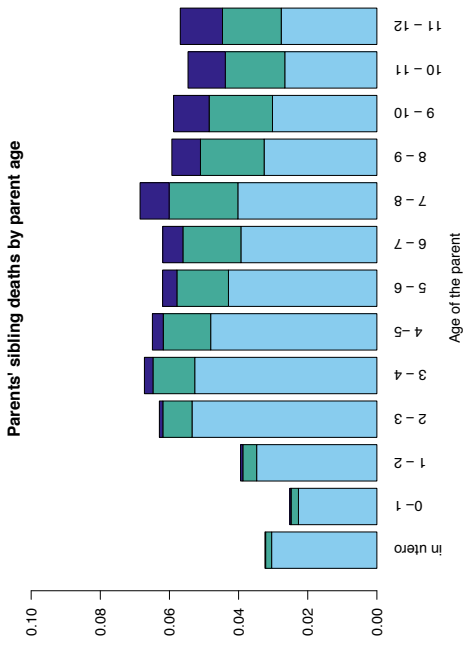


Figure 3 The distribution of parental sibling death experience by parental age (panel A) and occurrence of parental sibling deaths by parental age (panel B) in Zeeland (left) and Utah (right)



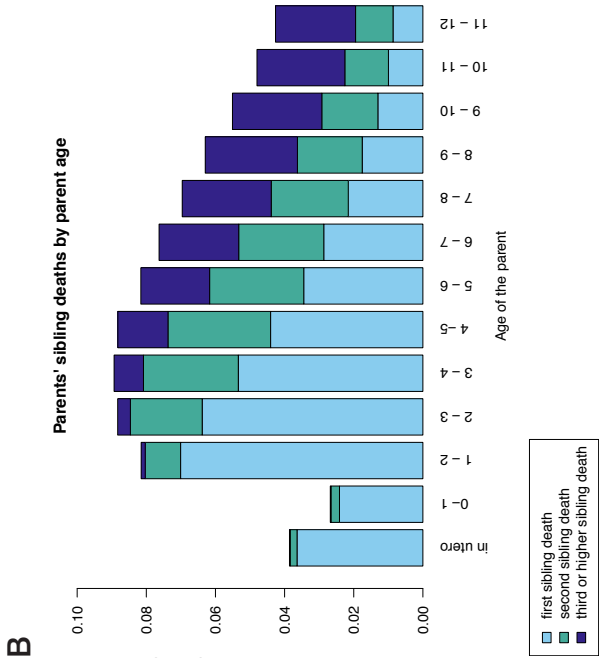


Figure 3 Continued

Several characteristics of the family of origin and the childhood circumstances of the parents are included. The descriptive statistics can be found in Table 1. The demographic characteristics of the children include the number of full siblings, the birth order, the age of the mother at birth, and twin births. For fathers and mothers separately, a distinction is made between first reproduction at a relatively young age (<21 years old) and reproduction at later ages. For Utah, an indicator of religious status is included. Active Mormons generally engage in lifestyle characteristics that are beneficial for survival, such as alcohol and tobacco avoidance, fasting, and participating in the social aspects of the religion. Inactive LDS members are baptized in the LDS church, generally at age 8, and share an early-life environment that may be beneficial for survival, but do not express their commitment to the LDS church in adulthood. Non-LDS members have no record of involvement with the LDS church. No individual-level indicator for denomination is available for predominantly Protestant Zeeland.

Other characteristics of interest concern whether or not the father or mother has their first child before the age of 21; whether the birth interval to the previous child was short (<18 months); the death of the father and mother (time-varying); and the highest socioeconomic status of the father, the maternal grandfather, and the paternal grandfather. Given the large fraction of the Utah and Zeeland population engaged in agriculture, farmers are distinguished from nonfarmers. For Utah, the Nam-Power SES score (1950 version) is used (Nam & Boyd, 2004), based on the highest-scoring occupation observed in the censuses and, if unavailable, the SES from the death certificate. For Zeeland, the highest HISCAM SES (Lambert et al., 2013) available from any linked vital event certificate is used to measure socioeconomic status, in conjunction with HISCLASS (Van Leeuwen & Maas, 2011), to distinguish farmers from nonfarmers. All codes were provided by the HSN standardized, HISCO-coded and classified occupational titles, release 2018.01 (Mandemakers et al., 2018). Socio-economic status was measured using the highest known occupational title of the father and the grandfather. In addition, a measurement for the farmer status of the father and the grandfather is included, which is based on the HISCLASS scheme code 8 (Van Leeuwen & Maas, 2011). For persons with a missing SES score, the mean SES score is imputed, as well as a dummy indicating the imputation. For this period, males were the primary breadwinners in both Utah and Zeeland, resulting in under-registration of female labor participation, with few recorded occupations for women. Therefore, for women, the socioeconomic status of the spouse is used.

Time-varying dummies indicate whether a parent dies in the first years of life of the child. Young children are highly dependent on family members for survival in their early life, and the death of a parent can therefore detrimentally affect the chances of survival of infants and young children. The date of death of the parent is used to assess whether

a parent died in the first five years of life. If no date of death is known, the last date of observation of the parent is used. For some children, observations on mothers or fathers end before the fifth birthday, sometimes with the birth of the child as the last observation. In these cases, it is assumed that the parents survived until the fifth birthday or—in the event that the child died before the fifth birthday—until the death of the child.

Clustered Cox proportional hazard models of survival between birth and the fifth birthday are estimated and observations are censored at age 5. The main independent variables are sibling deaths of the mother and the father. For Zeeland, 215,364 children are included in the analysis, who are the descendants of more than 40,000 mothers and fathers and more than 26,000 grandmothers. The analyses are clustered on the maternal and paternal grandmothers to estimate standard errors that are robust for multiple observations for children from the same household. In a next step, to estimate the effects of exposure within families, controlling for unobserved heterogeneity, fixed-effects models are used. These models are based on comparisons between siblings (fathers and mothers) originating from one family, who vary in timing and extent of exposure to sibling deaths but who share family characteristics, including the socioeconomic status, the age of the mother at first birth, and characteristics of the environment such as access to safe drinking water. Thus, these models allow for a distinction between effects as the result of sibling-shared characteristics and effects related to characteristics that are different between siblings, including exposure to sibling deaths. It should be noted that the mechanisms that link together sibling mortality exposure and offspring mortality indirectly—for instance, through trauma and behavioral change of the parents of deceased siblings—will be underestimated using models based on sibling comparisons. These models will only bring to light differences within the sibling set, and not the factors that affect all siblings simultaneously. Furthermore, to investigate whether intergenerational continuities exist beyond infancy, effects are estimated separately for infants between birth and the first birthday and young children between birth and age 5.

4. Results

In a first step, the effects of parental exposure to sibling mortality on survival are estimated. Clustered Cox regression models are employed, which account for the fact that the models are based on three-generation family trees, in which a single pair of grandparents can have multiple children—who shared similar exposures to sibling mortality—and multiple grandchildren in the analyses. The models are clustered at the level of the maternal and paternal grandparents, respectively, and estimated separately for mother and father effects and exposures in different periods of life. Parents may be exposed to sibling mortality in different stages of life: after conception but before birth—that is, *in utero*—between birth and age 3, age 3 and age 6 (early childhood), and between ages 6 and 12 (late childhood). The sex of the child and the parents' sibling group size at the beginning of the exposure period are controlled for. All models are stratified on the birth year to account for secular changes in infant and child mortality levels over time. The results are presented in Table 2 and Figure 4. In Figure 4, 95% confidence intervals are plotted for the mother and father effects, showing that in Zeeland mother effects tend to be stronger than father effects. The effects of exposure between birth and age three, strong exposure between ages 3 and 6, and strong exposure between age 6 and age 12 are significantly stronger for mothers than for fathers. For Utah, similar differences are found in the point estimates of the effects, but the differences between mothers and fathers are not significant. Furthermore, we find that effects are additive but not multiplicative. Both fathers' and mothers' sibling deaths contribute to higher child mortality simultaneously as well as separately, but do not interact: the effects add up, but neither exacerbate nor compensate each other in both Zeeland and Utah (results not shown).

For the Zeeland population, significant effects of parental exposure to sibling mortality in all stages of childhood (*in utero* up to age 12) are found, for both exposure of the mother and of the father. Regardless of the age at which parents have been exposed, exposure appears to be related to increased mortality hazards for their own offspring between birth and age 5. For Utah, the picture is more mixed. No significant effects are found for sibling deaths before birth. There exists an increased offspring death risk for maternal exposure between age 0 and 12, but not consistently. For fathers, effects are found especially for later-childhood exposures (age 6–12), but these are relatively weak in comparison to the Zeeland results.

In Figure 4, 95% confidence intervals are plotted for mother and father effects, showing that in Zeeland mother effects tend to be stronger than father effects. The effects of exposure between birth and age 3, strong exposure between ages 3 and 6, and strong exposure between age 6 and age 12 are significantly stronger for mothers

Table 2 The clustered model: the effects of mortality in the sibling group and the relative death risk of offspring, children 0–5

		Zeeland				Utah			
Maternal sibling deaths		**Model 1**		**Model 2**		**Model 1**		**Model 2**	
		RR	**CI**	**RR**	**CI**	**RR**	**CI**	**RR**	**CI**
Parent *in utero*									
	0 deaths	ref.		ref.		ref.		ref.	
	1 or more deaths	1.10***	1.06–1.15	1.07***	1.03–1.11	1.07	0.99–1.16	1.06	0.99–1.15
Parent 0-3									
	0 deaths	ref		ref		ref		ref	
	1 death	1.15***	1.12-1.17	1.13***	1.10-1.15	1.08***	1.04-1.12	1.07***	1.03-1.11
	2 or more deaths	1.29***	1.23-1.35	1.23***	1.18-1.29	1.20***	1.08-1.35	1.19***	1.06-1.32
Parent 3-6									
	0 deaths	ref		ref		ref		ref	
	1 death	1.10***	1.08-1.13	1.08***	1.06-1.11	1.04**	1.01-1.08	1.04	1.00-1.07
	2 or more deaths	1.27***	1.23-1.32	1.23***	1.18-1.27	1.01	0.93-1.10	0.99	0.92-1.09
Parent 6-12									
	0 deaths	ref		ref		ref		ref	
	1 death	1.06***	1.04-1.09	1.05***	1.02-1.07	1.07***	1.04-1.10	1.06***	1.03-1.09
	2 deaths	1.19***	1.15-1.23	1.15***	1.11-1.19	1.10***	1.04-1.16	1.08***	1.02-1.14
	3 or more deaths	1.41***	1.35-1.47	1.33***	1.28-1.39	1.08	0.98-1.20	1.06	0.96-1.18
Paternal sibling deaths		**Model 1**		**Model 2**		**Model 1**		**Model 2**	
		RR	**CI**	**RR**	**CI**	**RR**	**CI**	**RR**	**CI**
Parent in utero									
	0 deaths	ref		ref		ref		ref	
	1 or more deaths	1.07***	1.01-–1.10	1.05**	1.01–1.10	0.99	0.92–1.07	0.98	0.91–1.06
Parent 0–3									
	0 deaths	ref.		ref.		ref.		ref.	
	1 death	1.06***	1.02–1.07	1.04***	1.02–1.07	1.04*	1.00–1.09	1.04	1.00–1.08
	2 or more deaths	1.15***	1.07–1.17	1.12***	1.07–1.17	1.09	0.97–1.22	1.07	0.96–1.20
Parent 3–6									
	0 deaths	ref.		ref.		ref.		ref.	
	1 death	1.08***	1.03–1.07	1.05***	1.03–1.07	0.98	0.94–1.01	0.97	0.94–1.01
	2 or more deaths	1.20***	1.11–1.19	1.15***	1.11–1.20	1.02	0.94–1.11	1.00	0.93–1.09
Parent 6–12									
	0 deaths	ref.		ref.		ref.		ref.	
	1 death	1.07***	1.03–1.08	1.06***	1.03–1.06	1.03*	1.00–1.07	1.03	1.00–1.06
	2 deaths	1.15***	1.06–1.13	1.10***	1.06–1.13	1.07*	1.01–1.13	1.05	1.00–1.11
	3 or more deaths	1.24***	1.11–1.22	1.17***	1.11–1.22	1.13**	1.03–1.25	1.12*	1.02–1.24
Events		51,081				26,569			
N		215,364				385,879			

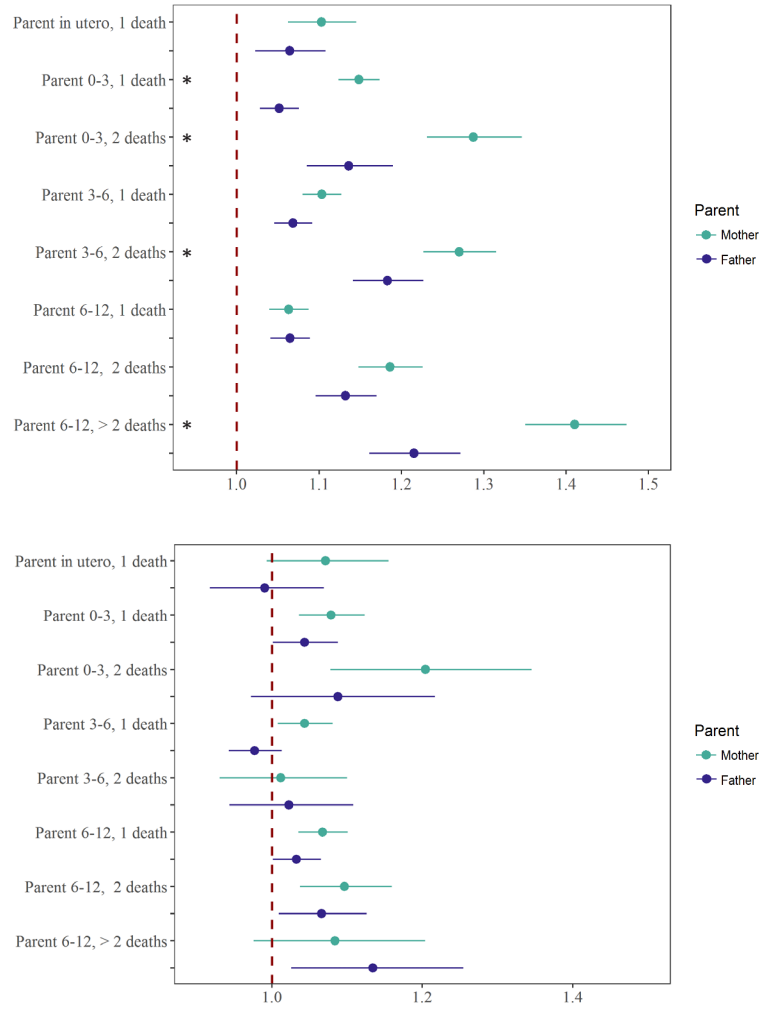
Notes: All models are controlled for parental family size in the exposure period and sex of the child, and stratified on the birth year of the child. Model 2 includes the birth order of the child; whether or not the previous birth interval was short (<18 months); whether the parent was young at birth of first child (<21 years old); the SES and farmer status of the parents and the grandparents; and the death of the parent in the first five years of the child's life (dynamic). The models use robust standard errors correcting for the inclusion of multiple grandchildren from the same maternal and paternal grandparents.
Statistical significance: ***, $p < 0.01$; **, $p < 0.05$; *, $p < 0.10$. CI refers to a 95% confidence interval.

than for fathers. For Utah, similar differences are found in the point estimates of the effects, but the differences between mothers and fathers are not significant. Furthermore, it is found that effects are additive but not multiplicative. Both fathers' and mothers' sibling deaths contribute to higher child mortality simultaneously as well as separately, but do not interact: the effects add up, but neither exacerbate nor compensate each other in both Zeeland and Utah (results not shown).

In Model 2, control variables are added. Effect sizes are reduced by the inclusion of the control variables, but the main effects can still be observed (effects of control variables not shown). In line with the literature, a higher birth order, short birth intervals, and death of the parent in the early years of the child all contribute to a higher likelihood of early child death. Children whose mother had her first child when she was relatively young (<21) have an increased risk of early death as well, even after controlling for the effect of her age at the birth of this child. In other words, regardless of her age at the birth of a child, if she started reproduction young, her children have an increased risk of death before the age of 5. These characteristics partially explain the relation between exposure to sibling mortality and offspring survival. Additional analyses (not shown here) reveal that the birth order of the child, a young age of the mother at first reproduction, and a short birth interval before the birth of the child in particular explain the relation between the exposure to sibling deaths of the parents and the child's likelihood of survival. This result applies to both the effect of paternal and maternal sibling deaths, and indicates that total family sizes, disease transmission, and resource competition in large families may play a role in increased child mortality among offspring of parents from high-mortality families. However, the effects of parental exposure to sibling mortality remain significant after including the explanatory characteristics, except for fathers exposed at young ages (0–3) in Utah.

In a next step, exposure is considered as a determinants of these differences. Alternatively, the effects may be caused by unobserved family characteristics, such as socioeconomic status, behavioral characteristics, or other factors that are shared within families. To control for unobserved family characteristics, within-family comparisons using fixed-effects models are conducted. Offspring of parents who have and have not been exposed to sibling mortality are compared. Because of differential ages and birth orders, multiple siblings from the same family may have a strikingly different experience with regard to exposure to death and disease in the household, while other family characteristics are shared (see Appendix, Figure 1). Models are estimated that are fixed on maternal grandmothers and paternal grandmothers, and that estimate the effect of exposure to sibling mortality of mothers and fathers, respectively, on the survival of offspring. Furthermore, the models control for the family size of the parent at the beginning of the time interval, the sex of the child, and the birth year.

Figure 4 The effect sizes of paternal and maternal exposure to sibling deaths on offspring survival by age of exposure of the parent and number of experienced deaths, Zeeland (upper panel) and Utah (lower panel)



* Effect differs significantly between mothers and fathers.

Children whose parents have same-sex, reproducing siblings are selected, in order to estimate the effects of maternal and paternal childhood exposure to sibling mortality. The survival of children whose mothers or fathers were exposed to sibling mortality is compared that of children of parents without this exposure. Children without maternal aunts or without paternal uncles in the dataset are therefore excluded from the current analysis. For Zeeland, 136,121 children are included for the analysis of father effects, and 134,442 children for the analysis of mother effects, who were born to about 25,000 pairs of fathers and mothers and more than 10,000 pairs of grandparents. For Utah, 304,147 children are included for father effects and 299,832 children for mother effects, who were born to about 60,000 pairs of fathers and mothers and more than 20,000 pairs of grandparents.

The results are shown in Table 3. The effects should be interpreted as follows. The relative risk represents the risk of mortality of offspring of exposed mothers and fathers in comparison to the risk of unexposed siblings of their parents; that is, the aunts (mother models) and uncles (father models). A visualization of the interpretation can be found in Figure 5.

For maternal exposure, it is found that exposure between birth and age 3 and between age 6 and 12 in particular has an effect on the survival of offspring. No significant effects of *in utero* exposure to sibling mortality are found. For paternal exposure, some effects of exposure on offspring mortality are found, but less consistently than for women. Strong exposure—more than two deaths—between age 3 and 6 or ages 6 and 12 is related to increased offspring mortality. Taken together, some support is found for the hypothesis that exposure to sibling mortality of both mothers and fathers is related to increased child mortality.

Furthermore, it is found that the effects that are described here are applicable to both infants under age 1 and young children between the ages of 1 and 5. In earlier work, the intergenerational transmission of young-child mortality was assessed for infant mortality only; here, it is confirmed that this relationship exists for infants as well as young children. The results are not driven by infant mortality, as effects of the parental sibling's mortality also exist for young children if infants are excluded. At the same time, the point estimates of the effect sizes are lower for children than for infants, and the effects all but disappear for paternal effects in Model 2, suggesting that, especially for father effects, family characteristics such as behavioral habits may play a determining role in intergenerational continuities for older children (results available on request).

Table 3 The fixed-effects model: parental exposure to sibling mortality and child mortality 0–5

	Maternal				**Paternal**			
	Zeeland		Utah		Zeeland		Utah	
Parent sibling deaths	RR	CI	RR	CI	RR	CI	RR	CI
Parent *in utero*								
0 deaths	*ref.*	*ref.*		*ref.*	*ref.*	*ref.*	*ref.*	*ref.*
1 or more deaths	0.99	0.93–0.97	1.01	0.91–1.13	1.00	0.94–0.97	0.99	0.88–1.10
Parent 0–3								
0 deaths	*ref.*	*ref.*		*ref.*	*ref.*	*ref.*	*ref.*	*ref.*
1 death	1.06***	1.02–1.11	1.07**	1.01–1.14	1.03	0.99–1.08	1.04	0.98–1.11
2 or more deaths	1.14***	1.04–1.25	1.10	0.93–1.30	1.06	0.96–1.16	1.05	0.88–1.25
Parent 3–6								
0 deaths	*ref.*	*ref.*		*ref.*	*ref.*	*ref.*	*ref.*	*ref.*
1 death	1.01	0.97–1.06	1.00	0.94–1.05	1.04*	1.00–1.09	0.97	0.92–1.03
2 or more deaths	1.07*	1.00–1.16	1.01	0.89–1.15	1.17***	1.08–1.27	1.00	0.88–1.13
Parent 6–12								
0 deaths	*ref.*	*ref.*			*ref.*	*ref.*	*ref.*	*ref.*
1 death	1.11***	1.06–1.17	1.06**	1.01–1.12	1.06**	1.01–1.12	1.03	0.98–1.08
2 deaths	1.18***	1.09–1.28	1.10**	1.01–1.21	1.07*	0.99–1.16	1.04	0.95–1.14
3 or more deaths	1.44***	1.28–1.63	1.09	0.91–1.29	1.18***	1.04–1.33	1.18*	1.00–1.39
Events	30,820			20,935	31,895			21,392
N	134,442			299,823 136,136				304,119

Notes: The table presents results from separate models for exposure in different age groups. The models are controlled for birth year of the child, sex of the child, and family size of the parent. Statistical significance: ***, $p < 0.01$; **, $p < 0.05$; *, $p < 0.10$. CI refers to a 95% confidence interval.

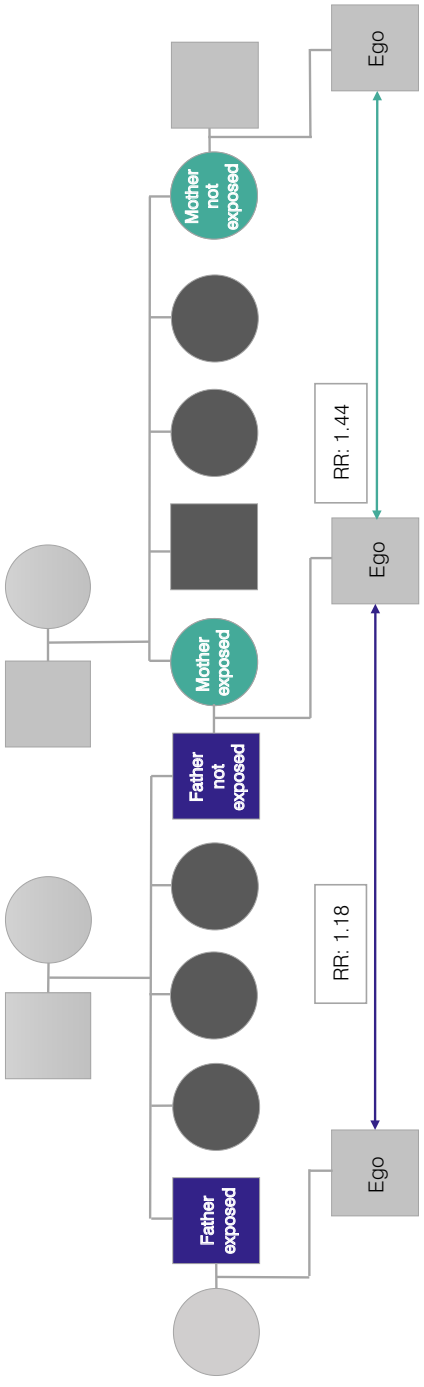


Figure 5 The effect of exposure to three or more sibling deaths between age 6 and age 12 and offspring survival between age 0 and age 5, Zeeland

Notes: Offspring of fathers who have been exposed to sibling mortality have an 18% higher risk of mortality between birth and age 5 than offspring of their uncles (brothers of the father) who have not been exposed to sibling deaths. In comparison to offspring of unexposed aunts, offspring of exposed mothers experience a 44% higher risk of mortality

Several robustness checks have been conducted. In the current analysis, birth year is controlled for, to ensure that our results are not driven by secular change in mortality risk over time. However, effects may still weaken over time due to contextual changes, leading to a lower mortlaity risk as the demographic transition progresses. Therefore, the effects have been analyzed separately for several time periods. For the Netherlands, analyses have been conducted separately for the period before the demographic transition (1833–70), during its onset (1870–90) (Wolleswinkel-Van den Bosch, 1998), and a period in which it was in full swing (1890–1917). Similarly, for Utah, the research period was divided into an early and late period (before and after 1900). The conclusions are not substantially different between these time periods. Second, to ensure that differences between the results by modeling strategy are not caused by a stricter selection of the included sample, clustered survival analysis is executed for all children who have a niece or nephew in the database (i.e., children who are included in the fixed-effects models). Although some effect sizes are slightly reduced, no substantial differences in the results for the full sample and the restricted sample are found (results available on request).

5. Discussion and conclusion

A high risk of infant and child mortality is transmitted between generations in both the paternal and the maternal line. This paper is a significant step forward in assessing intergenerational continuities in young-age survival and mortality, and in explaining how the survival and health of children are shaped within families. It is the first to show that besides mortality in the sibling set of the mother also mortality in the sibling set of the father is significantly related to offspring survival between birth and age 1, and between age 1 and age 5. Although earlier papers have looked at intergenerational transmission of infant mortality between generations, this is the first to explicitly address *exposure* to mortality, and to assess the explanatory mechanisms that link together high mortality across generations. A closer look at the mechanisms connecting mortality in several generations generates insights into processes of early-life exposures affecting health, morbidity, and mortality in the second generation. This has serious implications, not only for historical but also for modern-day populations. Here, despite rapid changes in health and mortality over recent decades, lasting effects of family histories of mortality may affect the health and survival of infants and small children up to the present day.

Earlier work on the role of the family in early-life mortality has shown that there exist intergenerational continuities in infant mortality between offspring of grandmothers and their surviving daughters (Lindkvist & Broström, 2006; Vandezande, 2012;

Quaranta et al., 2017; Broström et al., 2018; Donrovich et al., 2018; Quaranta, 2018; Sommerseth, 2018; Van Dijk & Mandemakers, 2018). In the current work, it has been shown that this relationship exists for both maternal and paternal grandmothers: infants born to surviving daughters and sons from a high-mortality family have a higher likelihood of dying before their first birthday. Furthermore, an increasing number of child deaths among the parents' siblings is related to increasing mortality chances for their offspring. Families in which mortality is increasingly high also show an increasing tendency to repeat that experience in the next generation.

Although the effects of paternal and maternal family history of mortality on infant and child survival are additive, they do not interact. In other words, mortality in the sibling set of the father and mother both matter for child survival, and do not compensate for each other. If the father was exposed to sibling mortality as a child, his offspring has increased mortality chances between birth and age 5, regardless of the mortality exposure of the mother. The paternal and maternal mortality exposure effects neither exacerbate nor ameliorate each other. It was expected that, given a mother from a high-mortality background, the mortality background of the father would possibly be less relevant, because of her major role in child-caring in historical populations. Indeed, exposure to sibling mortality for both fathers and mothers is related to increased mortality rates for their offspring, but at the same time, maternal effects are significantly stronger than paternal effects, which may be related to the highly gendered division of tasks within the household in 19th and early 20th century populations. The stronger association with the mother's family history of high mortality suggests that caretaking behaviors that are learned in families may be part of the explanation why levels of child mortality are related across generations. Furthermore, not only is infant mortality related within families, but effects can be detected both in infancy and in early childhood (up to age 5). This finding underlines the fact that familial relations in the likelihood of early-life mortality are not only determined by processes related to pregnancy and birth that affect highly vulnerable infants soon after birth, but also by processes that work later in childhood. These later periods are especially related to external causes of death, rather than to factors related to pregnancy and childbirth (Reid, 2002; van Dijk, 2018). Taken together, these findings indicate that there may be an important role for childcare habits and preferences in determining intergenerational processes of child mortality. This study is the first to show that there are intergenerational relations in early-life mortality beyond infancy. Nevertheless, the effects are stronger in infancy (before the first birthday) than in early childhood (between the first birthday and age 5).

The larger role of the maternal family history of mortality than the paternal family history of mortality may relate to the physiology of the reproductive system. Ova containing a women's egg cells that will ripen during her reproductive career are already in place in female fetuses *in utero*. Thus, *in utero* exposure to adversity and disease may therefore affect females' future offspring directly. In line with this explanation, it was found that for women but not for men, *in utero* exposure to sibling mortality matters for the survival of their children. In our analyses, sibling deaths occurring before conception were also included, and for these deaths it was found that there exists a relation to offspring mortality as well, which can be explained by indirect effects working through grandmothers, whose health and well-being may be affected by stress and infectious disease. For effects that work indirectly—via grandparents—a stronger effect in the maternal than the paternal line cannot be explained by the physiological mechanism explained above. However, here as well as for *in utero* exposure, the maternal side is more relevant than the paternal side, and this cannot be explained by the physiological mechanism described above.

In this work, a closer look has been taken at childhood mortality in the sibship and sibling mortality to which future parents were exposed, to control for unobserved heterogeneity and assess mechanisms related to exposure rather than all mortality occurring in the family. It has been found that a significant relation exists between exposure to sibling mortality and offspring mortality. This suggests that caretaking behaviors and other socially transmitted characteristics cannot be the sole explanations for intergenerational continuities in child mortality. Mechanisms that relate to mortality exposure itself are likely to play an important part in explaining these consistencies between generations. In line with life-history theory, birth order, a short birth interval preceding the birth, and a young age of the parents at first birth in particular—rather than at *this particular* birth—appear to be responsible for the relation between parental exposure to mortality and offspring mortality between birth and age 5. It should be noted that mechanisms that link together sibling mortality exposure and offspring mortality indirectly—for instance through trauma and behavioral change of the parents of deceased siblings—will be underestimated using models based on sibling comparisons. These models will only bring to light differences within the sibling set, and not the factors that affect all siblings simultaneously, as might be the case with explanations that focus on life histories.

Work in this field has shown that individuals exposed to adversity in early life, including absence of the father and extrinsic mortality risk, affect aspects of the life course, including age at menarche, early marriage and reproduction, and lower investment in individual offspring (Chisholm, 1993; Webster et al., 2014). Evidence for historical populations has shown that exposure to mortality in the sibship affects age at first

relationship formation, including marriage (Störmer & Lummaa, 2014). In this work, it has been established that in conjunction with these reproductive characteristics, the success of reproduction is also affected by the survival of children, as the likelihood of child survival is negatively affected by exposure of their parents to sibling mortality in childhood. However, more work needs to be done to further establish this relationship, paying attention to other characteristics that are considered part of this mechanism, including early onset of reproduction, short birth intervals in response to childhood adversity, and lower investments of parents in individual children. At the same time, the explanatory power of the suggested mechanism is weak. Furthermore, the results described here can alternatively be explained by deviating fertility and childhood mortality patterns in some subsets of society, including socioeconomic groups such as farmers and religious groups such as Catholic minorities, and in families affected by alcoholism, violence, and unhealthy living environments. In this study, at least part of the intergenerational effects in early-life mortality is related to all sibling deaths in sibships and not necessarily to those to which future parents are exposed, which indicates that family-level differences as well as direct exposure to sibling mortality both play a role in the intergenerational relation in early-life mortality. Furthermore, it appears that for offspring mortality, the window of exposure (the timing of exposure to sibling mortality in childhood) is less relevant than the total sibling mortality experienced in childhood.

Some critical remarks need to be made. Although many potential confounders of the relationship under scrutiny are controlled for and conservative methods and estimates were used, several alternative explanations may be at play. For instance, in sibling groups with relatively higher mortality, fewer siblings survive to adulthood, leading to a lower availability of kin to assist during reproductive years. As relatives other than mothers and fathers may significantly contribute to child well-being and survival (Sear et al., 2002; Sear & Mace, 2008), a smaller kin network may be negatively related to the survival of infants and small children. In addition, individuals may live in an environment similar to that of their family members and, thus, exposure to similar environmental conditions, including availability of nutrition and safe drinking water, may play a key role in intergenerational similarities in child mortality. At the same time, intergenerational relations in child mortality have been confirmed using within-family comparisons controlling for unobserved heterogeneity, indicating that there are other mechanisms besides environmental conditions at work increasing mortality at a young age in some families.

This study makes an important contribution to the field of child mortality and family relations in health and survival, by assessing the role of the family in infant and child mortality. Several methodologies were used to pinpoint the effect of family-level

sibling mortality and individual exposure to sibling mortality within families on offspring's survival. Mechanisms that possibly connect mortality in multiple generations were addressed, not only by taking apart mother and father effects and exposure- and nonexposure-related effects, but also by including observed heterogeneity through the inclusion of characteristics of families and individuals in our analyses. It has convincingly been shown that intergenerational correlations in child mortality exist for two historical populations in Utah (US) and Zeeland (the Netherlands). It is concluded that family characteristics and exposure to death and disease play an important role in family histories of high mortality, both in the maternal and the paternal line, and that early-life mortality in several generations of relatives plays a crucial in the survival of infants and young children.

Appendix

Table A.1 The selection of cases

Number of cases by selection criteria	Zeeland (LINKS) Number of cases	Utah (UPDB)
No criteria	1,050,202	1,526,054
Grandparents and parents are identified[1]	–790,204	–896,790
& parents have at least one known sibling	–12,572	–188,444
& removal of children death on registration[2]	–13,059	–17,083
& birth cohort[3]	–17,520	–34,270
& missing values on key variables and data cleaning	–1,483	–3,588
Total number of included cases	**215,364**	**385,879**

Notes:
[1] For Zeeland, the additional criterion is that grandparents are born after 1812, the start of the vital event registration system in the Netherlands.
[2] For Zeeland, this category includes stillbirths and infants who died before registration, which had to take place within three working days of the birth of the child.
[3] Zeeland, 1833–1912; Utah, 1874–1929.

Figure A.1 Variability within sibling groups in exposure to sibling deaths

Source: LINKS Zeeland.
Notes: Distribution of mother's sister's exposure to sibling deaths between ages 3 and 6, for mothers who experienced zero (left panel), one (middle panel), or two sibling deaths (right panel) between ages 3 and 6.

Paper V

The Long Harm of Childhood? Childhood Exposure to Mortality and the Subsequent Risk of Adult Mortality in Utah and the Netherlands

European Journal of Population (online first, 2018)

Authors
Van Dijk, I.K.
Janssens, A.A.P.O
Smith, K.R.

DOI 10.1007/s10680-018-9505-1

An earlier version of this paper has been presented at the Conference of the PAA, Chicago, US in April 2017 and at the International Population Conference, IUSSP, Cape Town, South Africa in October 2017. The paper has benefited from comments by Sören Edvinsson, Isabelle Devos, Jan Kok, Kees Mandemakers, Eline Slagboom, Niels van den Berg, and Rick Mourits.

Abstract

How do early-life conditions affect adult mortality? Research has yielded mixed evidence about the influence of infant and child mortality in birth cohorts on adult health and mortality. Studies rarely consider the specific role of mortality within the family. We estimated how individuals' exposure to mortality as a child is related to their adult mortality risk between ages 18 and 85 in two historical populations, Utah (US) 1874–2015 and Zeeland (the Netherlands) 1812–1957. We examined these associations for early community-level exposure to infant and early (before the sixth birthday) and late (before the 18th birthday) childhood mortality as well as exposure during these ages to sibling deaths. We find that exposure in childhood to community mortality and sibling deaths increases adult mortality rates. The effects of sibling mortality on adult all-cause mortality risk were stronger in Utah, where sibling deaths were less common in relation to Zeeland. Exposure to sibling death due to infection was related to the surviving siblings' risk of adult mortality due to cardiovascular disease (relative risk: 1.06) and metabolic disease (relative risk: 1.42), primarily diabetes mellitus, a result consistent with an inflammatory immune response mechanism. We conclude that early-life conditions and exposure to mortality in early life, especially within families of origin, contribute to adult mortality.

Keywords Human aging, infection, biodemography, epidemiology, early-life adversity, cause-specific mortality, exposure to disease, adult mortality

1. Introduction

The first birth cohorts that experienced decreasing mortality in childhood were also the first to experience an increase in life expectancy in adulthood (Crimmins & Finch, 2006), suggesting that health and survival in adulthood may be influenced by early-life exposure to disease and famine (Fridlizius, 1989; Bengtsson & Lindström, 2000; Barker et al., 2002; Finch & Crimmins, 2004). Support for this theory has been mixed, however. Research for rural Sweden showed that high infant and child mortality within birth cohorts is related to lower life expectancies of survivors of these cohorts (Bengtsson & Lindström, 2000; Costa, 2000; Quaranta, 2014), while in other cases no such relationship was found (Gagnon & Mazan, 2009; Cohen et al., 2010; Hayward et al., 2016). Evidence linking early-life exposure to infections to old-age mortality is often based on studies using (aggregate) cohort-level data. These studies could therefore represent overall shifts in mortality patterns in the population without demonstrating a causal link between early-life exposures and mortality patterns in adulthood. Moreover, as geographic and familial variation in health and survival exists, individuals born in the same period can experience heterogeneity in their exposure to mortality. Therefore, aggregate measures of exposure to infant and child mortality in large birth cohorts using excess mortality or crude death rates are likely to bias estimates relating exposure to mortality in early life to subsequent adult mortality. Exposure to mortality can be measured more precisely by using community-level measures, but especially by focusing on mortality among siblings. In this study, we focus on the relation between childhood exposure to mortality in communities and in families and subsequent adult mortality.

It is axiomatic that early-life conditions are shaped within families. Nonetheless, little consensus exists about the families' role in linking family deaths in childhood and subsequent adult mortality. Commonly, when research considers the role of families, it is to account for confounding factors (Öberg, 2015). For instance, Gagnon and Mazan (2009) controlled for family-level conditions in research regarding the relation between cohort mortality and old-age mortality and found that family-level characteristics have stronger effects than exposure to cohort-level mortality. However, as they measured mortality exposure through mortality in birth cohorts and not through more proximate exposures in families, exposure to mortality in childhood may, in fact, play a much larger role in adult mortality than previously thought. Other studies have looked at the influence of sibling mortality on the life course and found that sibling deaths affect life transitions such as the timing of marriage and fertility (Störmer & Lummaa, 2014). In addition, recent research has shown that in modern-day Scandinavia, childhood exposure to sibling mortality leads to an increased likelihood of death between birth and age 37 (Yu et al., 2017), suggesting that sibling deaths

have a lasting influence on health and mortality; however, it remains unknown whether survival is similarly affected in later adulthood.

With regard to the mechanisms linking early-life mortality exposure to later-life mortality, several mechanisms are relevant here: causal mechanisms, shared environments in childhood and adulthood, and mechanisms of scarring and selection. First, bacterial and viral infections are related to inflammatory immune responses and stunting in adulthood, indicators of adverse health (Finch & Crimmins, 2004). Sibling deaths indicate that individuals may have been directly exposed to these infections. Inflammatory immune responses and inadequate development of vital organs and the immune system may lead to reduced life expectancy (Barker et al., 2002; Lindeboom et al., 2010; Doblhammer et al., 2013) through increased risks of stroke, cardiovascular disease (Caruso et al., 2005), metabolic diseases (Crimmins & Finch, 2006), and other causes of death in adulthood (Vasto et al., 2007). In addition, sibling mortality may lead to chronically increased levels of stress (Yu et al., 2017), which, in turn, is a risk factor for diseases such as asthma, diabetes, gastrointestinal disorders, cancer, and myocardial infection (McEwen & Stellar, 1993; McEwen, 1998). Similarly, Norton et al. (2017) found that sibling deaths during childhood predict systemic inflammation in later life, linking stress responses to adult life expectancy. To gain further insights into the causal mechanisms linking sibling mortality to mortality hazards in adulthood, in this paper, we employ cause-of-death information and competing-risk models.

Second, individuals who experienced sibling deaths may originate from disadvantaged socioeconomic backgrounds related to poor living and nutritional circumstances, leading to both elevated infant and child mortality and a compromised start in life for survivors. When infants were not consistently breastfed, the use of contaminated water may have led to elevated mortality risks for infants and families and adverse health outcomes in adulthood. Access to nutrition is expected to affect only certain segments of the populations studied here, as the influence of nutrition on child survival decreased during the 19th century. Beyond the effects on health in childhood, childhood conditions are related to conditions in adulthood through processes of social reproduction. Life spans of individuals exposed to poor conditions in early life may therefore also be affected by the same adverse conditions in midlife, resulting in shortened life expectancy (Preston et al., 1998). Third, high cohort or family infant and child mortality may coexist with relatively favorable adult survival. Within high-mortality cohorts, frail individuals may be selected out, resulting in a relatively robust surviving cohort (Costa, 2000; Myrskylä, 2010; Doblhammer et al., 2013) or sibship (Alter et al., 2001). In addition, individuals from favorable childhood and adult environments may be spared (Preston et al., 1998). Apart from these selection effects, some exposures may lead to acquired immunity (Preston et al., 1998). Although some research has shown that scarring effects commonly dominate selection effects (Hatton, 2011), the extent to which this is true may vary by the strength of exposure to mortality in the population: Less lethal exposures may generate selection effects, while stronger exposures could have direct physiological effects on survivors, leading to scarring (Störmer & Willführ, 2010). In this paper, we take the degree of exposure into account in order to disentangle selection and scarring effects.

Deaths among siblings may have different effects within the sibling group, as individuals are at different ages when deaths occur (Störmer & Lummaa, 2014). There is no consensus in the literature about the age at which exposure to mortality may affect later health outcomes. Famines are related to shorter life expectancies for newborns at the time of the famine (Lindeboom et al., 2010) but also for children who are exposed *in utero* during mortality peaks (Barker et al., 2002; Bengtsson & Lindström, 2003; Doblhammer, 2003). Bruckner and Catalano (2009) found no effects of cohort infant mortality on adult mortality rates and hypothesize that the critical period may be at an older age. Other researchers hypothesized that exposure to mortality is significant up to age 5 (Fridlizius, 1989) or 15 (Crimmins & Finch, 2006). Alter et al. (2001) analyzed sibling mortality to assess the role of health behaviors in explaining excess mortality in early and later life. They found that sibling mortality affects mortality among other siblings during childhood but diminishes after age 15, with no effects after age 55, while demonstrating evidence consistent with acquired immunity after age 30. However, the authors used all sibling deaths in the family of origin, omitting the possibility that individuals may not have been born at the time of death of their sibling, or may have reached an age at which mortality exposure was less relevant for adult survival. In this paper, we use mortality when ego is *in utero*, birth to age 5, age 6 to age 18, and all mortality between birth and age 18.

In summary, the innovation of this paper is the estimation of adult mortality risks associated with mortality exposure in neighborhoods and communities while simultaneously measuring direct exposure to mortality as measured by sibling deaths in childhood. We use exposure to community mortality and sibling deaths to assess more fully if there is a causal link between early-life mortality exposure and subsequent adult mortality between age 18 and age 85. Furthermore, we assess the age-specific influences of mortality exposure and follow individuals through the life course. Early- and midlife characteristics such as socioeconomic status are included to control for mediating factors. We assess the effect of mortality exposure on all-cause and cause-specific mortality to better reveal possible causal mechanisms. Furthermore, we examine the role of mortality selection and investigate which exposure ages in childhood are most strongly associated with adult mortality. Finally, we compare two populations to assess the effect of exposure to mortality in different mortality regimes.

2. Data and methodology

We use two unique, large-scale historical demographic datasets arising from two widely differing environments: the Utah Population Database (UPDB) and LINKS Zeeland, the Netherlands. Both data sources contain large-scale demographic information about historical populations with multigenerational pedigrees, which have been linked to spatial information—census enumeration districts and municipalities, respectively—which allow us to follow individuals and their siblings, parents, and children over their life course. For Utah, we use birth cohorts from 1874–1929, following individuals until 1959–2015, and for the Netherlands birth cohorts 1811–72, following individuals until 1896–1957.

Utah, a western state in the US, was first settled by immigrants from North America and Europe after 1847. Initially, many migrants were associated with The Church of Jesus Christ of Latter-day Saints (i.e., the Mormon church), but over time the share of non-Mormons and inactive Mormons has grown. Despite poverty and harsh conditions in the early years of settlement, the population in Utah was relatively healthy (Mineau et al., 2004) in comparison with Zeeland, with lower infant and child mortality and higher life expectancy. Zeeland is a largely Protestant province in the southwestern corner of the Netherlands and consists of relatively isolated islands.

In Zeeland, a lack of potable water existed due to salinity. Adult life expectancy was low and infant and child mortality high, reaching 50% in some years and communities (Hoogerhuis, 2003; Van Dijk & Mandemakers, 2018). For both populations, high-quality family reconstitution databases exist: the Utah Population Database (UPDB), containing life courses of the population of Utah, and LINKS (LINKing System for historical family reconstruction) Zeeland, containing life histories based on linked civil certificates. Both databases contain large numbers of individuals and multi-generational pedigrees, with vital information linked to spatial information. These features make them highly suitable for assessing both contextual and familial effects on adult mortality. The UPDB has been used extensively to assess the effect of social, biological, and genetic factors on longevity and life spans (Kerber et al., 2001; Garibotti et al., 2006; Smith et al., 2009).

A portion of the UPDB is based on family group sheets on which the genealogies of the Utah founders and their descendants are recorded. These have been linked into intergenerational pedigrees with information from other official administrative and medical sources, including birth and death certificates and US Census records. The full database includes over 11 million individuals. Both individuals with and without an affiliation with The Church of Jesus Christ of Latter-day Saints (LDS, or Mormons) are included in the UPDB. For our analyses, we select individuals born in Utah between 1874 and 1929 for whom information is available about their parents and their own life course, and who have at least one sibling. Furthermore, eligible individuals must have survived to at least age 18 and are followed until death, out-migration, loss to follow-up, or survival to age 85. There are 348,964 eligible individuals in the UPDB (see Table 1), whom we call egos and individuals interchangeably.

LINKS is based on digitized vital event registration certificates that identify the individual and his or her parents. These have been used to reconstruct life courses by linking certificates pertaining to the same individuals using the first and last names of ego and his or her parents, spouses, and children. The full database contains information on almost 2 million individuals from Zeeland who experienced a vital event in that province between 1812 and 1912 for births, 1812 and 1927 for marriages, and 1812 and 1957 for deaths. For the analysis, we select 80,871 persons with a known date of birth and death, as life-course reconstructions have been shown to be reliable for these cases (Van den Berg et al., 2018), who were born between 1811 and 1872, whose parents are known and who have at least one sibling (see Table 1). Individuals are followed from age 18 until death or censoring at age 85.

In this paper, we examine mortality exposure in different phases of childhood. We consider exposure to sibling deaths during three phases: (1) from conception until birth, (2) from birth until the sixth birthday and (3) between the sixth and the 18th birthday. Additionally, we include sibling deaths before subjects are conceived, as these may still influence ego's adult mortality. Families may fail to provide children with adequate nutrition or care, leading to both increased childhood mortality and adult mortality among all offspring. Thus, it is suggested that the causal mechanism works broadly through adverse early-life conditions in the family rather than through direct exposure to sibling deaths.

2.1 Measuring community-level mortality

To assess exposure to infant and child mortality at the contextual level, we use two related approaches for Utah and the Netherlands. For Utah, we calculate infant and early and late child mortality rates for ten-year periods based on decennial census information aggregated into census enumeration districts (i.e., these approximate neighborhoods and represent a space that can be navigated by a census enumerator). For the Netherlands, we calculate these rates for municipalities by year. For both populations, community-level mortality rates are per 1,000 children. Variables are centered at the grand mean for their respective populations.

In Utah, birth certificates were not implemented before 1904. As the decennial US Census records for Utah are linked to the UPDB starting in 1880, we identify parents

Table 1 The descriptive statistics

	Utah (UDPB)				Zeeland (LINKS)			
	Min.	Max.	Mean, men	Mean, women	Min.	Max.	Mean, men	Mean, women
Age at last observation	18.01	85.00	67.40	69.65	18.01	85.00	62.69	62.72
Observed until death	0	1	0.72	0.55	0	1	0.88	0.88
Cause-specific deaths (ICD)	0	1	0.62	0.68	–	–	–	–
Cardiovascular disease	0	1	0.32	0.33	–	–	–	–
Cancer	0	1	0.12	0.15	–	–	–	–
Respiratory disease	0	1	0.07	0.06	–	–	–	–
Digestive system	0	1	0.03	0.04	–	–	–	–
Nervous system	0	1	0.03	0.04	–	–	–	–
Metabolic disease	0	1	0.02	0.03	–	–	–	–
Genito-urinal system	0	1	0.02	0.02	–	–	–	–
Infectious disease	0	1	0.01	0.02	–	–	–	–
Community mortality < 1	0.20	3.30	1.06	1.06	0.54	5.11	2.58	2.60
District mortality 1–4	0.04	2.35	0.44	0.44	0.00	2.92	1.05	1.06
District mortality 5–18	0.12	4.00	0.56	0.56	0.00	2.41	0.49	0.49
District mortality 0–18	0.66	5.71	1.93	1.92	1.22	8.30	4.13	4.15
Number of siblings	1	20	6.01	6.06	1	24	7.53	7.53
Sibling deaths								
Ego 0–18	0	12	0.52	0.50	0	13	1.81	1.82
Ego 0–5	0	7	0.17	0.17	0	7	0.78	0.79
Ego 6–18	0	8	0.34	0.34	0	10	1.03	1.03
Ego unexposed	0	16	0.40	0.41	0	16	0.97	1.01
Ego *in utero*	0	5	0.02	0.02	0	4	0.06	.06
Ego 0–18, to infection	0	4	0.04	0.04	–	–	–	–
Birth interval < 18 months	0	1	0.20	0.20	0	1	0.32	0.32
Birth order	1	20	4.04	4.06	1	24	4.28	4.32
Twin birth	0	1	0.02	0.02	0	1	0.01	0.01
Age mother at birth *12–19*	0	1	0.06	0.06	0	1	0.01	0.01
20–29	0	1	0.53	0.53	0	1	0.48	0.48
30–34	0	1	0.21	0.21	0	1	0.27	0.27
35–50	0	1	0.21	0.21	0	1	0.24	0.24
Mother dead < 45	0	1	0.07	0.08	0	1	0.13	0.13
Father dead < mother 45	0	1	0.08	0.08	0	1	0.12	0.13
SES of father	0	100	30	31	41	99	51	51
Father is a farmer	0	1	0.38	0.38	0	1	0.21	0.20
SES of father missing	0	1	0.02	0.03	0	1	0.00	0.00
SES	0	100	40	41	41	99	51	52
Farmer	0	1	0.22	0.19	0	1	0.17	0.15
SES missing	0	1	0.23	0.29	0	1	0.09	0.20
Inactive Mormon	0	1	0.30	0.28	–	–	–	–
Active Mormon	0	1	0.48	0.52	–	–	–	–
Share men/women			0.51	0.49			0.49	0.51
Total number men/women			178,404	170,560			39,849	41,022

and parental enumeration districts and find their children through the UPDB. For instance, children born between 1905 and 1915 were linked to their parental enumeration districts in the year 1910. For the years 1874–90, 1890–1905, 1905–15, 1915–25, and 1925–29, the censuses for the years 1880, 1900, 1910, 1920, and 1930 are used, respectively. Unfortunately, nearly the entire US census for 1890 was destroyed in a fire. Based on the census records and UPDB, we calculate the total number of children born and the percentage of deaths before the age of 1, ages 1–4, and 5–17 in each census enumeration district for ten-year periods. For Zeeland, we calculate comparable rates based on indexes of birth and death certificates per municipality per year. Mortality rates are averaged over five-year periods after birth, as there are small numbers of births and deaths in some municipalities. We use the municipality and year of birth to link individuals to the appropriate area-level mortality levels. Stillbirths are not included.

2.2 Sibling mortality

Sibling deaths are measured separately for different ages of ego. We include the total number of deaths among full siblings of ego, regardless of the sibling's age, between ego's birth and age 6; between ego's age 6 and 18, and the total number of sibling deaths between birth and the 18th birthday. Furthermore, we consider sibling deaths

to which ego was not exposed, occurring before ego's conception, and deaths occurring while ego was *in utero*. For Utah, we examine the number of sibling deaths from infectious disease (in parallel with noninfectious deaths) in 1904 or later, when death certificates with causes of death are available. In all analyses, we control for the number of siblings ever born.

2.3 Additional covariates

We control for several characteristics of the family of origin and circumstances in adulthood. Descriptive statistics can be found in Table 1. The demographic characteristics of individuals include the number of full siblings, the birth order, the age of the mother at birth, and the birth interval to prior birth. For Utah, we include an indicator of religious status. Active Mormons generally engage in lifestyle characteristics that are beneficial for survival, such as alcohol and tobacco avoidance, fasting, and participating in the social aspects of the religion. Inactive LDS members were baptized in the LDS church, generally at age 8, and shared an early-life environment that may be beneficial for survival, but did not express their commitment to the LDS church in adulthood. Non-LDS members have no record of involvement with the LDS church. No individual-level indicator for denomination was available for the predominantly Protestant Zeeland.

With regard to midlife characteristics, we include deaths of ego's parents before the mother reaches the end of her reproductive period (age 45), as the numbers of ego's siblings and deaths in the sibship are affected by the survival of the parents. We also include the highest socioeconomic status (SES) of the father and the individual and, given the large fraction of the populations engaged in agriculture, we distinguish farmers from nonfarmers as a dummy variable. For Utah, we use the Nam-Power SES score (1950 version) (Nam & Boyd, 2004), using the highest SES observed in the censuses and, if unavailable, the SES from the death certificate of the father. For Zeeland, we use the highest HISCAM SES (Lambert et al., 2013) from any linked vital event certificate to measure socioeconomic status, and HISCLASS (Van Leeuwen & Maas, 2011) to distinguish farmers from nonfarmers. For persons with a missing SES score, we impute the mean score and include a dummy indicating whether the SES score was imputed. For this period, males were the primary breadwinners for both Utah and Zeeland, resulting in under-registration of female labor participation, with few recorded occupations for women. Therefore, for women, we use the socio-economic status of the spouse. If there is no known spouse (Utah 13.7%, Zeeland 19.8%), we assign the mean male socioeconomic status and a dummy variable indicating whether missing values were imputed and included as a covariate.

3. Results

We estimate sex-specific Cox proportional hazard regressions. Individuals are followed from age 18 until death, last date of observation (Utah), or censoring at age 85, the oldest achievable age for the last Utah birth cohort (1929–2015). All models are estimated using R version 3.3.2. Men have higher mortality rates than women for sociocultural and biological reasons, including behavioral patterns, the protective role of estrogen, the suppressing effect of testosterone on immunity, and an unprotected X chromosome (Lindahl-Jacobsen et al., 2013). The effect of early-life conditions on survival may therefore be stronger for men than women (Smith et al., 2009; Doblhammer et al., 2013); on the other hand, cohort mortality may affect both women and men similarly (Lindeboom et al., 2010). We analyze the relative risk of mortality first by exposure to high infant mortality, early and late childhood mortality in the community and second by exposure to sibling mortality in the first six and 18 years of life. Finally, we analyze the cause-specific mortality risks using a competing-risk approach. We first present crude, unadjusted effects in Model 1. In Model 2, we add demographic characteristics of the family to control for possible confounders. In Model 3, we assess whether effects are mediated by a set of midlife characteristics. All analyses are stratified by year of birth to account for temporal changes in exposure to sibling mortality and adult mortality hazards. In these stratified models, the hazard of adult mortality as a function of exposure to sibling mortality is estimated based on comparisons with individuals from the same birth cohort.

3.1 Exposure to mortality in the community

We begin by reporting results that demonstrate crude (unadjusted) associations between exposure to community mortality during childhood and adult mortality past age 18. For Utah, we use all egos who can be linked to a parental enumeration district closest to their birth: 149,285 men and 141,992 women. For Zeeland, we use places of birth and municipal mortality rates in the first five years after birth. Estimates of relative risks are based on a community-level increase of mortality by 10%. The results are shown in Table 2.

Estimates of the relation between community mortality rates in childhood and adult mortality rates show that in Utah and Zeeland, there is a significant positive relation between community mortality rates for infants and for young and older children, and adult mortality rates for men and women. For females in Zeeland, only infant mortality, early child mortality, and total child mortality are related to adult mortality risk (Table 2, Model 1). After including controls for the demographic characteristics of the individual and the family of origin and socioeconomic characteristics of the family of origin, these effects remain significant. In Model 3, we add individual-level control

Table 2 Childhood community levels of mortality among infants, young, and older children and survival between ages 18 and 85 in Utah and Zeeland, relative risk, Cox regression

Men		Model 1				Model 2				Model 3			
	District mortality	Utah		Zeeland		Utah		Zeeland		Utah		Zeeland	
		HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI
	Infants	**1.09**	1.07–1.12	**1.01**	1.00–1.03	**1.08**	1.06–1.11	**1.02**	1.00–1.04	1.02	1.01–1.06	1.01	0.99–1.03
	Young children	**1.10**	1.05–1.14	**1.07**	1.03–1.11	**1.08**	1.02–1.10	**1.06**	1.02–1.10	0.99	0.95–1.03	**1.04**	1.00–1.08
	Older children	**1.17**	1.13–1.21	**1.09**	1.04–1.15	**1.15**	1.08–1.16	**1.07**	1.02–1.12	1.03	0.99–1.06	**1.06**	1.01–1.11
	All mortality	**1.08**	1.07–1.10	**1.03**	1.01–1.04	**1.07**	1.05–1.08	**1.03**	1.01–1.04	1.01	1.00–1.03	**1.02**	1.00–1.03
	N	149,285		39,849									
	Events	106,568		35,404									
Women		**Model 1**				**Model 2**				**Model 3**			
	District mortality	Utah		Zeeland		Utah		Zeeland		Utah		Zeeland	
		HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI
	Infants	**1.08**	1.05–1.12	**1.05**	1.03–1.07	**1.08**	1.05–1.11	**1.06**	1.04–1.07	**1.03**	1.02–1.08	**1.06**	1.04–1.08
	Young children	**1.13**	1.08–1.18	**1.03**	1.00–1.07	**1.12**	1.06–1.16	**1.03**	1.00–1.07	**1.05**	1.01–1.11	1.01	0.98–1.05
	Older children	**1.09**	1.05–1.14	1.00	0.96–1.05	**1.08**	1.02–1.10	1.00	0.96–1.05	0.99	0.96–1.04	0.96	0.92–1.01
	All mortality	**1.07**	1.05–1.09	**1.04**	1.02–1.05	**1.07**	1.04–1.08	**1.04**	1.02–1.05	1.02	1.01–1.05	**1.03**	1.03–1.05
	N	141,992		41,022									
	Events	77,449		36,289									

Notes: All models are stratified by birth year. Bold: statistical significance $p < 0.05$.
Model 1: uncontrolled effects. Model 2: controlled for number of siblings, birth interval, birth order, age of mother at birth and twin birth, and the socioeconomic characteristics of the family of origin. Model 3: controls Model 2, plus survival of the father and mother until the end of the reproductive period of the mother, religion (Utah), and socioeconomic characteristics in adulthood.

variables including the SES, the survival of parents and religious status (Utah). For men in Utah, we no longer find a significant association between childhood mortality rates in the community and mortality rates in adulthood. Most of the effects are mediated by religion in adulthood (results not shown). For women in Utah, the effect sizes are reduced but the effects remain significant for infant mortality rates and early child mortality rates. For Zeeland, early child mortality is no longer significantly associated with male adult mortality rates and infant mortality is no longer significantly associated with female adult mortality rates. For women, we find that infant and total child mortality affect adult mortality rates for Zeeland. Overall, for both Utah and Zeeland, the most consistent effect of community-level mortality when egos were children was infant mortality for women. For Zeeland, we also find an effect of all childhood mortality in their community for both males and females.

3.2 Exposure to sibling mortality

A proximate measure of exposure to death, most likely from infectious diseases, is childhood deaths among ego's siblings. In Table 3, results can be found for Cox proportional hazard models analyzing the mortality hazard rate between ages 18 and 85. Proportionality tests do not indicate that the effects of the covariates vary with age for either men or women in Zeeland or Utah.

The results show that sibling deaths are related to a higher relative risk of mortality for both men and women. In Utah, where relatively few sibling deaths occur, sibling mortality has a stronger effect on female mortality in adulthood than on male adult mortality. In Zeeland, where sibling deaths are relatively common, male adult mortality is particularly affected. Here, greater exposure in the family of origin—two or more sibling deaths—is significantly and adversely associated with adult mortality, whereas in lower-mortality Utah, one or two sibling deaths are also significantly related to higher mortality in adulthood. The effects of sibling mortality are not mediated by the demographic and socioeconomic characteristics of the family of origin or adulthood in either population. For both men and women, there is a dose–response pattern that indicates that with increasing sibling deaths there is increased adult mortality for both populations (see Table 4), controlling for sibship size.

Given that there is an effect of sibling deaths during the life of ego as well as before ego's conception, it is likely that the causal force underlying the observed association is based on family circumstances. These circumstances include dietary choices, lifestyle characteristics, and stress experienced by parents after child loss, as well as direct exposure to infections and disease of siblings, features that are shared within families.

Table 3 Sibling deaths in childhood and survival between ages 18 and 85

Men		Model 1				Model 2				Model 3			
Sibling deaths		**Utah**		**Zeeland**		**Utah**		**Zeeland**		**Utah**		**Zeeland**	
		HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI
Ego unexposed	1 or more deaths	**1.04**	1.02–1.05	**1.02**	1.00–1.05	**1.04**	1.02–1.05	1.02	0.99–1.05	**1.04**	1.02–1.05	1.02	0.99–1.04
Ego *in utero*	1 or more deaths	**1.05**	1.00–1.10	**1.05**	1.01–1.10	**1.05**	1.00–1.10	**1.05**	1.00–1.10	1.04	0.99–1.08	1.04	0.99–1.09
Ego 0–5	1 death	**1.05**	1.03–1.06	**1.03**	1.01–1.06	**1.04**	1.03–1.06	**1.03**	1.01–1.06	**1.04**	1.02–1.06	**1.03**	1.00–1.05
	2 deaths	**1.08**	1.03–1.12	**1.10**	1.06–1.13	**1.07**	1.03–1.11	**1.09**	1.06–1.13	**1.06**	1.02–1.10	**1.08**	1.05–1.12
	More than 2 deaths	1.08	0.98–1.18	**1.08**	1.03–1.13	1.07	0.98–1.17	**1.08**	1.03–1.13	1.07	0.97–1.17	**1.07**	1.02–1.12
Ego 6–18	1 death	**1.07**	1.06–1.09	1.02	0.99–1.05	**1.08**	1.06–1.09	**1.04**	1.01–1.06	**1.07**	1.05–1.08	**1.03**	1.01–1.06
	2 deaths	**1.08**	1.05–1.10	**1.07**	1.03–1.11	**1.08**	1.06–1.11	**1.10**	1.06–1.13	**1.07**	1.05–1.10	**1.09**	1.05–1.13
	More than 2 deaths	**1.11**	1.06–1.15	**1.07**	1.04–1.11	**1.11**	1.07–1.16	**1.11**	1.07–1.16	**1.10**	1.05–1.14	**1.10**	1.06–1.15
Ego 0–18	1 death	**1.08**	1.06–1.09	1.02	0.99–1.05	**1.08**	1.06–1.09	1.02	0.99–1.05	**1.07**	1.06–1.09	1.02	0.99–1.05
	2 deaths	**1.09**	1.07–1.12	**1.08**	1.05–1.12	**1.09**	1.07–1.12	**1.09**	1.06–1.13	**1.08**	1.06–1.11	**1.08**	1.05–1.12
	More than 2 deaths	**1.12**	1.09–1.15	**1.11**	1.07–1.15	**1.12**	1.09–1.16	**1.14**	1.10–1.18	**1.11**	1.08–1.14	**1.13**	1.09–1.17
	Events	127,749		35,404									
	N	178,404		39,849									
Ego 0–18	Infection, 1 or more deaths	**1.06**	1.02–1.10			**1.06**	1.02–1.10			**1.04**	1.00–1.08		
	Events	67,112											
	N	99,800											
Women		Utah		Zeeland		Utah		Zeeland		Utah		Zeeland	
Ego unexposed	1 or more deaths	**1.04**	1.03–1.06	**1.04**	1.02–1.06	**1.04**	1.02–1.06	**1.03**	1.00–1.05	**1.04**	1.02–1.06	**1.03**	1.01–1.06
Ego *in utero*	1 or more deaths	**1.04**	1.00–1.08	0.99	0.95–1.04	1.03	0.98–1.08	0.99	0.94–1.03	1.03	0.97–1.08	0.99	0.94–1.03
Ego 0–5	1 death	**1.12**	1.10–1.17	**1.03**	1.00–1.05	**1.11**	1.09–1.13	**1.03**	1.01–1.06	**1.11**	1.09–1.14	**1.03**	1.00–1.05
	2 deaths	**1.12**	1.07–1.17	1.02	0.98–1.05	**1.11**	1.07–1.16	1.02	0.99–1.05	**1.11**	1.06–1.16	1.02	0.99–1.05
	More than 2 deaths	**1.16**	1.05–1.28	0.99	0.95–1.04	**1.15**	1.04–1.27	1.00	0.95–1.04	**1.18**	1.07–1.30	0.99	0.95–1.04
Ego 6–18	1 death	**1.10**	1.08–1.11	1.02	0.99–1.04	**1.10**	1.08–1.11	**1.03**	1.00–1.06	**1.10**	1.08–1.12	**1.04**	1.01–1.06
	2 deaths	**1.12**	1.08–1.15	**1.04**	1.00–1.07	**1.12**	1.09–1.15	**1.06**	1.03–1.10	**1.12**	1.08–1.15	**1.07**	1.04–1.11
	More than 2 deaths	**1.09**	1.04–1.14	**1.05**	1.01–1.09	**1.10**	1.05–1.15	**1.10**	1.05–1.14	**1.10**	1.05–1.15	**1.12**	1.08–1.17
Ego 0–18	1 death	**1.14**	1.12–1.16	**1.03**	1.00–1.06	**1.14**	1.12–1.15	**1.03**	1.00–1.06	**1.14**	1.12–1.16	**1.04**	1.01–1.07
	2 deaths	**1.15**	1.12–1.17	1.02	0.99–1.06	**1.15**	1.12–1.18	**1.04**	1.00–1.07	**1.15**	1.12–1.18	**1.04**	1.00–1.07
	More than 2 deaths	**1.16**	1.12–1.20	**1.05**	1.02–1.08	**1.16**	1.13–1.20	**1.08**	1.04–1.11	**1.16**	1.13–1.20	**1.09**	1.05–1.13
	Events	94,047		36,289									
	N	170,560		41,022									
Ego 0–18	Infection, 1 or more deaths	**1.13**	1.08–1.18			**1.13**	1.08–1.18			**1.12**	1.07–1.17		
	Events	46,701											
	N	86,975											

Notes: See Table 2. Model 1 controls for the number of siblings.

Table 4 Test of trend for exposure to sibling deaths in childhood and adult mortality, ages 18–85

	Men			Women		
	HR	95% CI	No. of events	HR	95% CI	No. of events
Zeeland						
All-cause mortality	**1.05**	1.01–1.04	35,404	**1.03**	1.04–1.06	36,289
Utah						
All-cause mortality	**1.05**	1.04–1.05	127,749	**1.07**	1.06–1.08	94,047
Cardiovascular	**1.05**	1.04–1.07	40,781	**1.07**	1.06–1.09	30,875
Cancer	**1.03**	1.01–1.06	15,174	**1.07**	1.05–1.10	13,736
Respiratory system	**1.07**	1.05–1.10	9,305	**1.12**	1.08–1.16	5,497
Digestive system	**1.09**	1.05–1.14	4,297	**1.09**	1.04–1.14	3,635
Nervous system	**1.06**	1.02–1.11	3,444	**1.10**	1.06–1.15	3,602
Metabolic disease	**1.07**	1.01–1.13	2,532	**1.12**	1.07–1.17	3,282
Genitourinary system	**1.07**	1.01–1.13	1,980	1.05	0.99–1.12	1,928
Infectious disease	1.04	0.99–1.11	1,848	1.06	0.98–1.12	1,481

Notes: Significant effects ($p < 0.05$) are shown in bold. Models are stratified by birth year and control for number of siblings, birth interval, birth order, age of mother at birth and twin birth, socioeconomic characteristics of the family of origin, survival of the father and mother until the end of the reproductive period of the mother, and religion (Utah).

For Zeeland women, exposure to deaths during late childhood appears more lethal than those occurring earlier in life, while in Utah both are related to adult survival. For men, sibling deaths in both early and late childhood are related to adult mortality. Possibly, for women, caretaking tasks in late childhood play a role. As older girls were expected to care for sick siblings, there may be stronger exposure to their siblings' illnesses and deaths in this period of life. Alternatively, girls could be more susceptible to illness and disease in their teenage years, as research has found excess mortality among girls during this period of life (Devos, 2006). In line with these explanations, we find that siblings' deaths due to infectious disease affect female mortality in adulthood more than male mortality, although the difference between men and women is not significant (results not shown).

3.3 Causes of death

To investigate further the mechanism linking exposure to mortality in childhood and increased risks of adult mortality, we analyze the link between causes of death for ego and exposure to sibling mortality. Causes of death are not available for Zeeland.

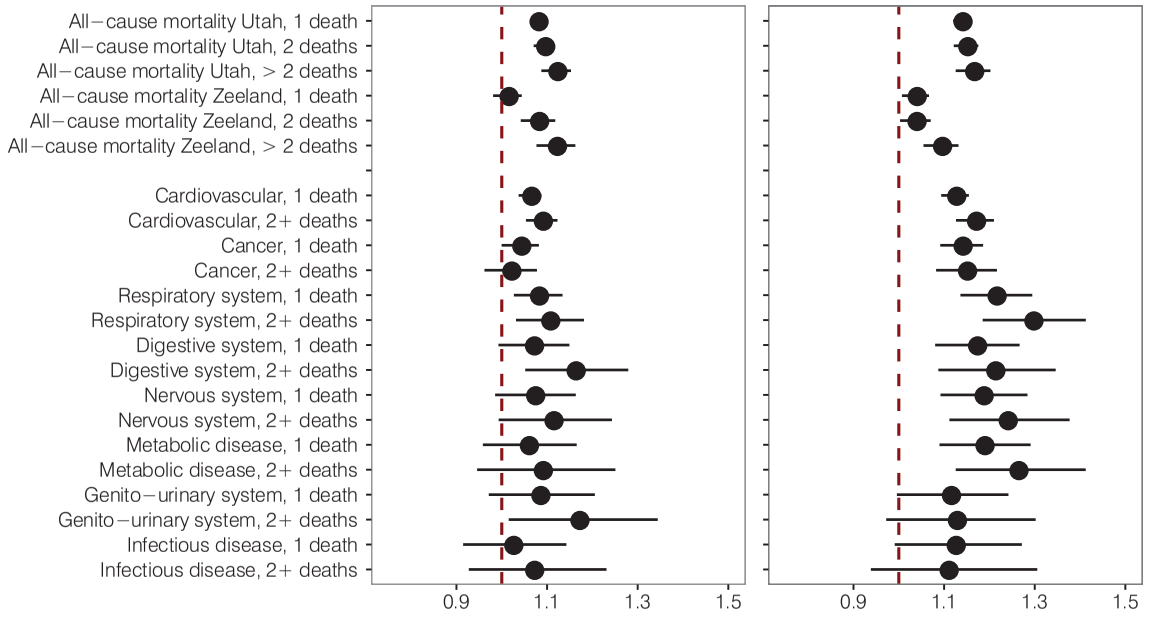
The causes of death have been collapsed into 17 broad categories, adopted by the US National Center for Health Statistics, using the International Classification of Diseases (ICD) codes. We include the following selected causes of death: infection, all cancers, cardiovascular disease, metabolic and endocrine systems, nervous system, respiratory system, digestive system, and diseases of the genitourinary system. We apply a competing competing-risk approach, estimating the risk of death before the age of 85 for each of these specific causes of death, where we treat deaths from all other causes as censored when analyzing any specific cause of death. Descriptive statistics for the causes of death can be found in Table 1; the results are summarized in Figure 1.

The results show that there is a significant relation between exposure to mortality of siblings in the family of origin and deaths due to cardiovascular disease, cancers, and diseases of the respiratory system and the digestive system. For women, the effect is also present for metabolic diseases and diseases of the nervous system (see Figure 1). Test of trend (see Table 4) shows that there is a dose–response pattern between the number of sibling deaths for men and women for all-cause mortality, cardiovascular disease, cancer, diseases of the respiratory system, digestive system, and nervous system, diseases of the genitourinary system (women), and metabolic diseases. In addition, sibling deaths due to infection are significantly related to ego all-cause mortality, and death from cardiovascular disease and metabolic disease for men and women, and to deaths from respiratory and digestive system diseases (see Table 5). A linear contrast between sibling deaths due to infectious disease and other causes of death indicates that all-cause mortality and mortality from cardiovascular disease and metabolic disease are affected differently by sibling deaths due to infectious disease than by sibling all-cause mortality (see Table 5). That also applies to systems of the digestive system, but here, sibling deaths due to infectious disease do not significantly affect the risk of death due to diseases of the digestive system.

3.4 Robustness checks

We conducted several sensitivity analyses. First, we selected only cases in which the parents lived until the end of the reproductive period of the mother, as the death of a parent likely affects sibling survival and adult mortality. We also selected only individuals with birth order two and up, as firstborn children may not be as likely to be exposed to sibling mortality and never had sibling mortality before their conception by definition. In addition, we conducted robustness checks estimating models with all effects included simultaneously, and using only individuals whose parents were located in the census records close to their year of birth. The results are available on request and were not substantively different from the analyses presented here.

Figure 1 Causes of death by number of sibling deaths, men (left) and women (right)



Notes: Relative risk, 95% confidence interval. Cause-specific mortality concerns Utah. The models are stratified by birth year. The models control for the number of siblings, the mother's age at birth, the birth order and interval, religion (Utah), and the socioeconomic characteristics of the father

Table 5 Cause-specific mortality age 18–85 by sibling death due to infection, Utah

	Deaths due to infectious disease		Other deaths				
	HR	95% CI	HR	95% CI	No. of events	Difference in deviance	*p*-value
All-cause mortality	**1.08**	**1.05–1.11**	**1.16**	**1.14–1.17**	113,813	**17.46**	**0.00**
Cardiovascular	**1.11**	**1.06–1.16**	**1.17**	**1.14–1.19**	35,286	**3.82**	**0.05**
Cancer	1.01	0.94–1.09	**1.16**	**1.12–1.20**	18,319	1.86	0.17
Respiratory system	**1.11**	**1.00–1.23**	**1.20**	**1.14–1.27**	8,103	0.51	0.47
Digestive system	1.12	0.97–1.30	**1.24**	**1.15–1.32**	4,086	**4.49**	**0.03**
Nervous system	**1.27**	**1.08–1.49**	**1.28**	**1.18–1.39**	3,532	2.49	0.11
Metabolic disease	**1.40**	**1.20–1.63**	**1.19**	**1.10–1.29**	3,485	**3.78**	**0.05**
Genitourinary system	1.10	0.83–1.32	**1.13**	**1.12–1.41**	1,798	0.76	0.38
Infectious disease	1.11	0.86–1.44	**1.22**	**1.07–1.39**	1,314	1.36	0.24

Notes: Total number of cases: 190,756 (men, 99,800; women, 95,083). The difference in deviance is based on nested models that estimate the effects of siblings who die from any cause first—treating the effects of sibling deaths from infectious and noninfectious deaths as having the same effect—compared to a model in which sibling deaths due to infectious disease and other causes of death are allowed to have distinct effects. The models are stratified by birth year and control for the number of siblings, the birth interval, the birth order, the age of the mother at birth and twin birth, the socioeconomic characteristics of the family of origin, the survival of the father and mother until the end of the reproductive period of the mother, and religion.

4. Discussion and conclusion

We have examined whether exposure to deaths during childhood may create enduring health effects. The premise of our approach is that families play a pivotal role by constructing the circumstances by which children were exposed to mortality among their siblings. In line with work by Finch and Crimmins (Finch & Crimmins, 2004; Crimmins & Finch, 2006), we showed that there is an association between exposure to mortality in communities and birth cohorts and subsequent adult mortality rates. Especially for infant and all-child mortality, we detect effects on adult mortality rates, consistent with some research (Schellekens & van Poppel, 2016) but only partially in line with the findings of Finch and Crimmins (2004), who suggest that, primarily, infant mortality is related to the risk of adult mortality.

Research shows that exposure to sibling deaths is related to later-life stress (Norton et al., 2017) and increases mortality in early and late childhood and early adulthood (Yu et al., 2017). In line with these results, we found that sibling deaths are related to reduced survival at adult ages. Indeed, we found stronger evidence for negative long-term effects of exposure to sibling deaths than community-level exposure to mortality. After taking family-level characteristics and individual-level characteristics into account, we found that exposure to sibling deaths has enduring effects on adult mortality rates. Community-level mortality has a persistent effect on adult mortality as well, but the effects are weaker, reduced after including family characteristics and became insignificant after including midlife characteristics.

We did not find evidence that weak mortality crises have a selection effect that overwhelms potential scarring effects. Putative scarring effects appear to dominate selection effects, even if there is only a single exposure to sibling death (Hatton, 2011). Despite the potential ameliorating effects of immunization and access to modern health technologies and care, including antibiotics, we observed that even a single exposure to a sibling death in Utah is related to adult mortality rates. In Zeeland—where child mortality was very high—exposure to multiple sibling deaths is adversely related to adult survival, while exposure to only one sibling death is not. Possibly, in environments with high mortality, the event of a sibling death does not necessarily contribute to adult mortality risks, as individuals are likely to be exposed to deaths in childhood in general. Considering lower adult life expectancies in Zeeland, the cumulative damage of adverse exposures is stronger in Zeeland than in Utah, an explanation for weaker effects of sibling deaths in Zeeland. In line with this explanation, we found that in Zeeland and Utah, the effects of sibling mortality are stronger in later, lower-mortality birth cohorts (results available on request).

In contrast to work that has found no link between exposure to infectious disease and all-cause mortality, deaths due to cardiovascular disease, cancer, and stroke (Hayward et al., 2016), we found that exposure to sibling deaths is related to all-cause adult mortality as well as mortality due to cardiovascular disease, cancer, diseases of the respiratory, digestive, and nervous systems (women), and metabolic disease (women). Thus, exposure to all-cause early-life mortality among siblings is related to an array of causes of death and appears to be a general risk factor for premature adult mortality. In addition, the mortality risk of these diseases follows a dose–response pattern. The extent to which individuals are exposed to sibling mortality is important for understanding adult survival: the greater the number of sibling deaths, the stronger are the effects on adult all-cause mortality and cause-specific mortality. On the other hand, exposure to sibling deaths due to infection is related to cardiovascular disease and metabolic diseases for men and women. Thus, deaths attributable to infections appeared to be specifically related to chronic inflammatory responses. Sibling all-cause mortality is related to several common causes of death, suggesting that there may be a mechanism beyond inflammatory immune responses at play relating early-life exposure to mortality to adult mortality rates.

A few limitations are noteworthy regarding the study design. The measurement of infant and early and late childhood mortality constituted a challenge both for the Utah and Zeeland data. For the Netherlands, out-migration information is unknown and, therefore, we used death risks relative to birth rates. For Utah, we necessarily relied on census data, which requires us to accept ten years between measurements of several key variables. As a result, mortality conditions for individuals and families are measured at different ages of the egos and estimates of the relation between mortality exposure and adult mortality are conservative. Furthermore, while we were able to distinguish exposure to sibling deaths at several ages of the ego, we included mortality among siblings at all ages, while mortality among siblings of comparable age could be more influential than deaths among siblings who differed in age. Finally, although we were better able than existing research to distinguish between individuals who were most likely and less likely to have been exposed to illness and disease, it remains unknown whether individuals fell ill themselves. Using modern data, it should be possible to determine the extent to which earlier illnesses are related to lower adult life expectancies.

Frailty (random effects) models constitute an alternative modeling approach to the family-clustered survival models that were employed in this study. These models have the advantage of comparing adult mortality between exposed and unexposed individuals within families, thus controlling for family-unobserved heterogeneity. These models rely on a strict assumption that family-specific frailty is independent of

the other exposures, notably sibling deaths. It is unlikely that frailty, as an indicator of family-specific risk of death, is independent of exposure to sibling deaths. Moreover, it is likely that exposure to sibling mortality is a sibling-shared experience because of the included age ranges of egos in relation to ages of their siblings. Alternatively, to estimate family fixed-effects models, a much more select sample would have been chosen for analysis, as at least three siblings need to be observed: two siblings who survive into adulthood and a third sibling to put them at risk of sibling mortality in childhood. This restriction would have resulted in a much smaller sample size (a 40% and 30% reduction of the sample size in Utah and Zeeland, respectively). We therefore included age-specific exposure to sibling deaths in our main models rather than using frailty models, and assessed the effects of family-shared characteristics by incorporating the effect of sibling deaths that occur before the birth of the ego. The results show that at least part of the effect of sibling mortality on adult mortality works at the family level rather than the individual level.

The strengths of this paper are, first, that we have been able to show convincingly the enduring effects of sibling deaths on adult mortality risk over the life course. Second, we have used two large high-quality databases on historical populations from Zeeland and Utah, with follow-up until 2015 for Utah, which has enabled us to compare across mortality regimes and living conditions. This implies that the relationship that we have identified between mortality exposure and adult mortality risk is applicable to populations across contexts. Furthermore, the results that we have found were subjected to several robustness checks, which adds to the strength of our analysis. We were able to assess mechanisms linking early-life mortality exposure to adult mortality risk by studying micro-level causes of death. We jointly assessed community and family effects. Although research has suggested that period influences are more important than early-life exposures (Myrskylä, 2010), we have been able to show that early-life exposures do matter greatly for adult health and mortality and that the family plays an important role in shaping and mediating these influences.

Part III

Summary: “Death and the Family”

This dissertation discusses the themes of infant and child mortality, and the consequences of individuals’ exposure to disease and mortality on the life course and child mortality in the next generation. Earlier research has shown that child mortality is related to family background: the chances of child mortality are higher in some families than in others, the so-called high-risk families. In this dissertation, a set of research questions related to these high-risk families is explored. The central aim of the dissertation is to explore whether and how exposure to child mortality affects the further lives of their siblings, and if and how this exposure is related to elevated infant and child mortality in the second generation.

A review of the existing literature on high mortality in families—childhood death clustering—has been conducted. In existing research to childhood mortality clustering a variety of methods is used to determine whether deaths cluster in certain families. The review shows that there is evidence that several mechanisms are involved in shaping high mortality in high-risk families. These concern, among others, the effects of scarring: children who are exposed to sibling mortality have a greater likelihood of childhood mortality than children who have not been exposed to sibling mortality. This explanation, however, does not provide insights into the mechanisms that connect the increased likelihood of mortality between siblings. Other research has shown that short birth intervals, related to maternal depletion syndrome, and competition over limited resources in the family of origin play an important role in mortality of young children. Maternal education, sources of income, and poverty are important too. Finally, little research has been done on the role of the wider family network and fathers, and on mortality clustering after infancy.

The dissertation focuses on the province of Zeeland, in the southwest of the Netherlands, which consists of a group of islands and coastal regions, mainly dependent on agriculture and fishery. This province had high child mortality for a long time, higher than other regions of the Netherlands, Europe, and the rest of the world at the time. It was only late in the 19th century that child mortality started to decrease rapidly, to a level lower than in other regions. Infectious disease played a large role in the high child mortality and in its decline to relatively low levels in a later stage. For the province of Zeeland, several data sources are available for historical demographic research. Indexes of certificates of birth, marriage, and death certificates have been computerized over recent decades. Using name-based algorithms, individuals and their parents—whose names were mentioned on the birth and marriage certificates, and for individuals who died relatively young also on the death certificate—have been linked together, reconstructing families from generation to generation. These reconstruc-

tions of families allow us to conduct research about the demographic history of this province, including questions that concern the whole life course and intergenerational change. The resulting database—the LINKing System for historical family reconstitution, or LINKS—is unique to the Netherlands, with only a few projects worldwide that are comparable in size and setup.

At the start of the doctoral project, the quality of the family reconstructions was unclear. To check the reliability of the life course and family reconstructions, results from LINKS have been compared with family reconstructions of the same individuals in an established data source for historical demographic research, the Historical Sample of the Netherlands (HSN). Given the fact that the HSN is a population sample, and that all individuals experiencing a vital event in Zeeland between 1812 and 1912 were present in LINKS—which contains all civil certificates for that period—all individuals born in Zeeland between 1812 and 1912 and present in the HSN had to be present in LINKS. Life-course and family reconstructions from both databases have been compared, to assess differences that have been introduced as a consequence of the employment of the separate registration systems and two different ways of conducting family reconstruction. Based on this exploration, we conclude that that although, in comparison to LINKS, the HSN underestimates the total number of siblings of research persons, especially if they died at a young age, in general, the life-course and family reconstructions are very similar. Furthermore, in LINKS in particular, life-course reconstructions for individuals for whom a death certificate in the province of Zeeland is available are of a good quality. Out-migration is, as expected, an important source of a failure to follow life courses of Zeeland-born individuals over time. The end of observations on birth certificates in 1912, of marriage certificates in 1937, and of death certificates in 1962 is another factor that has to be taken into account when conducting research on the LINKS database.

In the research papers at the core of the dissertation, I take a closer look at causes of child mortality at the level of the family, but also at the effects of high mortality in the family of origin on the life course. In a first step, I show that there is a stronger concentration of infant and child mortality in some families than chance would predict. In the papers, I further show that there is a significant intergenerational correlation in infant and child mortality. A variety of methods has been used to establish that this correlation does exist and is not a product of the chosen methodology. Controlling for a large number of family characteristics does not change this conclusion. Both in the paternal and the maternal line there is an increased risk of mortality if the parent lost more siblings when he or she was a child. However, the role of the mother is stronger than the role of the father. In considering the intergenerational correlation in high early-life (infant and child) mortality and its causes, both causal mechanisms—independent, causal effects of mortality exposure on the survival of one's children—as well as mechanisms that are related to socially transmitted characteristics related within families, including family socioeconomic and behavioral characteristics, were found to be relevant. As the socioeconomic structure of the society changed over time and behavior and knowledge on caring for infants, hygiene, medical care, and nutrition became more widespread, infant and child mortality slowly declined. Fewer children were exposed to mortality among their brothers and sisters, and consequently, fewer children's health and survival was negatively affected by exposure to sibling mortality, thus reinforcing the transition toward better health for young children.

Gender discrimination and unequal treatment of boys and girls in the household may also have played a role in the concentration of infant and child mortality in families. Because the unequal treatment of boys and girls is culturally determined, these patterns should be visible from generation to generation, while girls would be more strongly affected by mortality in these families than boys. However, here, no proof for gender discrimination is found. Girls are no more strongly affected in families with a history of high infant and child mortality than boys. Furthermore, socioeconomic status affects the chance that children might die. Not only is the socioeconomic status of the family of origin relevant, but also the socioeconomic status of the grandparents. In particular, the children and grandchildren of farmers have an advantage with regard to survival in infancy and in childhood in comparison to other socioeconomic groups.

In the last paper, I explore the effects of exposure to infant and child mortality on survival in adulthood. It is found that exposure to infant and child mortality in the family of origin contributes to adult mortality risks in several age groups. Mortality due to cardiovascular disease and diabetes in particular is higher for individuals from high-mortality environments, which is in line with the inflammatory immune response hypothesis, which suggests that disease exposure is related to chronic inflammatory immune responses that negatively affect cardiologic and metabolic health over time. Effects of exposure to death in the family of origin appear to be stronger than effects within regions of origin, which further supports the idea that families play an important role in exposing family members to disease and death, thus affecting the health and survival of individuals.

Nederlandstalige samenvatting

In dit proefschrift staat de rol van de familie in het bepalen van overlevingskansen centraal. De thema's zuigelingen- en kindersterfte en de gevolgen van blootstelling aan ziekte en sterfte voor de latere levensloop en kindersterfte in de volgende generatie worden behandeld. Uit eerder onderzoek is gebleken dat kindersterfte geen op zichzelf staande gebeurtenis is: er zijn families, waar de kans op kindersterfte aanzienlijk groter is dan in andere families. Deze families worden ook wel hoogsterftefamilies genoemd. In dit proefschrift worden verschillende vragen rondom deze hoogrisicofamilies behandeld. Het centrale doel van het proefschrift is onderzoeken hoe blootstelling aan kindersterfte de latere levensloop van hun broertjes en zusjes beïnvloedt, en of, en zo ja hoe, deze blootstelling ook in de tweede generatie weer leidt tot een hogere kindersterfte.

Het proefschrift richt zich op de provincie Zeeland. Deze provincie kende lange tijd een hoge kindersterfte, hoger dan destijds op veel andere plaatsen in de wereld en in Europa en Nederland in het bijzonder. Pas laat in de 19e eeuw begon de kindersterfte langzaam af te nemen, om daarna relatief snel te dalen naar een niveau dat lager was dan elders. Infectieziekten speelden een grote rol in de aanvankelijk hoge kindersterfte, maar ook in de rappe daling van het sterfteniveau. Daarnaast vormde de beschikbaarheid van zuiver drinkwater een probleem. Verschillende databronnen zijn beschikbaar over het 19e en 20e eeuwse Zeeland. Het proefschrift is gebaseerd op gedigitaliceerde indexen van de geboorte-, overlijdens en huwelijksregisters van Zeeland, die in de afgelopen decennia zijn gedigitaliseerd. Met behulp van de namen van individuen en hun ouders—die vermeld werden op geboorte- en huwelijkscertificaten, en voor jonggestorvenen ook op de overlijdenscertificaten—zijn families gereconstrueerd. Hierdoor kunnen families van generatie op generatie worden gevolgd. De resulterende database—LINKS, ofwel LINKing System for historical family reconstruction —is uniek voor Nederland, en kent qua omvang slechts enkele vergelijkbare projecten in de wereld. Door het in kaart brengen van de familienetwerken kunnen we vragen beantwoorden over levenslopen van mensen uit Zeeland, en de intergenerationele samenhang van geboortepatronen, huwelijkspatronen en sterfte binnen families.

Bij aanvang van het doctoraatsproject was het nog onduidelijk wat de kwaliteit was van de levensloop- en familiereconstructies in LINKS. Om de kwaliteit van de data te toetsen, zijn de levensloop- en familiereconstructies binnen LINKS vergeleken met familiereconstructies van dezelfde individuen binnen een bekende en veelgebruikte historisch-demografische steekproef van levenslopen, de Historische Steekproef Nederland (HSN). Hierbij werd gebruik gemaakt van het gegeven dat de volledige populatie van Zeeland die tussen 1812 en 1912 geboren werd, overleed of trouwde

terug te vinden is in LINKS. Daardoor zijn alle personen in de HSN die in Zeeland woonde en hier een levenslooptransitie beleefden—een huwelijk, sterfgeval of geboorte—ook in LINKS te vinden. 485 personen geboren in Zeeland konden worden geïdentificeerd in zowel LINKS als de HSN. De levensloop- en familiereconstructies van dezelfde personen zijn vergeleken tussen de beide databanken, om zo te onderzoeken of de verschillende registratiesystemen en technieken voor het reconstrueren van levenslopen en families tot andere uitkomsten hebben geleid. We concluderen op basis van deze vergelijking dat hoewel er een onderschatting is van het aantal geboren kinderen, en met name van het aantal jonggestorven kinderen, in de HSN ten opzichte van LINKS, de familie- en levensloopreconstructies verder sterk op elkaar lijken. Verder zijn in LINKS met name de levenslopen van personen voor wie in Zeeland een sterftecertificaat wordt gevonden van uitstekende kwaliteit. Uitmigratie is, zoals verwacht, de belangrijkste oorzaak van het niet koppelen van de verdere levensloop aan geboortecertificaten, evenals het aflopen van de onderzoeksperiode na 1912 (geboorte), 1937 (huwelijk) en 1962 (overlijden).

De kern van de dissertatie begint met een review van de bestaande literatuur over sterfteclustering op jonge leeftijd. Naar dit fenomeen wordt onderzoek gedaan zowel in historische als in de contemporaine populaties en worden verschillende methoden toegepast, die voortkomen uit verschillende wetenschapstradities. Uit de inventarisatie van de literatuur blijkt dat er in hoogsterftefamilies waarschijnlijk veel verschillende mechanismen samenwerken die samen leiden tot een hogere kindersterfte in deze gezinnen. Het gaat dan onder andere om effecten van *scarring:* kinderen die zijn blootgesteld aan sterfte van broertjes en zusjes hebben een grotere sterftekans dan kinderen die daaraan niet zijn blootgesteld. Deze verklaring biedt echter weinig inzicht in het precieze mechanisme waardoor deze kinderen een hogere sterftekans hebben. Verder blijkt dat korte geboorte-intervallen tussen de verschillende kinderen leiden tot fysieke uitputting van de moeder, en de aanwezigheid van veel jonge kinderen in het gezin samenhangt met sterkte competitie over schaarse hulpbronnen in het ouderlijke gezin. Deze factoren spelen vervolgens een rol in sterfte van jonge kinderen. Ook de opleiding van moeders, het huishoudinkomen en armoede spelen een belangrijke rol. Tot slot is er nog weinig onderzoek gedaan naar de rol van vaders, en naar sterfteclustering onder kinderen ouder dan een jaar.

In de overige drie papers in de dissertatie wordt dieper ingegaan op oorzaken voor kindersterfte op familieniveau, maar ook op de effecten van hoge kindersterfte in het gezin van herkomst op de verdere levensloop. In de eerste paper laat ik zien dat er inderdaad een sterkere concentratie is van zuigelingen- en kindersterfte in een klein aantal families dan het toeval voorspelt. Verder blijkt dat er inderdaad een significante intergenerationele samenhang bestaat in zuigelingen- en kindersterfte. Ik heb verschillende methodes gebruikt om vast te stellen dat sterfte clustert op het familieniveau, en te laten zien dat deze clustering geen product is van de gekozen methodologie. Gecontroleerd voor een groot aantal familiekenmerken blijft deze intergenerationele samenhang overeind. Zowel de familieachtergrond van de vader als van de moeder is hierbij belangrijk: als zij meer broertjes en zusjes verloren als kind, is de kans van hun eigen kinderen groter om te sterven in hun kindertijd. De rol van de moeder is echter groter dan de rol van de vader. Bij een nadere beschouwing van de intergenerationele samenhang en haar oorzaken, blijkt dat het gaat om een mix van verklarende mechanismen. Het gaat dan enerzijds om causale mechanismen —blootstelling heeft een onafhankelijk, causaal effect op sterfte in de volgende generatie—als anderzijds mechanismen die te maken hebben met kenmerken die samenhangen binnen families, zoals sociaaleconomische en gedragsfactoren. Met de tijd veranderde de sociaaleconomische structuur van de samenleving en nam de kennis over een goede hygiene en voeding voor zuigeling toe, en werd de medische zorg beter, waardoor de zuigelingen- en kindersterfte geleidelijk afnam. Steeds minder kinderen werden blootgesteld aan sterfte van hun broertjes en zusjes. Omdat sterfte van broertjes en zusjes het overleven van kinderen negatief beïnvloed kunnen we daarom verwachten dat de verminderde blootstelling aan ziekte en sterfte in het gezin leidde tot een betere gezondheid van de andere kinderen, waarmee de afname van kindersterfte zichzelf verder versterkte.

Andere oorzaken van zuigelingen- en babysterfte, waaronder een ongelijke behandeling van meisjes en jongens, kunnen daarnaast ook een rol hebben gespeeld in de (intergenerationele) concentratie van zuigelingensterfte en kindersterfte in families. Bevoordeling van jongens bij verdeling van voedsel, huishoudelijke taken en medische zorg kan hun gezondheid positief beïnvloeden, mogelijk ten koste van meisjes in het huishouden. Daarnaast kan selectieve verwaarlozing van meisjes een rol spelen. Omdat een dergelijke ongelijke behandeling van jongens en meisjes door socialisatie wordt overgedragen tussen generaties, zouden patronen van seksediscriminatie van generatie op generatie zichtbaar moeten zijn, waarbij meisjes van generatie op generatie in het nadeel zijn. Bewijs voor een dergelijk patroon is hier echter niet gevonden. In gezinnen waarbij de ouders zelf uit een gezin komen met een hoge sterfte sterven weliswaar meer kinderen, maar meisjes worden hierdoor niet sterker beïnvloed dan jongens. Verder blijkt dat sociaaleconomische status de stervenskansen van jonge kinderen beïnvloed. Hierbij gaat het niet alleen om de sociaaleconomische status van hun gezin van herkomst, maar ook om die van hun grootouders, waarbij vooral het nageslacht van boeren een voordeel heeft ten opzichte van personen met een andere sociaaleconomische achtergrond. Deze sociaaleconomische kenmerken verklaren echter niet waarom sterfte zich concentreert in een klein deel van de families in de populatie.

In de laatste paper ga ik nader in op de effecten van blootstelling aan hoge kindersterfte op overleven in de volwassen jaren. Het blijkt dat blootstelling aan zuigelingen- en kindersterfte in het gezin een bijdrage levert aan de relatieve sterftekansen op volwassen leeftijd. Dat geldt voor sterfte op verschillende volwassen leeftijden tussen 18 en 80 jaar oud. Daarnaast zien we dat vooral de sterfte aan hartziekten en aan diabetes beïnvloed wordt door blootstelling aan sterfte in de kindertijd. Deze bevinding komt overeen met theorieën uit de literatuur die voorspellen dat blootstelling aan ziekte in het vroege leven samenhangt met een overactief immuunsysteem op volwassen leeftijd, wat vervolgens kan leiden tot cardiologische problemen en diabetes. Voorts zijn de effecten van blootstelling binnen het gezin sterker dan de effecten van blootstelling in de regio van herkomst. Deze bevinding illustreert dat het gezin van grote invloed is de gezondheid van gezinsleden, zowel in de kindertijd als ook op volwassen leeftijd.

References

Adams, J.W., Kasakoff, A., & Kok, J. (2002). Migration over the life course in XIXth century Netherlands and the American North: A comparative analysis based on genealogies and population registers. Paper presented at the Annales de démographie historique.

Alam, N. (1995). Birth spacing and infant and early childhood mortality in a high fertility area of Bangladesh: Age-dependent and interactive effects. Journal of Biosocial Science, 27(4), 393–404. doi:10.1017/S0021932000023002

Alam, N., & David, P.H. (1998). Infant and child mortality in Bangladesh: Age-specific effects of previous child's death. Journal of Biosocial Science, 30(3), 333–348. doi:10.1017/S0021932098003332

Almond, D. (2006). Is the 1918 influenza pandemic over? Long-term effects of *in utero* influenza exposure in the post-1940 US population. Journal of Political Economy, 114(4), 672–712.

Almond, D., & Currie, J. (2011). Killing me softly: The fetal origins hypothesis. The Journal of Economic Perspectives, 153–172.

Alter, G., & Mandemakers, K. (2014). The Intermediate Data Structure (IDS) for longitudinal historical microdata, version 4. Historical Life Course Studies, 1, 1–26. doi:http://www.ehps-net.eu/article/intermediate-data-structure-ids-longitudinal-historical-microdata-version-4

Alter, G., Mandemakers, K., & Gutmann, M.P. (2009). Defining and distributing longitudinal historical data in a general way through an intermediate structure. Historical Social Research/Historische Sozialforschung, 34(3), 78–114. doi:10.12759/hsr.34.2009.3.78-114

Alter, G., & Oris, M. (2005). Childhood conditions, migration, and mortality: Migrants and natives in 19th–century cities. Social Biology, 52(3–4), 178–191. doi:10.1080/19485565.2005.9989108

Alter, G., Oris, M., & Broström, G. (2001). The family and mortality: A case study from rural Belgium. Annales de démographie historique, 101(1), 11–31. doi:10.3917/adh.101.0010

Andersson, T., Hogberg, U., & Åkerman, S. (1996). Survival of orphans in 19th century Sweden—the importance of remarriages. Acta Paediatrica, 85(8), 981–985. doi:10.1111/j.1651-2227.1996.tb14198.x

Anderton, D.L., Tsuya, N.O., Bean, L.L., & Mineau, G.P. (1987). Intergenerational transmission of relative fertility and life course patterns. Demography, 24(4), 467–480. doi:10.2307/2061386

Antonovsky, A., & Bernstein, J. (1967). Social class and infant mortality. Social Science & Medicine, 11(8–9), 453–470. doi:10.1016/0037-7856(77)90022-1

Arulampalam, W., & Bhalotra, S. (2006). Sibling death clustering in India: State dependence versus unobserved heterogeneity. Journal of the Royal Statistical Society: Series A (Statistics in Society), 169(4), 829–848. doi:10.1111/j.1467-985X.2006.00425.x

Arulampalam, W., & Bhalotra, S. (2008). The linked survival prospects of siblings: Evidence for the Indian states. Population Studies, 62(2), 171–190. doi:10.1080/00324720802000176

Barker, D.J., Eriksson, J.G., Forsen, T., & Osmond, C. (2002). Fetal origins of adult disease: Strength of effects and biological basis. International Journal of Epidemiology, 31(6), 1235–1239. doi:10.1093/ije/31.6.1235

Beekink, E., Boonstra, O., Engelen, T., & Knippenberg, H. (2003). Nederland in verandering: Maatschappelijke ontwikkelingen in kaart gebracht. Amsterdam: Aksant.

Beekink, E., Van Poppel, F., & Liefbroer, A.C. (1999). Surviving the loss of the parent in a nineteenth-century Dutch provincial town. Journal of Social History, 32(3), 641–669. doi: 10.1353/jsh/32.3.641

Beise, J., & Voland, E. (2002). A multilevel event history analysis of the effects of grandmothers on child mortality in a historical German population (Krummhörn, Ostfriesland, 1720–1874). Demographic Research, 7(13), 469–498. doi:10.4054/DemRes.2002.7.13

Bengtsson, T., & Lindström, M. (2000). Childhood misery and disease in later life: The effects on mortality in old age of hazards experienced in early life, southern Sweden, 1760–1894. Population Studies, 54(3), 263–277. doi:10.1080/713779096

Bengtsson, T., & Lindström, M. (2003). Airborne infectious diseases during infancy and mortality in later life in southern Sweden, 1766–1894. International Journal of Epidemiology, 32(2), 286–294. doi:10.1093/ije/dyg061

Bengtsson, T., & Van Poppel, F. (2011). Socioeconomic inequalities in death from past to present: An introduction. Explorations in Economic History, 48(3), 343–356. doi:10.1016/j.eeh.2011.05.004

Bhalotra, S., & Van Soest, A. (2008). Birth-spacing, fertility and neonatal mortality in India: Dynamics, frailty, and fecundity. Journal of Econometrics, 143(2), 274–290. doi:10.1016/j.jeconom.2007.10.005

Blom, J.C., & Lamberts, E. (2006). Geschiedenis van de Nederlanden. Amsterdam: Bert Bakker.

Bolstad, W.M., & Manda, S.O. (2001). Investigating child mortality in Malawi using family and community random effects: A Bayesian analysis. Journal of the American Statistical Association, 96(453), 12–19. doi:10.1198/016214501750332659

Bongaarts, J., & Guilmoto, C.Z. (2015). How many more missing women? Excess female mortality and prenatal sex selection, 1970–2050. Population and Development Review, 41(2), 241–269. doi:10.1111/j.1728-4457.2015.00046.x

Brändström, A. (1984). "De kärlekslösa mödrarna": Spädbarnsdödligheten i Sverige under 1800-talet med särskild hänsyn till Nedertorneå: Universitetet i Umeå.

Bras, H., & Neven, M. (2007). The effects of siblings on the migration of women in two rural areas of Belgium and the Netherlands, 1829–1940. Population Studies, 61(1), 53–71. doi:10.1080/00324720601048319

Bras, H., Van Bavel, J., & Mandemakers, K. (2013). Unraveling the intergenerational transmission of fertility: genetic and shared-environment effects during the demographic transition in the Netherlands, 1810–1910. The History of the Family, 18(2), 116–134. doi:10.1080/1081602X.2013.803491

Breschi, M., & Manfredini, M. (2002). Parental loss and kin networks: Demographic repercussions in a rural Italian village. Bern: Peter Lang.

Broström, G., Edvinsson, S., & Engberg, E. (2018). Intergenerational transfers of infant mortality in 19th century northern Sweden. Historical Life Course Studies. doi:http://hdl.handle.net/10622/23526343-2018-0005?locatt=view:master

Bruckner, T.A., & Catalano, R.A. (2009). Infant mortality and diminished entelechy in three European countries. Social Science & Medicine, 68(9), 1617–1624. doi:10.1016/j.socscimed.2009.02.011

Burgerlijke Stand Gemeente Nieuwvliet [Civil Registry Municipality of Nieuwvliet], Zeeuws Archief [Zeeland Archives] (2018). Civil Certificates of Death 1796–1970.

Caruso, C., Candore, G., Colonna-Romano, G., Lio, D., & Franceschi, C. (2005). Inflammation and life-span. Science, 307(5707), 208–209. doi:10.1126/science.307.5707.208

Carvalho, L. (2012). Childhood circumstances and the intergenerational transmission of socioeconomic status. Demography, 49(3), 913–938. doi:10.1007/s13524-012-0120-1

Centraal Bureau voor Geneaologie (2019). Uittreksels PK/PL's. Retrieved from cbg.nl/diensten/uittreksels-pkpl.

Chisholm, J.S. (1993). Death, hope, and sex: Life-history theory and the development of reproductive strategies [and comments and reply]. Current Anthropology, 1–24. doi:jstor.org/stable/2743728

Chisholm, J.S., Quinlivan, J.A., Petersen, R.W., & Coall, D.A. (2005). Early stress predicts age at menarche and first birth, adult attachment, and expected lifespan. Human Nature, 16(3), 233–265. doi:10.1007/s12110-005-1009-0

Cohen, A.A., Tillinghast, J., & Canudas-Romo, V. (2010). No consistent effects of prenatal or neonatal exposure to Spanish flu on late-life mortality in 24 developed countries. Demographic Research, 22(20), 579. doi:10.4054/DemRes.2010.22.20

Costa, D.L. (2000). Understanding the twentieth-century decline in chronic conditions among older men. Demography, 37(1), 53–72. doi:10.2307/2648096

Crimmins, E.M., & Finch, C.E. (2006). Infection, inflammation, height, and longevity. Proceedings of the National Academy of Sciences of the United States of America, 103(2), 498–503. doi:10.1073/pnas.0501470103

Curtis, S.L., Diamond, I., & McDonald, J.W. (1993). Birth interval and family effects on postneonatal mortality in Brazil. Demography, 30(1), 33–43. doi:10.2307/2061861

Curtis, S.L., & Steele, F. (1996). Variations in familial neonatal mortality risks in four countries. Journal of Biosocial Science, 28(2), 141–159. doi:10.1017/S0021932000022203

Das Gupta, M. (1990). Death clustering, mothers' education and the determinants of child mortality in rural Punjab, India. Population Studies, 44(3), 489–505. doi:10.1080/0032472031000144866

Das Gupta, M. (1997). Socio-economic status and clustering of child deaths in rural Punjab. Population Studies, 51(2), 191–202. doi:10.1080/0032472031000149906

De Meere, J. (1982). Economische ontwikkeling en levensstandaard in Nederland gedurende de eerste helft van de negentiende eeuw: Aspecten en trends. Den Haag: M. Nijhoff.

Delger, H., & Kok, J. (1998). Bridegrooms and biases: a critical look at the study of intergenerational mobility on the basis of marriage certificates. Historical Methods: A Journal of Quantitative and Interdisciplinary History, 31(3), 113–121. doi:10.1080/01615449809601194

Derosas, R. (2002). Infant mortality broken into pieces: A case study on Venice in mid-nineteenth century. Società Italiana di Statistica, Atti della XLI Riunione Scientifica. Università di Milano-Bicocca, 5–7.

Devos, I. (2000). Te jong om te sterven: De levenskansen van meisjes in België omstreeks 1900. Tijdschrift voor sociale geschiedenis(1), 55–75. doi:http://hdl.handle.net/1854/LU-121876

Devos, I. (2006). Allemaal beestjes: Mortaliteit en morbiditeit in Vlaanderen, 18de–20ste eeuw. Gent: Academia Press.

Doblhammer, G. (2003). Commentary: Infectious diseases during infancy and mortality in later life. International Journal of Epidemiology, 32(2), 294–295. doi:10.1093/ije/dyg066

Doblhammer, G., Van den Berg, G., & Lumey, L. (2013). A re-analysis of the long-term effects on life expectancy of the Great Finnish Famine of 1866–68. Population Studies, 67(3), 309–322. doi:10.1080/00324728.2013.809140

Donrovich, R., Puschmann, P., & Matthijs, K. (2018). Mortality clustering in the family: Fast life history trajectories and the intergenerational transfer of infant death in late 19th- and early 20th-century Antwerp, Belgium. Historical Life Course Studies. doi:http://hdl.handle.net/10622/23526343-2018-0006?locatt=view:master

Edvinsson, S., Brändström, A., Rogers, J., & Broström, G. (2005). High-risk families: The unequal distribution of infant mortality in nineteenth-century Sweden. Population Studies, 59(3), 321–337. doi:10.1080/00324720500223344

Edvinsson, S., & Janssens, A. (2012). Clustering of Deaths in Families: Infant and Child Mortality in Historical Perspective. Biodemography and Social Biology, 58(2), 75–86. doi:10.1080/19485565.2012.738575

Ekamper, P., Van Duin, C., Van Poppel, F., & Mandemakers, K. (2010). Heat waves and cold spells and their effect on mortality: An analysis of micro-data for the Netherlands in the nineteenth and twentieth centuries. Annales de démographie historique, 120(2), 55–104. doi:10.3917/adh.120.0055

Ekamper, Van Poppel, F. & Mandemakers, K. (2011). Widening horizons? The geography of the marriage market in the nineteenth and early-twentieth century Netherlands. In E.R. Merchant, G.D. Deane, M.P. Gutmann & K.M. Sylvester (Eds.): Navigating time and space in population studies (pp. 115–160). Dordrecht: Springer.

Finch, C.E., & Crimmins, E.M. (2004). Inflammatory exposure and historical changes in human life-spans. Science, 305(5691), 1736–1739. doi:10.1126/science.1092556

Floud, R., Fogel, R.W., Harris, B., & Hong, S.C. (2011). The changing body: Health, nutrition, and human development in the Western world since 1700. Cambridge: Cambridge University Press.

Fogel, R.W. (2004). The escape from hunger and premature death, 1700–2100: Europe, America, and the Third World (Vol. 38). Cambridge: Cambridge University Press.

Fogel, R.W., & Costa, D. L. (1997). A theory of technophysio evolution, with some implications for forecasting population, health care costs, and pension costs. Demography, 34(1), 49–66.

Fridlizius, G. (1989). The deformation of cohorts: Nineteenth century mortality decline in a generational perspective. Scandinavian Economic History Review, 37(3), 3–17. doi:10.1080/03585522.1989.10408152

Gagnon, A., & Mazan, R. (2009). Does exposure to infectious diseases in infancy affect old-age mortality? Evidence from a pre-industrial population. Social Science & Medicine, 68(9), 1609–1616. doi:10.1016/j.socscimed.2009.02.008

Gagnon, A., Smith, K.R., Tremblay, M., Vézina, H., Paré, P.P., & Desjardins, B. (2009). Is there a trade–off between fertility and longevity? A comparative study of women from three large historical databases accounting for mortality selection. American Journal of Human Biology, 21(4), 533–540.

Garðarsdóttir, Ó. (2002). Saving the child: Regional, cultural and social aspects of the infant mortality decline in Iceland, 1770–920 (Vol. 19): Demographic Data Base, Umeå University.

Garibotti, G., Smith, K.R., Kerber, R.A., & Boucher, K.M. (2006). Longevity and correlated frailty in multigenerational families. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences, 61(12), 1253–1261. doi:10.1093/gerona/61.12.1253

Gavrilov, L.A., & Gavrilova, N.S. (2001). Biodemographic study of familial determinants of human longevity. Population: An English Selection, 197–221. doi:jstor.org/stable/3030266
Guo, G. (1993). Use of sibling data to estimate family mortality effects in Guatemala. Demography, 30(1), 15–32. doi:10.2307/2061860
Gyimah, S.O. (2009). Polygynous marital structure and child survivorship in sub-Saharan Africa: Some empirical evidence from Ghana. Social Science & Medicine, 68(2), 334–342. doi:10.1016/j.socscimed.2008.09.067
Hagen, E.H., Barrett, H.C., & Price, M.E. (2006). Do human parents face a quantity–quality tradeoff?: Evidence from a Shuar community. American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists, 130(3), 405–418.
Hammel, E.A. (2005). Chayanov revisited: A model for the economics of complex kin units. Proceedings of the National Academy of Sciences, 102(19), 7043–7046.
Hanson, H.A., Smith, K.R., & Hasstedt, S. (2013). Heritability of longevity and the role of early and mid-life environments. Paper presented at the SSHA, Chicago.
Hatton, T.J. (2011). Infant mortality and the health of survivors: Britain, 1910–50. The Economic History Review, 64(3), 951–972. doi:10.1111/j.1468-0289.2010.00572.x
Hatton, T.J., & Martin, R.M. (2010). Fertility decline and the heights of children in Britain, 1886–1938. Explorations in Economic History, 47(4), 505–519.
Hawkes, K., O'Connell, J.F., Jones, N.B., Alvarez, H., & Charnov, E.L. (1998). Grandmothering, menopause, and the evolution of human life histories. Proceedings of the National Academy of Sciences, 95(3), 1336–1339. doi:10.1073/pnas.95.3.1336
Hayward, A.D., Rigby, F.L., & Lummaa, V. (2016). Early-life disease exposure and associations with adult survival, cause of death, and reproductive success in preindustrial humans. Proceedings of the National Academy of Sciences, 113(32), 8951–8956. doi:10.1073/pnas.1519820113
Hobcraft, J., McDonald, J., & Rutstein, S. (1985). Demographic determinants of infant and early child mortality: A comparative analysis. Population Studies, 39(3), 363–385. doi:10.1080/00324720310001415 76
Hofstee, E. (1974). De demografische ontwikkeling van Nederland sinds 1800. In H. Heeren & P. van Praag (Eds.), Van nu tot nul: Bevolkingsgroei en bevolkingspolitiek in Nederland (pp. 36–75). Utrecht: Spectrum.
Hofstee, E. (1981). Korte demografische geschiedenis van Nederland van 1800 tot heden [Short demographic history of The Netherlands 1800–today]. Haarlem: Fibula-Van Dishoeck.
Holmberg, H., & Broström, G. (2012). On statistical methods for clustering: A case study on infant mortality, Northern Sweden, 1831–1890. Biodemography and Social Biology, 58(2), 173–184.
Hoogerhuis, O.W. (2003). Baren op Beveland. Vruchtbaarheid en zuigelingensterfte in Goes en omliggende dorpen gedurende de 19e eeuw [Giving birth on Beveland. Fecundity and infant mortality in Goes and surrounding villages during the 19th century]. Wageningen: Wageningen Universiteit.
HSN. (2010a). Historical Sample of the Netherlands (HSN): Data set civil certificate release 2010.01 [Data file and code book].
HSN. (2010b). Historical Sample of the Netherlands (HSN): Data set life courses release 2010.01 [Data file and code book].
Hussain, R., Bittles, A., & Sullivan, S. (2001). Consanguinity and early mortality in the Muslim populations of India and Pakistan. American Journal of Human Biology, 13(6), 777–787. doi:10.1002/ajhb.1124
IISG. (2018a). About HSN: Sampling. Retrieved from https://socialhistory.org/en/hsn/sampling
IISG. (2018b). LINKS Publications. Retrieved from https://socialhistory.org/en/hsn/links-publications
Ikamari, L. (2000). Sibling mortality correlation in Kenya. Journal of Biosocial Science, 32(2), 265–278. doi:10.1017/S0021932000002650
Janssens, A. (1993). Family and social change. The household as a process in an industrializing community. Cambridge, UK: Cambridge University Press.
Janssens, A., Messelink, M., & Need, A. (2010). Faulty genes or faulty parents? Gender, family and survival in early and late childhood in the Netherlands, 1860–1900. The History of the Family, 15(1), 91–108. doi:10.1016/j.hisfam.2010.01.005
Janssens, A., & Pelzer, B. (2012). Did factory girls make bad mothers? Women's labor market experience, motherhood, and children's mortality risks in the past. Biodemography and Social Biology, 58(2), 133–148. doi:10.1080/19485565.2012.720451
Kaplanis, J., Gordon, A., Shor, T., Weissbrod, O., Geiger, D., Wahl, M., . . . Gymrek, M. (2018). Quantitative analysis of population-scale family trees with millions of relatives. Science, 360(6385), 171–175. doi:10.1126/science.aam9309
Kerber, R.A., O'Brien, E., Smith, K.R., & Cawthon, R.M. (2001). Familial excess longevity in Utah genealogies. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences, 56(3), B130-B139. doi:10.1093/gerona/56.3.B130
Kippen, R. (2011). A pestilence stalks abroad': familial clustering of deaths during the Tasmanian scarlet fever, measles and influenza epidemics of 1852–1854. Genus, 67(2), 57–75.
Kippen, R., & Walters, S. (2012). Is sibling rivalry fatal? Siblings and mortality clustering. Journal of Interdisciplinary History, 42(4), 571–591. doi:10.1162/JINH_a_00305
Knigge, A. (2016). Beyond the parental generation: The influence of grandfathers and great-grandfathers on status attainment. Demography, 53(4), 1219–1244. doi:10.1007/s13524-016-0486-6
Knotter, A., & Meijer, A. (1995). De gemeentelijke bevolkingsregisters, 1850–1920. Den Haag: Instituut voor Nederlandse Geschiedenis.
Kok, J., Mandemakers, K., & Bras, H. (2009). Van geboortebank tot collaboratory: Een reflectie op twintig jaar dataverzameling en onderzoek met de HSN. Tijdschrift voor sociale en economische geschiedenis, 6(4), 3–36. doi:10.18352/tseg.468
Kok, J., Vandezande, M., & Mandemakers, K. (2011a). Household structure, resource allocation and child well-being. A comparison across family systems. Tijdschrift voor Sociale en Economische Geschiedenis/The Low Countries Journal of Social and Economic History, 8(4), 76–101. doi:10.18352/tseg.346
Kok, J., Vandezande, M., & Mandemakers, K. (2011b). Household structure, resource allocation and child well-being. A comparison across family systems. Journal of Family History, 25, 269–290. doi:hdl.handle.net/1765/89462
Kuate-Defo, B., & Diallo, K. (2002). Geography of child mortality clustering within African families. Health & Place, 8(2), 93–117. doi:10.1016/S1353-8292(01)00038-7
Kuyper, J. (1868). *Gemeente-atlas van Nederland*. Leeuwarden: Hugo Suringar.
Lalou, R. (1997). Endogenous mortality in New France: At the crossroads of natural and social selection. In A. Bideau, B. Desjardins, & H. P. Brignoli (Eds.), Infant and child mortality in the past (pp. 203–215). Oxford: Clarendon Press.
Lambert, P.S., Zijdeman, R.L., Van Leeuwen, M.H., Maas, I., & Prandy, K. (2013). The construction of HISCAM: A stratification scale based on social interactions for historical comparative research. Historical Methods: A Journal of Quantitative and Interdisciplinary History, 46(2), 77–89. doi:10.1080/01615440.2012.715569
Laslett, P. (1980). The bastardy prone sub-society. In P. Laslett, K. Oosterveen, & R. M. Smith (Eds.), Bastardy and its comparative history: Studies in the history of illegitimacy and marital nonconformism in Britain, France, Germany, Sweden, North America, Jamaica and Japan (pp. 217–246). Cambridge, MA: Harvard University Press.
Lassetter, J.H., & Callister, L. C. (2009). The impact of migration on the health of voluntary migrants in western societies: A review of the literature. Journal of Transcultural Nursing, 20(1), 93–104. doi:10.1177/1043659608325841
Last, M. (1992). The importance of extremes: The social implications of intra-household variation in child mortality. Social Science & Medicine, 35(6), 799–810. doi:10.1016/0277-9536(92)90080-A
Liefbroer, A.C., & Elzinga, C.H. (2012). Intergenerational transmission of behavioural patterns: How similar are parents' and children's demographic trajectories? Advances in Life Course Research, 17(1), 1–10. doi:10.1016/j.alcr.2012.01.002
Lindahl-Jacobsen, R., Hanson, H.A., Oksuzyan, A., Mineau, G.P., Christensen, K., & Smith, K.R. (2013). The male–female health-survival paradox and sex differences in cohort life expectancy in Utah, Denmark, and Sweden 1850–1910. Annals of Epidemiology, 23(4), 161–166. doi:10.1016/j.annepidem.2013.02.001
Lindeboom, M., Portrait, F., & Van den Berg, G.J. (2010). Long-run effects on longevity of a nutritional shock early in life: The Dutch Potato famine of 1846–1847. Journal of Health Economics, 29(5), 617–629. doi:10.1016/j.jhealeco.2010.06.001

Lindkvist, M., & Broström, G. (2006). Clustered infant mortality in an intergenerational perspective, Skellefteå 1831–1890. Paper presented at the SSHA, Minneapolis.
Lovejoy, C.O. (1981). The origin of man. Science, 211(4480), 341–350. doi:10.1126/science.211.4480.341
Lu, Y., & Qin, L. (2014). Healthy migrant and salmon bias hypotheses: A study of health and internal migration in China. Social Science & Medicine, 102, 41–48. doi:10.1016/j.socscimed.2013.11.040
Lundevaller, E., & Edvinsson, S. (2012). The effect of Rh-negative disease on perinatal mortality: Some evidence from the Skellefteå region, Sweden, 1860–1900. Biodemography and Social Biology, 58(2), 116–132. doi:10.1080/19485565.2012.720450
Lynch, K.A., & Greenhouse, J.B. (1994). Risk factors for infant mortality in nineteenth-century Sweden. Population Studies, 48(1), 117–133. doi:10.1080/0032472031000147506
Lyngstad, T.H., & Prskawetz, A. (2010). Do siblings' fertility decisions influence each other? Demography, 47(4), 923–934. doi:10.1007/BF03213733
Madise, N.J., & Diamond, I. (1995). Determinants of infant mortality in Malawi: An analysis to control for death clustering within families. Journal of Biosocial Science, 27(01), 95–106. doi:10.1017/S0021932000007033
Manda, S.O.M. (1998). Unobserved family and community effects on infant mortality in Malawi. Genus, 54(1/2), 143–164.
Mandemakers, K. (2000). Historical sample of the Netherlands. In P.K. Hall, R. McCaa, & G. Thorvaldsen (Eds.), Handbook of international historical microdata for population research (pp. 149–177). Minneapolis: Minnesota Population Center.
Mandemakers, K. (2002). Building life course datasets from population registers by the Historical Sample of the Netherlands (HSN). History and Computing, 14(1–2), 87–107. doi:10.3366/hac.2002.14.1-2.87
Mandemakers, K., Bloothooft, G. (2015). 'Record linkage with the LINKS project: From Marriage Certificates to Pedigrees' paper session Challenges, Methods and Solutions in Record Linkage of Micro-data 40th Social Science History Conference, 12–15 November 2015, Baltimore.
Mandemakers, K., & Laan, F. (2017a). LINKS dataset Genes Germs and Resources. WieWasWie Zeeland. Civil Certificates. Version 2017.01 [Data file and code book].
Mandemakers, K., & Laan, F. (2017b). LINKS dataset Genes Germs and Resources. WieWasWie Zeeland. Civil Certificates. version 2017.01. IDS version. [Data file and code book].
Mandemakers, K., Mourits, R.J., Muurling, S., Boter, C., Van Dijk, I.K., Maas, I., . . . Miles, A. (2018). HSN standardized, HISCO-coded and classified occupational titles, release 2018.01.
Mandemakers, K., Muurling, S., Maas, I., Van de Putte, B., Zijdeman, R.L., Lambert, P., . . . Miles, A. (2013). HSN standardized, HISCO-coded and classified occupational titles.
Massey, C.G. (2017). Playing with matches: An assessment of accuracy in linked historical data. Historical Methods: A Journal of Quantitative and Interdisciplinary History, 50(3), 129–143. doi:10.1080/01615440.2017.1288598
McEwen, B.S. (1998). Stress, adaptation, and disease: Allostasis and allostatic load. Annals of the New York Academy of Sciences, 840(1), 33–44. doi:10.1111/j.1749-6632.1998.tb09546.x
McEwen, B.S., & Stellar, E. (1993). Stress and the individual: Mechanisms leading to disease. Archives of Internal Medicine, 153(18), 2093–2101. doi:10.1001/archinte.1993.00410180039004
McFarland, M.J., Taylor, J., Hill, T.D., & Friedman, K.L. (2018). Stressful life events in early life and leukocyte telomere length in adulthood. Advances in Life Course Research, 35, 37–45.
McKeown, T. (1976). The modern rise of population, London: Edward Arnold.
McMurray, C. (1997). Measuring excess risk of child mortality: An exploration of DHS I for Burundi, Uganda and Zimbabwe. Journal of Biosocial Science, 29(1), 73–91. doi:10.1017/S0021932097000734
Miller, J.E., Trussell, J., Pebley, A.R., & Vaughan, B. (1992). Birth spacing and child mortality in Bangladesh and the Philippines. Demography, 29(2), 305–318. doi:10.2307/2061733
Mineau, G., Smith, K.R., & Bean, L. (2004). Adult mortality risks and religious affiliation: The role of social milieu in biodemographic studies. Annales de démographie historique, 108(2), 85–104. doi:10.3917/adh.108.0085
Mitchell, S.A. j. (Cartographer). (1868). A new universal atlas containing maps of the various empires, kingdoms, states and republics of the world.
Mosley, W.H., & Chen, L.C. (1984). An analytical framework for the study of child survival in developing countries. Population and Development Review, 10(s1), 25–45. doi:10.2307/2807954

Mourits, R. (2017). Later-life mortality and longevity in late-18th and 19th-century cohorts. Where are we now, and where are we heading? Historical Life Course Studies, 4, 1–19.
Murphy, M. (2013). Cross-national patterns of intergenerational continuities in childbearing in developed countries. Biodemography and Social Biology, 59(2), 101–126. doi:10.1080/19485565.2013.833779
Myntti, C. (1993). Social determinants of child health in Yemen. Social Science & Medicine, 37(2), 233–240. doi:10.1016/0277-9536(93)90457-F
Myrskylä, M. (2010). The relative effects of shocks in early–and later–life conditions on mortality. Population and Development Review, 36(4), 803–829. doi:10.1111/j.1728-4457.2010.00358.x
Nam, C.B., & Boyd, M. (2004). Occupational status in 2000; Over a century of census-based measurement. Population Research and Policy Review, 23(4), 327–358. doi:10.1023/B:POPU.0000040045.51228.34
Nault, F., Desjardins, B., & Légaré, J. (1990). Effects of reproductive behaviour on infant mortality of French-Canadians during the seventeenth and eighteenth centuries. Population Studies, 44(2), 273–285. doi:10.1080/0032472031000144596
Nettle, D., Coall, D.A., & Dickins, T.E. (2011). Early-life conditions and age at first pregnancy in British women. Proceedings of the Royal Society B: Biological Sciences, 278(1712), 1721–1727. doi:10.1098/rspb.2010.1726
Nieuwvliet, G. (1796–1970). Burgerlijke Stand gemeente Nieuwvliet 1796–1870 [Civil Registry Municipality of Nieuwvliet 1796–1870]. from Zeeuws Archief
Nonyane, B.A., Williams, E.K., Blauvelt, C., Shah, M.R., Darmstadt, G.L., Moulton, L., & Baqui, A.H. (2013). Clustering of neonatal deaths in Bangladesh: Results from the Projahnmo studies. Paediatric and Perinatal Epidemiology, 27(2), 165–171.
Norton, M.C., Hatch, D.J., Munger, R.G., & Smith, K.R. (2017). Family member deaths in childhood predict systemic inflammation in late life. Biodemography and Social Biology, 63(2), 104–115. doi:10.1080/19485565.2017.1281099
Öberg, S. (2015). The direct effect of exposure to disease in early life on the height of young adult men in southern Sweden, 1814–1948. Population Studies, 69(2), 179–199. doi:10.1080/00324728.2015.1045545
Öberg, S. (2018). Changing bodies: The role of nutritional status in the great mortality decline. Seminar presentation, Radboud University, 24-10-2018.
Omariba, D.W.R., Beaujot, R., & Rajulton, F. (2007). Determinants of infant and child mortality in Kenya: an analysis controlling for frailty effects. Population Research and Policy Review, 26(3), 299–321. doi:10.1007/s11113-007-9031-z
Omariba, D.W.R., Rajulton, F., & Beaujot, R. (2008). Correlated mortality of siblings in Kenya: The role of state dependence. Demographic Research, 18(11), 311–336. doi:10.4054/DemRes.2008.18.11
Pakot, L. (2015). Maternal health and infant mortality in rural Transylvania. A case study of Vlăniă and Căpâlniă, 1850–1939. Romanian Journal of Population Studies, 9(1).
Pavard, S., Gagnon, A., Desjardins, B., & Heyer, E. (2005). Mother's death and child survival: The case of early Quebec. Journal of Biosocial Science, 37(2), 209–227. doi:10.1017/S0021932004006571
Pelzer, B., & Janssens, A. (2014). Lovely little angels in Heaven? The influence of religiously determined cultural life scripts on infant survival in the Netherlands, 1880–1920. Historical Social Research, 39(1), 19-47.
Preston, S.H., & Haines, M.R. (2014). Fatal years: Child mortality in late nineteenth-century America. Princeton, NJ: Princeton University Press.
Preston, S.H., Hill, M.E., & Drevenstedt, G.L. (1998). Childhood conditions that predict survival to advanced ages among African-Americans. Social Science & Medicine, 47(9), 1231–1246. doi:10.1016/S0277-9536(98)00180-4
Priester, P.R. (1998). Geschiedenis van de Zeeuwse landbouw circa 1600–1910. Wageningen: Landbouwuniversiteit.
Puschmann, P., Donrovich, R., & Matthijs, K. (2017). Salmon bias or red herring? Human Nature, 28(4), 481–499. doi:10.1007/s12110-017-9303-1
Quaranta, L. (2013). Scarred for life. How conditions in early life affect socioeconomic status, reproduction and mortality in Southern Sweden, 1813–1968. Lund University,
Quaranta, L. (2014). Early life effects across the life course: The impact of individually defined exogenous measures of disease exposure on mortality by sex in 19th-and 20th-century Southern Sweden. Social Science & Medicine, 119, 266–273. doi:10.1016/j.socscimed.2014.04.007

Quaranta, L. (2016). STATA Programs for using the Intermediate Data Structure (IDS) to construct files for statistical analysis. Historical Life Course Studies, 3, 1–19. doi:http://hdl.handle.net/10622/23526343-2016-0001?locatt=view:master

Quaranta, L. (2018). Intergenerational transfers in infant mortality in southern Sweden, 1740–1968. Historical Life Course Studies. doi:http://hdl.handle.net/10622/23526343-2018-00013?locatt=view:master

Quaranta, L. (2018). Program for studying intergenerational transmissions in infant mortality using the Intermediate Data Structure (IDS). Historical Life Course Studies.

Quaranta, L., Broström, G., Van Dijk, I.K., Donrovich, R., Edvinsson, S., Engberg, E., . . . Sommerseth, H. L. (2017). Intergenerational transfers of infant mortality in historical contexts: a comparative study of five European populations. Paper presented at the PAA, Chicago.

Ranjan, M., Dwivedi, L.K., & Mishra, R. (2018). Caste differentials in death clustering in entral and eastern Indian states. Journal of Biosocial Science, 50(2), 254–274. doi:10.1017/S0021932017000219

Rao, S.R., Pandey, A., & Shajy, K. (1997). Child mortality in Goa: A cross-sectional analysis. Social Biology, 44(1–2), 101–110.

Reher, D.S. (2001). In search of the "urban penalty": Exploring urban and rural mortality patterns in Spain during the demographic transition. International Journal of Population Geography, 7(2), 105–127.

Reid, A. (2001). Neonatal mortality and stillbirths in early twentieth century Derbyshire, England. Population Studies, 55(3), 213–232. doi:10.1080/00324720127696

Reid, A. (2002). Infant feeding and post-neonatal mortality in Derbyshire, England, in the early twentieth century. Population Studies, 56(2), 151–166. doi:10.1080/00324720215926

Risse, G.B. (1997). Cause of death as a historical problem. Continuity and Change, 12(2), 175–188.

Riswick, T. (2018). Testing the conditional resource-dilution hypothesis: The impact of sibship size and composition on infant and child mortality in the Netherlands, 1863–1910. The History of the Family, 1–33. doi:10.1080/1081602X.2018.1532310

Riswick, T., & Engelen, T. (2018). Siblings and life transitions: Investigating the resource dilution hypothesis across historical contexts and outcomes. History of the Family, 23(4), 521–532. doi: 10.1080/1081602X.2018.1532309

Ronsmans, C. (1995). Patterns of clustering of child mortality in a rural area of Senegal. Population Studies, 49(3), 443–461.

Rosenbaum-Feldbrügge, M. (2018). Gender differences in response to family crisis: Changes in household composition and migration of widowed parents with minor children in the Netherlands, 1863–1910. The History of the Family, 1–27. doi:10.1080/1081602X.2018.1537881

Ruggles, S. (1992). Migration, marriage, and mortality: Correcting sources of bias in English family reconstitutions. Population Studies, 46(3), 507–522. doi:10.1080/0032472031000146486

Ruggles, S., Fitch, C.A., & Roberts, E. (2018). Historical census record linkage. Annual Review of Sociology(44), 19–37. doi:10.1146/annurev-soc-073117-041447

Saha, U.R., & van Soest, A. (2011). Infant death clustering in families: Magnitude, causes, and the influence of better health services, Bangladesh 1982–2005. Population Studies, 65(3), 273–287. doi:10.1080/00324728.2011.602100

Saha, U.R., van Soest, A., & Bijwaard, G.E. (2014). Cause-specific neonatal deaths in rural Bangladesh, 1987–2005: Levels, trends, and determinants. Population Studies, 68(3), 247–263. doi:10.1080/00324728.2014.902094

Sastry, N. (1997). Family-level clustering of childhood mortality risk in northeast Brazil. Population Studies, 51(3), 245–261. doi:10.1080/0032472031000150036

Scalone, F., Agati, P., Angeli, A., & Donno, A. (2017). Exploring unobserved heterogeneity in perinatal and neonatal mortality risks: The case of an Italian sharecropping community, 1900–39. Population Studies, 71(1), 23–41. doi:10.1080/00324728.2016.1254812

Schellekens, J., & Van Poppel, F. (2012). Marital fertility decline in the Netherlands: Child mortality, real wages, and unemployment, 1860–1939. Demography, 49(3), 965–988. doi:10.1007/s13524-012-0112-1

Schellekens, J., & Van Poppel, F. (2016). Early-life conditions and adult mortality decline in Dutch cohorts born 1812–1921. Population Studies, 70(3), 327–343. doi:10.1080/00324728.2016.1223336

Sear, R., & Coall, D. (2011). How much does family matter? Cooperative breeding and the demographic transition. Population and Development Review, 37(s1), 81–112.

Sear, R., & Mace, R. (2008). Who keeps children alive? A review of the effects of kin on child survival. Evolution and Human Behavior, 29(1), 1–18. doi:10.1016/j.evolhumbehav.2007.10.001

Sear, R., Steele, F., McGregor, I.A., & Mace, R. (2002). The effects of kin on child mortality in rural Gambia. Demography, 39(1), 43–63. doi:10.1353/dem.2002.0010

Shah, B.D., & Dwivedi, L.K. (2011). Causes of neonatal deaths among tribal women in Gujarat, India. Population Research and Policy Review, 30(4), 517–536. doi:10.1007/s11113-010-9199-5

Smith, K.R., Gagnon, A., Cawthon, R.M., Mineau, G.P., Mazan, R., & Desjardins, B. (2009). Familial aggregation of survival and late female reproduction. Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences, 64(7), 740–744. doi:10.1093/gerona/glp055

Smith, K.R., Mineau, G.P., Garibotti, G., & Kerber, R. (2009). Effects of childhood and middle-adulthood family conditions on later-life mortality: Evidence from the Utah Population Database, 1850–2002. Social Science & Medicine, 68(9), 1649–1658. doi:10.1016/j.socscimed.2009.02.010

Sommerseth, H.L. (2018). Intergenerational transfer in infant mortality in northern Norway during the 19th- and early 20th-century. Historical Life Course Studies. doi:http://hdl.handle.net/10622/23526343-2018-0008?-locatt=view%3Amaster

Stadsarchief Amsterdam. (2018). Handleiding Gezinskaarten 1893–1939. Retrieved from https://archief.amsterdam/indexen/gezinskaarten_1893-1939/handleiding/index.nl.html

Störmer, C., & Lummaa, V. (2014). Increased mortality exposure within the family rather than individual mortality experiences triggers faster life-history strategies in historic human populations. PloS One, 9(1), e83633. doi:10.1371/journal.pone.0083633

Störmer, C., & Voland, E. (2014). Lebensgeschichtsevolution–Variation von Lebensstrategien evolutionär erklären. Darwin heute–Evolution als Leitbild in den modernen Wissenschaften, 171–201.

Störmer, C., & Willführ, K. P. (2010). Mortality crises and their consequences for human life histories. In Homo Novus—A human without illusions (pp. 95–109). Berlin Heidelberg: Springer.

Stradford, L., Van Poppel, F., & Lumey, L. (2017). Can resource dilution explain differences in height by birth order and family size? A study of 389,287 male recruits in twentieth-century Netherlands. The History of the Family, 22(2–3), 214–235.

Turra, C.M., & Elo, I.T. (2008). The impact of salmon bias on the Hispanic mortality advantage: New evidence from social security data. Population Research and Policy Review, 27(5), 515. doi:10.1007/s11113-008-9087-4

UNICEF, The World Bank, and the United Nations Population Division. (2014). Levels and Trends in Child Mortality 2014. New York: Unicef.

Van Bodegom, D., Eriksson, U.K., Houwing-Duistermaat, J.J., & Westendorp, R.G. (2012). Clustering of child mortality in a contemporary polygamous population in Africa. Biodemography and Social Biology, 58(2), 162–172. doi:10.1080/19485565.2012.720445

Van Cruyningen, P. (2005). Vrouwenarbeid in de Zeeuwse landbouw in de achttiende eeuw. Tijdschrift voor sociale en economische geschiedenis, 2(3), 43.

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

Van den Berg, N., Van Dijk, I.K., Mourits, R.J., Janssens, A., Slagboom, P.E., & Mandemakers, K. (2018). Families in comparison: An individual-level comparison of life course and family reconstructions between population and vital event registers. SocArXiv. doi:https://10.31235/osf.io/h2w8t

Van der Kaaden, J. (2003). Geschiedenis van de inheemse malaria in Nederland. Infectieziekten Bulletin, 10, 388–393.

Van Dijk, I.K. (2018). Early-life mortality clustering in families: A literature review. Population Studies, 1–21. doi:10.1080/00324728.2018.1448434

Van Dijk, I.K., & Mandemakers, K. (2018). Like mother, like daughter. Intergenerational transmission of infant mortality clustering in Zeeland, the Netherlands, 1833–1912. Historical Life Course Studies. doi:http://hdl.handle.net/2066/187251

Van Leeuwen, M.H., & Maas, I. (2011). HISCLASS: A historical international social class scheme. Leuven: University Press.

Van Poppel, F. (2000). Long-term trends in relative health differences between men and women. In: European Journal of Obstetrics & Gynecology and Reproductive Biology, 93(2), 119-122

Van Poppel, F., Bijwaard, G.E., Ekamper, P., & Mandemakers, K. (2012). Historical Trends in the Correlation of Sibling Deaths in Infancy in The Netherlands. Biodemography and Social Biology, 58(2), 87–115. doi:10.1080/19485565.2012.720448

Van Poppel, F., Jonker, M., & Mandemakers, K. (2005). Differential infant and child mortality in three Dutch regions, 1812–1909. The Economic History Review, 58(2), 272–309. doi:10.1111/j.1468-0289.2005.00305.x

Van Poppel, F., & Mandemakers, K. (2002). Sociaal-economische verschillen in zuigelingen-en kindersterfte in Nederland, 1812–1912 [Socioeconomic differences in infant and child mortality in The Netherlands, 1812–1912]. Bevolking en gezin, 31(2), 5–40.

Van Poppel, F., Monden, C., & Mandemakers, K. (2008). Marriage timing over the generations. Human Nature, 19(1), 7–22. doi:10.1007/s12110-008-9033-5

Van Poppel, F., Schellekens, J., & Walhout, E. (2009). Oversterfte van jonge meisjes in Nederland in de negentiende en eerste helft twintigste eeuw. Tijdschrift voor sociale en economische geschiedenis, 6.

Vandezande, M. (2012). Born to die. Death clustering and the intergenerational transmission of infant mortality, the Antwerp district, 1846–1905. Leuven: Centrum voor Sociologisch Onderzoek.

Vandezande, M. (2013). Using death clustering in an attempt to reassess the impact of men on infant mortality. WOG Working paper, 2013(2).

Vandezande, M., & Matthijs, K. (2013). Inherited dimensions of infant mortality. Detecting signs of disproportionate mortality risks in successive generations. The History of the Family, 18(2), 169–186. doi:10.1080/1081602X.2012.755131

Vandezande, M., Moreels, S., & Matthijs, K. (2010). Explaining death clustering. Intergenerational patterns in infant mortality, Antwerp, 1846–1905. Working paper of the Scientific Research Community Historical Demography, 13.

Vasto, S., Candore, G., Balistreri, C. R., Caruso, M., Colonna-Romano, G., Grimaldi, M. P., . . . Caruso, C. (2007). Inflammatory networks in ageing, age-related diseases and longevity. Mechanisms of Ageing and Development, 128(1), 83–91. doi:10.1016/j.mad.2006.11.015

Vaupel, J.W. (1988). Inherited frailty and longevity. Demography, 25(2), 277–287. doi:10.2307/2061294

Vaupel, J.W., Manton, K.G., & Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography, 16(3), 439–454. doi:10.2307/2061224

Voland, E. (2000). Contributions of family reconstitution studies to evolutionary reproductive ecology. Evolutionary Anthropology: Issues, News, and Reviews, 9(3), 134–146. doi:10.1002/1520-6505(2000)9:3<134::AID-EVAN3>3.0.CO;2-M

Vulsma, R. (1988). Burgerlijke stand en bevolkingsregister [Civil Register and Population Register]. Den Haag: Centraal Bureau voor Genealogie.

Webster, G.D., Graber, J.A., Gesselman, A.N., Crosier, B.S., & Schember, T.O. (2014). A life history theory of father absence and menarche: A meta-analysis. Evolutionary Psychology, 12(2), 273–294. doi:10.1177/147470491401200202

Willführ, K.P., & Gagnon, A. (2012). Are stepmothers evil or simply unskilled? Infant death clustering in recomposed families. Biodemography and Social Biology, 58(2), 149–161. doi:10.1080/19485565.2012.734745

Wisselgren, M.J., Edvinsson, S., Berggren, M., & Larsson, M. (2014). Testing methods of record linkage on Swedish censuses. Historical Methods: A Journal of Quantitative and Interdisciplinary History, 47(3), 138–151. doi:10.1080/01615440.2014.913967

Wolfinger, N.H. (2011). More evidence for trends in the intergenerational transmission of divorce: A completed cohort approach using data from the general social survey. Demography, 48(2), 581–592. doi:10.1007/s13524-011-0025-4

Wolleswinkel-Van den Bosch, J.H. (1998). The epidemiological transition in the Netherlands. Rotterdam: Erasmus University.

Woods, R. (2003). Urban–rural mortality differentials: An unresolved debate. Population and Development Review, 29(1), 29–46.

Yashin, A.I., Vaupel, J.W., & Iachine, I.A. (1995). Correlated individual frailty: An advantageous approach to survival analysis of bivariate data. Mathematical Population Studies, 5(2), 145–159. doi:10.1080/08898489509525394

Yu, Y., Liew, Z., Cnattingius, S., Olsen, J., Vestergaard, M., Fu, B., . . . Li, J. (2017). Association of mortality with the death of a sibling in childhood. Jama Pediatrics. doi:10.1001/jamapediatrics.2017.0197

Zaba, B., & David, P.H. (1996). Fertility and the distribution of child mortality risk among women: An illustrative analysis. Population Studies, 50(2), 263–278. doi:10.1080/0032472031000149346

Zenger, E. (1993). Siblings' neonatal mortality risks and birth spacing in Bangladesh. Demography, 30(3), 477–488. doi:10.2307/2061652

Zhao, Z. (1994). Demographic conditions and multi-generation households in Chinese history. Results from genealogical research and microsimulation. Population Studies, 48(3), 413–425. doi:10.1080/0032472031000147946

Zhao, Z. (2001). Chinese genealogies as a source for demographic research: A further assessment of their reliability and biases. Population Studies, 55(2), 181–193. doi:10.1080/00324720127690

Zijdeman, R.L. (2009). Like my father before me: Intergenerational occupational status transfer during industrialization (Zeeland, 1811–1915). Continuity and Change, 24(03), 455–486. doi:10.1017/S0268416009990166

List of Abbreviations

CI	confidence interval
CMR	child mortality rate
HSN	Historical Sample of the Netherlands
ICD	International Classification of Diseases
IDS	Intermediate Data Structure
IMR	infant mortality rate
LINKS	LINKing System for historical family reconstruction
RP	research person
RR	relative risk
SES	socioeconomic status
UPDB	Utah Population Database
US	United States
WHO	World Health Organization

List of Figures and Boxes

Part I

Figure 1. Mechanisms that link together sibling mortality and the mortality risk of ego and ego's children 16
Box 1. An overview of the hypotheses on exposure to mortality in families and effects across life courses and generations 28
Figure 2. Figure 2. A map of the province of Zeeland, 1866 31
Figure 3. Stillbirths and infant and child mortality in Zeeland, 1833–1912 32
Figure 4. The clustering of child deaths in families with three, five, seven, nine, 11, and 14 children in Zeeland, 1833–1912 34
Figure 5. Examples of death certificates 36
Figure 6. A map of the state of Utah, 1868 43
Figure 7. Infant and child mortality in Utah, 1874–1929 44
Figure 8. The effect of maternal infant deaths on the death risk of her daughter's infant children in five historical European populations 51

Part II

Paper I

Figure 1. The data structure for HSN and LINKS 81
Figure 2. Matching for the number of siblings and children between the HSN and LINKS: more in LINKS (green) or HSN (blue) 85
Figure 3. A comparison between HSN and LINKS for mortality data 86
Figure 4. The HISCLASS score for research persons in HSN and LINKS 87
Figure A.1. The HISCLASS score for research persons in HSN and LINKS 98

Paper II

-

Paper III

Figure 1. Multigenerational influences on infant mortality 140
Figure 2. IMR Zeeland population and sample LINKS, 1833–1912 141
Figure 3. The expected and observed numbers of families with a given number of deaths, by parity 146
Figure 4. The cumulative hazard by number of infant deaths of the grandmother 147

Paper IV
Box 1. An overview of the hypotheses on exposure to mortality in the second and third generations 165
Figure 1. The child mortality rate (CMR), the infant mortality rate (IMR), and the stillbirth rate: (a) for the province of Zeeland, 1833–1912; (b) for the State of Utah, 1874–1929 171
Figure 2. The design of the study 172
Figure 3. The distribution of parental sibling death experience by parental age (panel A) and occurrence of parental sibling deaths by parental age (panel B) in Zeeland (left) and Utah (right) 175
Figure 4. The effect sizes of paternal and maternal exposure to sibling deaths on offspring survival by age of exposure of the parent and number of experienced deaths, Zeeland (upper panel) and Utah (lower panel) 183
Figure 5. The effect of exposure to three or more sibling deaths between age 6 and age 12 and offspring survival between age 0 and age 5, Zeeland 186
Figure A.1. Variability within sibling groups in exposure to sibling deaths 193

Paper V
Figure 1. Causes of death by number of sibling deaths, men (left) and women (right) 212

List of Tables

Part I
Table 1. An overview of the numbers of cases in the LINKS data files 39
Table 2. An overview of the generations and dates of birth, marriage, or death 39
Table 3. An overview of succesful linking of civil certificates by cohort 40

Part II
Paper I
Table 1. The expected availability of demographic indicators in the HSN and LINKS 79
Table 2. The available family ties in the HSN and LINKS for the selected 495 RPs from the 1862–71 Zeeland cohort 82
Table 3. The number of available cases and the mean score on demographic indicators after selection of best cases in the HSN and LINKS separately (A) and simultaneously (B) 83

Paper II
Table 1. An overview of the methodological approaches to mortality clustering 106
Table 2. An overview of publications on early-life mortality clustering 110

Paper III
Table 1. The descriptive statistics 143
Table 2. The effects of the grandmother's infant deaths on infant survival, Zeeland, 1833–1912 149
Table 3. The effect of the grandmother's infant deaths on infant survival, extended models, Zeeland, 1833–1912: Cox models with mother shared frailty 151

Paper IV
Table 1. The descriptive statistics 173
Table 2. The clustered model: the effects of mortality in the sibling group and the relative death risk of offspring, children 0–5 180
Table 3. The fixed-effects model: parental exposure to sibling mortality and child mortality 0–5 185
Table A.1. The selection of cases 192

Paper V
Table 1. The descriptive statistics 202
Table 2. Childhood community levels of mortality among infants, young, and older children and survival between ages 18 and 85 in Utah and Zeeland, relative risk, Cox regression 206
Table 3. Sibling deaths in childhood and survival between ages 18 and 85 208
Table 4. Test of trend for exposure to sibling deaths in childhood and adult mortality, ages 18–85 210
Table 5. Cause-specific mortality age 18–85 by sibling death due to infection, Utah 213

List of Publications

This doctoral thesis is based on the following journal articles:

I. Van Dijk, I.K.*, Mourits, R.*, Van den Berg, N.*, Slagboom, P.E., Janssens, A.A.P.O., & Mandemakers, K. (revised and resubmitted). Families in Comparison: An individual-level comparison of life course and family reconstructions between population and vital event registers. *Shared first authorship.
II. Van Dijk, I.K. (2019). Early-life mortality clustering in families: A literature review. Population Studies, 72(1), 79–99.
III. Van Dijk, I.K. & Mandemakers, K. (2018). Like mother, like daughter. Intergenerational Transmission of Mortality Clustering in Zeeland, The Netherlands, 1833–1912. Historical Life Course Studies, 7(2), 28–46 . Special Issue on Intergenerational transmissions of infant mortality using the Intermediate Data Structure (IDS). Editors: Luciana Quaranta and Hilde L. Sommerseth.
IV. Van Dijk, I.K. (under review). Repeating a Lethal Experience: Mother's and Father's Exposure to Sibling Mortality and Offspring Mortality.
V. Van Dijk, I.K., Janssens, A.A.P.O., & Smith, K.R. (2018). The long harm of childhood? Childhood exposure to mortality and subsequent risk of adult mortality in Utah and The Netherlands. European Journal of Population (online first).

Other publications that have resulted from this doctoral project:

VI. Mandemakers, K., Mourits, R.J., Muurling, S., Boter, C., **Van Dijk, I.K.**, Maas, I., Van de Putte, B., Zijdeman, R.L., Lambert, P., Van Leeuwen, M.H.D., Van Poppel, F., & Miles, A. (2018). HSN standardized, HISCO-coded and classified occupational titles, release 2018.01.
VII. Van den Berg, N., Rodriguez-Girondo, M., **Van Dijk, I.K**., Mourits, R., Janssens, A., Mandemakers, K., Beekman, M., Smith, K.R.*, & Slagboom, P.E.* (2019). Longevity among top 15% survivors is an additive quantitative trait: a novel study of large 3-generation Dutch and US datasets. Nature Communications, 10(35). *Shared last authorship.

Funding and acknowledgments

This work was supported by the grant "Genes, Germs and Resources" (project number 360-53-180) awarded to prof. dr. A. Janssens, financed by the Netherlands Organisation for Scientific Research (NWO). A three-month research visit to Population Sciences, Huntsman Cancer Institute, the University of Utah (Salt Lake City, Utah, US) was supported by a teaching and internationalization grant of the Radboud University Nijmegen, the professor van winterfonds, and a Prins Bernhard Cultuurfonds grant.

Work on the UPDB was supported by the National Institutes of Health (grant number AG022095, Early Life Conditions, Survival and Health; Smith PI). I thank the Pedigree and Population Resource of the Huntsman Cancer Institute, University of Utah (funded in part by the Huntsman Cancer Foundation) for its role in the ongoing collection, maintenance, and support of the Utah Population Database (UPDB). Partial support for the UPDB (grant number P30 CA2014) from the National Cancer Institute, University of Utah and from the University of Utah's Program in Personalized Health and Center for Clinical and Translational Science is also acknowledged.

Preparatory courses were supported by a DANS grant for attending the ICPSR Summer School 2016 at the University of Michigan (Ann Arbor, Michigan, US) and a Erasmus+ grant for attending a Summer School at CEDAR, University of Umeå (Umeå, Sweden). A conference visit was supported by the SSHA Tilly Graduate Student Travel Award for the Annual Meeting of the Social Science History Association (Phoenix, Arizona, US).

A word of thanks

Four delightful years were spend working on this project. My words of thanks are devoted to those who made the project worthwhile, and especially to those who reminded me that I was not my project, and that the project was not my life, however important it was to me. I'm thankful to those who encouraged me to start this project, to finish it, and to continue in academia. Not all of you are mentioned here by name, but I am sure that you know who you are.

First of all, I want to offer a warm thank you to my supervisors Angelique Janssens and Jan Kok, without whom this project would not have been possible. Angelique, thank you for unfaltering support throughout the process, and cheering me on in the more frustrating periods. Jan, thank you for your worthwhile comments on earlier versions of this work. In addition, a warm word of thanks is offered to the members of the Advisory Board of the "Genes, Germs and Resources" project—Isabelle Devos, Sören Edvinsson, and Ken Smith—as well as to the other members of the project team, Eline Slagboom and Marian Beekman. A further word of thanks goes to Kees Mandemakers, for delivering a thousand versions of the data files necessary for our work—I'm sorry that we were such a pain.

Further I want to mention Niels van den Berg and Rick Mourits, my colleagues in the *Genes, Germs and Resources* project, without whom the project would not have been the same. I enjoyed our discussions on the possibilities and impossibilities of the data, collective bickering on slow academic processes, and our very reliable instant-support-hotline on social media. The absolute highlight of our cooperation was our research visit to Utah, with lots of basketball, biking around town, super hero series, camping trips in the South of Utah and Yellowstone—and discussions about our work. It was great to share this journey with you and I hope we can keep working together in the future. Further, a warm thanks to Ken Smith and everyone at UPDB for the warm welcome to Utah, and showing Rick, Niels and me around in Salt Lake City. The work that we did in Utah, and the possibility to compare results between the Utah Population Database, greatly contributed to the project, and I am very grateful about this opportunity.

Furthermore, it was a delight to conduct my research at Economic, Social and Demographic History at the Radboud University. I enjoyed the warm and cooperative environment, the support, and discussing worthless and worthwhile research ideas and topics completely unrelated. At the department of History in general, I found a warm environment that I will dearly miss. Thanks to my colleagues for sharing a million coffee breaks, cookie breaks, lunch breaks, ice cream breaks and outings in

the city. I truly regret to lose you as my daily colleagues but I hope we will stay in touch and continue our friendship.

I am very thankful for the support of my parents, siblings and friends for their support throughout my studies and doctoral project. Specifically, I want to mention Number Two, my identical twin, whom I keep around as a potential donor and stand-in for the defense. Karin, thanks for your support and please be careful with my spare liver at the defense party. Also I'd like to thank my other siblings and parents for their support in this endeavor. Your practical and emotional support means a lot to me.

Joost, Bente, Lieke, Dominique, Henri: collectively you represent happiness in high definition, and I´m thankful to have you in my life. The last ten years wouldn't have been the same without the five of you. I know we have not been seeing each other too much lately and this project was certainly to blame for that! I hope we will be able to spend a bit more time together in the coming months now that this is off my plate, even if we will have a bit more travelling time between Sweden and The Netherlands.

Finally, thank you Kai, for being around. I´m very glad that we found each other. Off to our next adventure!

Lund, April 2019

Curriculum Vitae

Ingrid van Dijk was born on 17 November 1989, in Groningen. She received her bachelor degree in Sociology in 2012 and her master degree in comparative sociological research in 2014 at the Radboud University Nijmegen, The Netherlands. Both degrees were awarded *cum laude.* In the same year, she took on her doctoral project within the *Genes, Germs and Resources* project at the Department of Economic, Social and Demographic History, at Radboud University. In the course of her PhD dissertation, she was a guest researcher at Population Studies, University of Utah, US. From 2019, she was employed as a postdoctoral researcher at the Centre for Economic Demography, Department of Economic History, University of Lund, Sweden.