**Social Sciences: Social Sciences**

# EARLY-LIFE DISEASE EXPOSURE AND ASSOCIATIONS WITH ADULT SURVIVAL, CAUSE OF DEATH AND REPRODUCTIVE SUCCESS IN PRE-INDUSTRIAL HUMANS

**Short title: Early-life disease and later fitness in humans**

**Accepted for publication in** ***Proceedings of the National Academy of Sciences***

Adam D. Hayward[1,2], Francesca L. Rigby[3] & Virpi Lummaa[3,4]

1. Department of Biological and Environmental Sciences, School of Natural Sciences, University of Stirling, Stirling, FK9 4LA, UK.
2. Institute of Evolutionary Biology, Ashworth Laboratories, University of Edinburgh, Charlotte Auerbach Road, Edinburgh, EH9 3FL, UK.
3. Department of Animal and Plant Sciences, Alfred Denny Building, University of Sheffield, Western Bank, Sheffield, S10 2TN, UK.
4. Department of Biology, University of Turku, FIN-20014 Turku, Finland

**Corresponding author:**

Adam Hayward

Department of Biological and Environmental Sciences, School of Natural Sciences, University of Stirling, Stirling, FK9 4LA, UK.

Tel: +44(0) 1786 467757

Email: adam.hayward@stir.ac.uk

**Keywords:** Inflammation, stress, infection, life-history, fitness

## ABSTRACT

A leading hypothesis proposes that increased human lifespan since 1850 has resulted from decreased exposure to childhood infections, which has reduced chronic inflammation and later-life mortality rates, particularly from cardiovascular disease, stroke and cancer. Early-life cohort mortality rate often predicts later-life survival in humans, but such associations could arise from factors other than disease exposure. Additionally, the impact of early-life disease exposure on reproduction remains unknown, and thus previous work ignores a major component of fitness, through which selection acts upon life-history strategy. We collected data from seven 18$^{th}$ and 19$^{th}$ century Finnish populations experiencing naturally-varying mortality and fertility levels. We quantified early-life disease exposure as the de-trended child mortality rate from infectious diseases during an individual's first five years, controlling for important social factors. We found no support for an association between early-life disease exposure and all-cause mortality risk after age 15 or 50. We also found no link between early-life disease exposure and probability of death specifically from cardiovascular disease, stroke or cancer. Independently of survival, there was no evidence to support associations between early-life disease exposure and any of several aspects of reproductive performance, including lifetime reproductive success and age at first birth, in either males or females. Our results do not support the prevailing assertion that exposure to infectious diseases in early life has long-lasting associations with later-life all-cause mortality risk or mortality putatively linked to chronic inflammation. Variation in adulthood conditions could therefore be the most likely source of recent increases in adult lifespan.

## SIGNIFICANCE STATEMENT

Why has human life expectancy increased since 1850? A leading hypothesis proposes that limited exposure to childhood infections has reduced lifelong inflammation and enhanced survival, but tests of this hypothesis typically use all-cause mortality rates to estimate disease exposure. Meanwhile, links between early-life disease and reproduction have been neglected. We used data from pre-industrial Finnish populations to show that early-life disease exposure was not associated with all-cause mortality, mortality from cardiovascular disease, stroke and cancer, or reproductive success. Our study therefore does not support the prevailing contention that reduced exposure to early-life infections has increased life expectancy in modern populations.

\body

## INTRODUCTION

In industrialized nations, human life expectancy at birth has increased by around three months per year between 1840-2002 (1). While this is linked to improvements in child survival rates, there have been considerable increases in life expectancy of adults: in the UK in 1843, the mean expected age at death of a 20-year old was 60, while in 2011 it was 81 (2, 3), a 50% increase in remaining lifespan. Longer adult life expectancy has been accompanied by increased incidence of previously rare chronic health problems and thus identifying the drivers of enhanced lifespan is an important question for global health. Declines in adult mortality rate since the 18$^{th}$ century in Europe were more closely linked to year of birth than to the year in which mortality was assessed, leading to the hypothesis that: "...the expectation of life was determined by the conditions which existed during the child's earlier years" (4). The 'cohort morbidity phenotype' hypothesis suggests that the link between declining birth year mortality rate and increasing adult survival is due to reduced exposure to infectious diseases in early life and reduced lifelong burden of chronic inflammation (5–8). Thus, individuals experiencing exposure to infectious disease in early life are expected to have higher adulthood mortality risk.

It is hypothesized that early-life exposure to infectious diseases could increase later-life mortality risk through a number of physiological mechanisms (6). Bacterial and viral infections elicit inflammatory immune responses, and chronic inflammation is associated with atherosclerosis, increased risk of stroke, cardiovascular disease (CVD) and mortality (9). Thus, early-life disease exposure is expected to be linked to adult mortality risk from these causes in particular (6). Support for this hypothesis comes from studies showing that exposure to high levels of early-life cohort mortality or infection is associated with middle-

age mortality from CVD (10, 11); that exposure to infection is associated with chronically high levels of inflammation (12, 13); that high levels of inflammation are associated with CVD (14). This physiological mechanism is underpinned by the fact that natural selection should favour the expression of robust immune responses at young ages, even at the cost of lower later-life survival (15).

Building on such knowledge, studies have tested the 'cohort morbidity phenotype' hypothesis by showing that early-life infections are associated with chronic inflammation and impaired later health (16, 17). A drawback is that these studies are performed on modern populations with good health care and therefore cannot estimate associations between early-life disease exposure and natural mortality risk. A second approach is to use historical data on natural-mortality pre-industrial populations and to link early-life cohort mortality (a proxy for early-life disease exposure) with later-life mortality (18–20). Studies of historical populations have however experienced at least one of several shortcomings. First, many studies are based on population-level statistics, meaning they cannot identify factors allowing some individuals to live long lives, despite being born in high-mortality cohorts. Second, most studies only consider all-cause cohort mortality rate as a measure of disease exposure, which may be driven by nutrition and social factors as well as disease. Third, it is often unknown which types of deaths are associated with early stressors, obfuscating the mechanism through which early-life disease exposure is associated with later mortality.

The majority of studies of early-life disease exposure have focused solely on associations with later-life mortality, despite the fact that selection on life-history strategy should act through both survival and reproductive success. Life-history theory predicts that high mortality risk in early life should lead to earlier age at maturity (21), but associations between

early-life disease exposure and between-individual variation in reproductive success have been neglected. A series of studies have shown that early life trauma, such as sibling mortality, parental separation, disadvantageous birth location and famine (22–25) are associated with earlier age at first reproduction and/or larger family size, suggesting adoption of a 'fast' life-history in high-stress, high-mortality environments. In contrast, childhood infections may have a non-lethal but damaging impact upon physiology which may impede reproduction during adulthood (26). These could include infections after birth leading to stunted growth in childhood and eventually reduced adult size (27), or even infections of focal individuals' mothers leading to reduced birth weight, which may lead to reduced birth weight of focal individuals' offspring (28, 29). Nevertheless, it has been shown that the direction of the association between final adult height and fertility is variable across populations (30).

We used church register data collected from seven Finnish populations across the 18$^{th}$ and 19$^{th}$ centuries to determine the association between early-life exposure to infectious diseases and (1) later-life mortality risk; (2) cause of death; (3) reproductive success. These populations had no access to modern contraception or healthcare (31) and as a result ~40% of children died of disease before age 15. Having survived to age 15, mean lifespan in our sample was ~56 years and individuals produced an average of ~4 surviving offspring, with >99% of individuals born to parents aged <50. We used cause-of-death data recorded by churches to adopt a measure of disease exposure based on the local probability of child death from infectious diseases during the early years of an individual's life. Crucially, we controlled for other salient variables reflecting early-life conditions and known to influence fitness in these populations (SI Appendix, Table S1 and Fig. S1), including social class at birth (32), twin status (33), inter-birth interval (34) and season of birth (35).

## RESULTS

There was considerable between-year variation in the prevalence of deaths from infectious diseases in the early years of our study individuals' lives. We used these data to calculate a measure of early-life exposure to infectious diseases for each of the first five years of life of each individual, based on the proportion of children dying from infectious diseases in their parish in each year. We recorded the number of children aged 0-10 alive in each parish in each year, 1751-1855 (a total of 195,170 person-years). In 194,182 (>99.5% of) cases, the child had a known cause of death or survived, while in the remaining 988 cases, the child died of an unknown cause; these 988 cases were not included in our calculations of early-life disease exposure (SI Appendix, Table S2). During these years, 4780 child deaths of known cause were recorded, with 73% due to infectious diseases (SI Appendix, Table S3). We calculated annual parish-specific variation in child mortality rate from infectious diseases by dividing the number of child (aged 0-10) deaths in each parish due to infections in each year, 1751-1855, by the number of children alive in each parish in the same year. The child mortality rate from infectious diseases varied substantially across the study period in all seven parishes (Fig. 1). We de-trended the child mortality rate from infectious diseases by applying a Hodrick-Prescott filter to data from each parish using the R package 'mFilter', with a filtering term of 100. This decomposed variation in child mortality rate from infectious diseases in each parish into a long-term smoothed 'trend' component (Fig. 1) and the short-term deviation from the trend, or the 'cycle' (Fig. 2). The 'cycle' component for each year was multiplied by 100 to aid model convergence, and we refer to this measure herein as annual 'disease exposure' (Fig. 2). Each individual was therefore defined by their potential early exposure to infections, rather than simply the level of all-cause cohort mortality (8, 18,

19), on a continuous quantitative scale, accounting for temporal changes in conditions, and specific to the parish in which they were born. We designed analyses in order to determine whether our measure of disease exposure in each of the first five years of life was associated with later-life survival and reproductive success.

**Mortality risk**

Our mixed-effects Cox model analyses showed that greater disease exposure in each of the first five years of life was not associated with increased mortality risk in adulthood (N=7283). Controlling for a higher mortality risk among poor individuals and males, and variation between parishes, as well as other fixed and random effects (Table 1), disease exposure measured in each of the first five years of life was not significantly associated with hazard of mortality after age 15 and nor was the association between early-life disease exposure and mortality risk dependent upon parish, social class, or sex (SI Appendix, Table S4). When we restricted our analyses to individuals who survived to at least the age of 50 (N=3822) and who were therefore most likely to experience stroke, CVD and cancer, we again found no evidence for an association between early-life disease exposure and mortality risk and no indication of interactions with individual parish, social class or sex (SI Appendix, Table S4).

**Cause of death**

Mixed-effects Cox models showed that there was no significant association between early-life disease exposure and cause of death. We tested for an association between disease exposure in each of the first five years of life and death from CVD, stroke and cancer, on which the 'cohort morbidity phenotype' focuses (8). Individuals were scored as 1 (a 'failure')

if they died of CVD, stroke or cancer, and 0 if they died of other causes or were censored without dying (SI Appendix, Table S5 & Fig. S2). As above, we first performed the analysis on all individuals who survived to at least the age of 15. There was no statistical support for an association between early-life disease exposure in any of the first five years of life and death from CVD, stroke or cancer (SI Appendix, Table S6). Although there was some evidence to support an interaction between early disease exposure and social class (Disease Year 4:Social class (Poor) Hazard = 0.79; 95% C.I.=0.62-1.00; $\chi^2_1$=4.17; P=0.0.041) suggesting that Poor individuals with higher early-life disease exposure in the fourth year of life were less likely to later die of infection (Table S7), this interaction was only weakly statistically supported over the base model with ΔAIC=-2.02 (36). Nevertheless, it suggested a trend for a mildly protective effect of early disease exposure and later death risk from CVD, stroke or cancer in Poor individuals. We also repeated this analysis on individuals who survived to at least age 50, since the majority (263/346) of individuals affected by these death causes were aged over 50. Once again, there was no evidence that early-life disease exposure was associated with cause of death, nor any evidence for interactions between disease exposure and parish, social class or sex (SI Appendix, Table S6).

**Lifetime reproductive success**

We next assessed links between early-life disease exposure and later-life reproductive success in individuals who married and survived to at least age 50 and who had therefore completed their reproductive lifespans with the opportunity to reproduce. We analysed data from females and males separately, due to their differing social and physiological factors affecting reproductive schedules. First, we used linear mixed-effects models to show that there was no association between early-life disease exposure and Box-Cox transformed

lifetime reproductive success (LRS) in females (N=1512), defined as the number of offspring an individual produced which survived to age 15 (best-fitting single year, Disease year 1 estimate = -0.0202±0.0112; $\chi^2_{(1)}$=3.25; P=0.071; SI Appendix, Table S8). This model controlled for variation between parishes ($\chi^2_{(6)}$=68.60; P<0.001), lower LRS in Poor compared to Rich individuals (estimate=-0.24±0.04; $\chi^2_1$=44.77; P<0.001) and non-significant effects of birth interval, birth season and twinning status. There was also no statistical support for interactions between early disease exposure and social class or parish (SI Appendix, Table S8). In males (N=1333), there was once again no support for an association between early-life disease exposure and LRS (best-fitting single year, Disease year 4 estimate=0.0068±0.0118; $\chi^2_1$=0.33; P=0.563), nor any support for interactions with social class or parish (SI Appendix, Table S8).

We next determined whether there were any links between early-life disease exposure and other components of reproduction which constitute LRS, namely age at first reproduction, the number of children born, or child survival rate. In neither sex, however, did we find any support for associations between early-life disease exposure and any of these measures of reproductive success, either as a main effect or in interactions with social class or parish (SI Appendix, Tables S9-S11).

## DISCUSSION

It has been hypothesized that recent increases in human lifespan have been brought about by reduced exposure to infectious diseases in early life and reduction of lifelong inflammation and associated chronic diseases such as CVD, stroke and cancer (5–8). Tests of this hypothesis have used data from historical populations to support a link between cohort all-

cause mortality rate and later-life survival, but have not assessed later-life survival in relation to measures of early-life disease exposure based on disease prevalence, or determined how early disease exposure is linked to mortality risk specifically from CVD, stroke and cancer. Studies of how early-life disease exposure may impact upon later-life reproductive success are also extremely rare. We used data from seven pre-industrial Finnish populations and addressed these shortcomings by (1) quantifying variation in early-life disease exposure as the proportion of children alive in each of the first five years of an individual's life who died of infectious diseases; (2) assessing links between disease exposure and cause of death; (3) determining the association between early-life disease exposure and reproduction. Our results show that individuals experiencing higher early-life disease exposure do not experience higher all-cause mortality risk or risk of mortality from CVD, stroke or cancer during adulthood. We also found no support for associations between early-life disease exposure and reproductive performance in either males or females. Our results do not support the link between early-life disease exposure and chronic inflammation-associated mortality risk which forms the crux of the 'cohort morbidity phenotype' hypothesis.

Our analyses showed that, controlling for individual parish, social class, sex, birth season, twinning status, birth interval and temporal changes in disease exposure (Fig. 1) there was no statistical support for an association between early-life disease exposure and mortality risk after age 15 (Table 1 and SI Appendix, Table S4). Previous studies have found evidence for links between early-life cohort mortality rate and later survival (4, 5, 8), but without considering that such a measure of disease exposure could be confounded by spatial variation, social status, other intrinsic and extrinsic conditions as well as disease exposure, and improvements in living conditions and hygiene across time. Demographic studies have applied statistical techniques such as Hodrick-Prescott decomposition to trending

environmental data and found generally weak associations between early- and later cohort mortality (37–40), though some of these studies have still detected significant correlations (19, 41). By applying Hodrick-Prescott decomposition we were able to account for declines in death rates from infection across time by using deviations from this trend as our measure of early-life disease exposure (Fig. 2). Unlike these studies, however, we use a measure of disease exposure explicitly based on prevalence of deaths from the infectious diseases most likely to affect infants (SI Appendix, Tables S2 & S3).

Early-life disease exposure may influence later-life health in a number of ways, and it is this complexity that may have prevented us from detecting significant associations. For example, early infections may have negative associations with later health through chronic physiological 'scarring', but may also have positive associations due to acquired immunity or selection of the most robust individuals who were able to survive the harsh early environment (42). There may also be changes in the factors which influence mortality in different parts of the life course, such that the relative importance of scarring and selection may vary with age (43) and the effects of early adversity may be short-lived, or be negligible in comparison to events which occur in adult life (39). A study using data from Sweden, collected 1813-1968 and calculating early-life disease risk based on neonatal mortality rates, showed that effects of exposure on mortality varied with age: individuals experiencing higher neonatal mortality rates showed higher mortality initially, but lower mortality in later life (44). These results suggest initial scarring effects of disease at younger ages, followed by a strengthening influence of acquired immunity and/or selection of robust individuals at later ages. Thus, rather than being unimportant for later-life survival, early-life infections may have both negative (through scarring) and positive (through acquired immunity and/or selection)

associations with later survival which may combine to culminate in no detectable overall effect.

The ‘cohort morbidity phenotype’ (5–8) proposes a link between infection, chronic inflammation, and risk of CVD, stroke and cancer (14). The exact prevalence of these diseases in our study population is unknown: although 9.6% of deaths were due to CVD or stroke, a further 13.7% of deaths were due to ‘old age’ and could have been related to chronic inflammation (SI Appendix, Table S5). Exposure to infection in early life has also been linked to respiratory disorders in later life (16, 17, 45), which is noteworthy given that 14.9% of deaths in our studied individuals were from tuberculosis (SI Appendix, Table S5). Focusing on deaths from CVD, stroke and cancer, we did not find evidence that early-life disease exposure was associated with these causes of death, linked to chronic inflammation by the ‘cohort morbidity phenotype’ (SI Appendix, Table S6). It should be noted that, with only 346/7283 (5% of) individuals scored as dying of CVD, stroke or cancer, our statistical power to detect links between early disease exposure and cause of death may have been low. Meanwhile, studies of contemporary populations experiencing high exposure to infection and high levels of inflammatory markers rarely show signs of CVD despite their short lifespans, suggesting that a low level of CVD, stroke and cancer is simply the ancestral state and that high prevalence of these conditions may be a consequence of longer lifespans in industrialized populations (46). Finally, it may be possible that adverse effects of chronic inflammation are only apparent in populations experiencing resource-rich or pathogen-poor environments, and that in our (relatively) resource-poor and pathogen-rich environments, heightened inflammatory responses are advantageous or neutral. Our study individuals were born before 1850, and thus it is possible that our results may not be generalizable to later cohorts, which experienced increased adult nutrition and healthcare. On the other hand, if

chronic inflammation does not lead to CVD, stroke and cancer in low-nutrition, high-pathogen environments, then the 'cohort morbidity phenotype' would not offer an explanation for longer modern lifespans being due to reduced or delayed instances of these disorders.

The majority of studies of early-life disease exposure have focused solely on associations with mortality, and our results are among the first to address links between early-life disease exposure and reproductive success. We found, however, that associations between early-life disease exposure and lifetime reproductive success (LRS), age at first reproduction, lifetime breeding success and child survival rate were absent in both sexes. Life-history theory predicts that high mortality risk in early life should lead to earlier age at maturity (21), which should be associated with higher LRS, predictions which have been supported by population- (23, 47) and individual-level studies showing that adverse early-life conditions such as experience of mortality, a deprived upbringing, experience of famine or chronic childhood illness are associated with earlier age at first birth and higher fertility (22, 23, 25). In contrast, early-life stressors such as low food availability have been linked to later-life health problems (48) and, in our study population, reduced reproductive success (49, 50). These effects may be mediated through neonatal infections which stunt growth and diminish later reproductive success (26, 27), or maternal infections which reduce the birthweight of focal individuals (28, 29). In a historical German population, birth during a high-mortality period was associated with lower child survivorship in males and lower lifetime births in females (51). Thus, adverse conditions in early pre- or post-natal life may slow growth and development, affect reproductive physiology and postpone first birth (52). As with survival, the conflicting negative (scarring) and positive (selection) effects of disease exposure may underpin our finding of no overall association between early-life disease exposure and later-life

reproductive success. Alternatively, the lack of association in the present study could indicate that links detected in previous studies were due to factors other than disease. A recent study on data from 251 20th century English women used path analysis to show that although social class was linked to age at first birth, and illness in the first year of life was linked to adult height, there was no direct or indirect effect of childhood illness on age at first birth or LRS (53). Our results, with N>1000 are therefore unlikely to reflect lack of statistical power and it may therefore simply be that effects of early-life infections on reproduction are weak and overshadowed by socio-economic factors.

Unlike many previous studies testing for links between early-life disease exposure and later-life fitness, we (1) incorporated data on disease prevalence into our measure of early-life disease exposure; (2) determined associations between disease exposure and cause of death; (3) assessed associations between disease exposure and reproductive success; (4) controlled for several important social factors; (5) de-trended our measure of early-life disease exposure to account for temporal trends; (6) examined early-life effects not only around birth but up to age five. An 'ideal' study would link full health records from each individual documenting previous illness with lifetime follow-up on survival and reproduction. Our results should be interpreted with the acknowledgment that infection of any individual in our population at any time is unknown. Instead, our measure of disease exposure assumes that a higher child mortality rate from infection meant that, on average, individuals experiencing such conditions were more likely to be exposed to infections. In addition, we included the 22% of child deaths due to 'unknown diseases' in the 'infectious' category since deaths from infection were generally more common than those due to other causes, especially in children (Tables S3 & S5), but this may have introduced noise into our calculation of disease exposure if not all such unknown diseases were indeed infectious. It therefore remains possible that inability

to detect an association between early disease exposure and fitness is due to our way of calculating early disease exposure. Ability to accurately determine cause of death with poor medical knowledge is a general caveat of using historical data. These two caveats aside, a strength is that we have used parish-specific records rather than nationwide figures, thereby gaining the most geographically-restricted estimate of disease exposure possible.

The results of our study strongly challenge the hypothesis that links between early and later-life mortality rates are mediated through effects of early-life infectious diseases on chronic inflammation. Longitudinal study of contemporary populations will be essential for detecting the mechanisms underpinning such effects (53); for example, for disentangling the role of early and later conditions in influencing mortality risk. With the increasing availability of national, individual-based health data, this endeavour will be achievable in a more explicit manner than shown here. Our study, however, uses data from 100 cohorts of a pre-industrial population and controls for variation in social class and birth circumstances to provide a unique snapshot of how early-life conditions may have influenced human populations for much of our evolutionary history (6, 7). We conclude, in agreement with several other authors (37, 39, 40) that the influence of improved adult, rather than childhood, conditions are likely to have driven recent improvements in life expectancy.

## MATERIALS AND METHODS

### Study population and data collection

We investigated associations between early-life disease exposure and later-life survival and reproductive success using longitudinal data collected from 18$^{th}$-19$^{th}$ century church records

in seven populations, termed ‘parishes’, located across the southern half of Finland. Prior to the 1870s, these populations had high birth and death rates, primitive agricultural technology, unreliable contraception and medical care (31) and were strictly monogamous. Individuals were divided into two social classes: “Rich” individuals included farm owners, merchants and craftsmen; “Poor” individuals included crofters and labourers (32).

**Causes of death and early-life disease exposure**

We have collected data including birth, parentage, marriage, reproduction and death for a total of 72,564 individuals in the parishes of Hiittinen, Ikaalinen, Jaakkima, Kustavi, Pulkkila, Rautu and Rymättylä born 1702-2012. We restricted our analyses to 21,539 individuals born after 1750, since before this records were patchy, and born before 1851, ensuring that we only used data from individuals with natural mortality and fertility (31). We further restricted our analyses to individuals who survived to at least age 15 (15,237) and those of known maternal identity and social class (N = 7327). This was done to ensure we captured longer-term effects of disease exposure; and to control for variation in survival and reproductive success between families and social classes (32). Finally, because birth order (34), twinning status (33) and birth season (35) are all linked to fitness in this population, we restricted our analyses to individuals for whom this information was recorded, leaving a final sample size of 7283 individuals, born 1751-1850. None of our study individuals moved between parishes between birth and adulthood.

Based on previous studies examining links between early exposure to mortality and later-life health and fitness (19, 38–40), we aimed to estimate an individual’s experience of infectious diseases in their first five years of life, as illustrated by the prevalence of death from

infectious diseases recorded by the church (54). Since children and adults are vulnerable to different diseases in this population (SI Appendix, Tables S3 & S5), we recorded the number of individuals aged 0-10 years who were alive in the parish in each year, 1751-1855, a total of 194,182 person-years (SI Appendix, Table S2), during which time 4780 children died of a known cause, 73% of which were due to infectious diseases (SI Appendix, Table S3). 22% of deaths were from an 'unknown disease', which we included in the 'infectious' category. We gained a parish-specific measure of 'disease exposure' for each year by dividing the number of children dying of infectious diseases in a given year in a given parish by the number of children alive in the same parish in the same year (Fig. 1). Thus, our measure of early-life disease exposure is based on the child mortality rate from infectious diseases during an individual's early life. We then de-trended this measure using Hodrick-Prescott decomposition, and used the deviations from the trend (the 'cycle' component) as our measure of disease exposure (Fig. 2). Early-life disease exposure was multiplied by 100 prior to use in analysis to aid model convergence.

**Statistical analysis**

We determined the association between disease exposure in each of the first five years of life and later-life fitness, controlling for other potentially confounding effects (SI Appendix, Table S1). All models described below were mixed-effects models with maternal identity and birth year as random effects, to account for between-family and -cohort variation in fitness. All models estimated the association between early-life disease exposure and fitness after controlling for the following fixed effects, which were retained in all models. Parish (seven levels), social class (Rich or Poor) and sex (Male or Female) were all included as categorical fixed effects. Since child survival is higher for those born in August, September and October

compared to other months (35), birth season was included as a two-level categorical variable (High or Low survival). First-born males have higher prospects of wealth inheritance, marriage and reproduction in this population, and individuals born with a longer intervening period after their previous sibling may have higher fitness (34). We incorporated both of these effects into a four-level categorical variable, Inter-Birth Interval (IBI: Firstborn; born <2 years after sibling; born <3 years after sibling; born >3 years after sibling). Finally, all models accounted for the link between twin status at birth and later-life fitness (33) by including twin status at birth (singleton or twin) as a categorical fixed effect. All analyses were conducted in R version 3.2.3.

**Mortality risk**

We determined the association between early-life disease exposure and mortality risk after the age of 15 in 7283 individuals using mixed-effects Cox models in the R package 'coxme'. We constructed a 'base' model containing the fixed and random effects described above (model 0) and then built a series of models testing the association between early-life disease exposure and mortality risk. We considered disease exposure in each of the first five years of an individual's life as separate variables: Year 0 (birth year), Year 1, Year 2, Year 3 and Year 4. We first tested for main effects of disease exposure in each of the first five years of life separately by (models 1-5) comparing the 'base' model with separate models containing each of the first five years' disease exposure, fitted as covariates, using likelihood ratio test (LRTs), where the $\chi^2$-distributed test statistic is calculated as $-2*(\text{LogLik}_{\text{model1}} - \text{LogLik}_{\text{model2}})$. We next (model 6) fitted all five disease exposure years as covariates in the same model and compared model 6 to model 0 using an LRT. We determined which of models 1-5 fitted best by comparing their Akaike Information Criterion (AIC) values, where

the lowest value provided the best fit (36). We did not encounter any situation where more than one of models 1-5 outperformed model 0. This allowed us to determine which disease exposure year was most strongly associated with mortality risk. We took the best-fitting disease exposure year forwards into model 7, which tested for a non-linear association between the best-fitting disease exposure year and survival by comparing the model with a linear term of disease exposure to one with both linear and quadratic terms. Finally, we tested for interactions involving the best-fitting disease exposure year by comparing the model with the main effect of the best-fitting disease exposure year to models with interactions between disease exposure and parish (model 8), social class (model 9) and sex (model 10). We subsequently restricted the analyses to individuals who lived to at least age 50 (N=3822), since these individuals are most likely to fall victim to CVD, stroke and cancer.

**Cause of death**

We next aimed to determine whether early-life disease exposure was associated with increased risk of death from stroke, CVD or cancer versus other causes. Each individual was scored as dying from stroke, CVD or cancer (1; 346 individuals) with the remainder (6937 individuals) scored as 0. Cause-of-death analyses were performed using mixed-effects Cox models, following the same fixed and random effects structure and model selection procedure as described above. We then repeated this analysis on individuals who survived until at least the age of 50, since 263/346 (76%) of deaths from these causes occurred in individuals aged 50 and over.

**Lifetime reproductive success**

We next determined the association between early-life disease exposure and later-life reproductive success by quantifying lifetime reproductive success (LRS) as the number of children an individual produced that survived to age 15. From the original sample size of 7283, we removed data from 1200 individuals who were unmarried or whose marriage status was unknown; 488 individuals who had unknown LRS; 1417 individuals whose reproductive success was unknown or whose children had unknown fates to age 15; 1333 individuals who failed to survive to the age of 50 and therefore complete their reproductive lifespan (49). We therefore focused on fertility differences within marriage to determine differences in reproductive capacity, rather than 'mating success' (or marriage) which in influenced by a host of social and cultural factors which contribute to finding a partner. The final data set for analysis of LRS contained 2845 individuals, with 1512 females and 1333 males. The analyses were separated by sex due to the biological and social differences governing variation in reproductive success between the sexes in this population. We visually inspected diagnostic plots of several different modelling structures, including linear mixed-effects models (LMMs) of untransformed LRS; LMMs of log-transformed LRS; LMMs of Box-Cox transformed LRS; generalized linear mixed effects models of untransformed LRS with Poisson, Poisson-lognormal and negative binomial error structures. In both males and females, Box-Cox transformation provided the most favourable diagnostics, so LRS was modelled using linear mixed-effects models (LMMs) of $(LRS+1)^{0.5}$ in the R package 'lme4'. The fixed and random effect structures of models 1-9 were the same as for the Cox models described above, although since we performed separate analysis for males and females, we did not fit model 10, which included the interaction between early-life disease exposure and sex.

Finally, we tested for associations between early-life disease exposure and the reproductive traits which contribute to LRS. We therefore performed analyses determining whether early-life disease exposure was associated with later age at first reproduction, reduced lifetime breeding success (number of children born) and/or child survival rate to age 15. These analyses are described in full in the SI Appendix.

**ACKNOWLEDGEMENTS**

We thank Kimmo Pokkinen for collection of the Ikaalinen church record data, and three reviewers for their comments on previous versions which have greatly improved the manuscript. This research was funded by the Academy of Finland (V.L.) and Wellcome Trust (A.D.H.).

## REFERENCES


1. Oeppen J, Vaupel JW (2002) Demography. Broken limits to life expectancy. *Science* 296(5570):1029–31.
2. Kertzer DI, Laslett P (1995) *Aging in the Past: Demography, Society and Old Age* (University of California Press, Berkeley, California).
3. Roser M (2015) Life expectancy. Available at: http://ourworldindata.org/data/population-growth-vital-statistics/life-expectancy/ [Accessed February 3, 2016].
4. Kermack WO, McKendrick AG, McKinlay PL (1934) Death-rates in Great Britain and Sweden. Some general regularities and their significance. *Lancet* 223(5770):698–703.
5. Crimmins EM, Finch CE (2006) Infection, inflammation, height, and longevity. *Proc Natl Acad Sci U S A* 103(2):498–503.
6. Finch CE (2010) Evolution of the human lifespan and diseases of aging: roles of infection, inflammation, and nutrition. *Proc Natl Acad Sci U S A* 107 Suppl(suppl_1):1718–24.
7. Finch CE (2012) Evolution of the human lifespan, past, present and future: phases in the evolution of human life expectancy in relation to the inflammatory load. *Proc Am*

549 *Philos Soc* 156(1):9–44.
550 8. Finch CE, Crimmins EM (2004) Inflammatory Exposure and Historical Changes in
551 Human Life-Spans. *Science (80- )* 305(5691):1736–1739.
552 9. Vasto S, et al. (2007) Inflammatory networks in ageing, age-related diseases and
553 longevity. *Mech Ageing Dev* 128(1):83–91.
554 10. Mazumder B, Almond D, Park K, Crimmins EM, Finch CE (2010) Lingering prenatal
555 effects of the 1918 influenza pandemic on cardiovascular disease. *J Dev Orig Health*
556 *Dis* 1(1):26–34.
557 11. Myrskylä M, Mehta NK, Chang VW (2013) Early life exposure to the 1918 influenza
558 pandemic and old-age mortality by cause of death. *Am J Public Health* 103(7):e83–90.
559 12. Gurven M, Kaplan H, Winking J, Finch C, Crimmins EM (2008) Aging and
560 Inflammation in Two Epidemiological Worlds. *Journals Gerontol Ser A Biol Sci Med*
561 *Sci* 63(2):196–199.
562 13. Vasunilashorn S, et al. Blood lipids, infection, and inflammatory markers in the
563 Tsimane of Bolivia. *Am J Hum Biol* 22(6):731–40.
564 14. Danesh J, et al. (2000) Low grade inflammation and coronary heart disease:
565 prospective study and updated meta-analyses. *BMJ* 321(7255):199–204.
566 15. Williams GC (1957) Pleiotropy, natural selection, and the evolution of senescence.
567 *Evolution (N Y)* 11:398–411.
568 16. Barker DJ, et al. (1991) Relation of birth weight and childhood respiratory infection to
569 adult lung function and death from chronic obstructive airways disease. *BMJ*
570 303(6804):671–5.
571 17. Case A, Fertig A, Paxson C (2005) The lasting impact of childhood health and
572 circumstance. *J Health Econ* 24(2):365–89.
573 18. Beltrán-Sáncheza H, Crimmins EM, Finch CE (2012) Early cohort mortality predicts

the rate of aging in the cohort: a historical analysis. *J Dev Orig Health Dis* 3(5):380–6.

19. Bengtsson T, Lindström M (2003) Airborne infectious diseases during infancy and mortality in later life in southern Sweden, 1766–1894. *Int J Epidemiol* 32(2):286–294.

20. 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, 1760-1984. *Popul Stud (NY)* 54(3):263–277.

21. Stearns SC (1992) *The Evolution of Life Histories* (Oxford University Press, Oxford).

22. 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.

23. Nettle D (2010) Dying young and living fast: variation in life history across English neighborhoods. *Behav Ecol* 21(2):387–395.

24. Nettle D, Coall DA, Dickins TE (2011) Early-life conditions and age at first pregnancy in British women. *Proc Biol Sci* 278(1712):1721–7.

25. Painter RC, et al. (2008) Increased reproductive success of women after prenatal undernutrition. *Hum Reprod* 23(11):2591–2595.

26. Khan AD, Schroeder DG, Martorell R, Haas JD, Rivera J (1996) Early childhood determinants of age at menarche in rural guatemala. *Am J Hum Biol* 8(6):717–723.

27. Beard AS, Blaser MJ (2002) The Ecology of Height: The Effect of Microbial Transmission on Human Height. *Perspect Biol Med* 45(4):475–498.

28. Nohr EA, Vaeth M, Rasmussen S, Ramlau-Hansen CH, Olsen J (2009) Waiting time to pregnancy according to maternal birthweight and prepregnancy BMI. *Hum Reprod* 24(1):226–32.

29. Ekholm K, Carstensen J, Finnström O, Sydsjö G (2005) The probability of giving birth among women who were born preterm or with impaired fetal growth: a Swedish

599 population-based registry study. *Am J Epidemiol* 161(8):725–33.
600 30. Sear R (2010) Height and reproductive success: is bigger always better? *Homo Novus:*
601 *A Human Withough Illusions*, eds Frey U, Stoermer C, Willfuehr K (Heidelberg), pp
602 127–275.
603 31. Korpelainen H (2003) Human life histories and the demographic transition: A case
604 study from Finland, 1870–1949. *Am J Phys Anthropol* 120(4):384–390.
605 32. Pettay JE, Helle S, Jokela J, Lummaa V (2007) Natural selection on female life-history
606 traits in relation to socio-economic class in pre-industrial human populations. *PLoS*
607 *One* 2(7):e606.
608 33. Lummaa V, Jokela J, Haukioja E (2001) Gender difference in benefits of twinning in
609 pre-industrial humans: boys did not pay. *J Anim Ecol* 70(5):739–746.
610 34. Faurie C, Russell AF, Lummaa V (2009) Middleborns disadvantaged? Testing birth-
611 order effects on fitness in pre-industrial Finns. *PLoS One* 4(5):e5680.
612 35. Lummaa V, Lemmetyinen R, Haukioja E, Pikkola M (1998) Seasonality of births in
613 Homo sapiens in pre-industrial Finland: maximisation of offspring survivorship? *J*
614 *Evol Biol* 11(2):147–157.
615 36. Burnham KP, Anderson DR (2002) *Model Selection and Multimodel Inference: A*
616 *Practical Information Theoretic Approach* (Springer, New York). 2nd Ed.
617 37. Gagnon A, Mazan R (2009) Does exposure to infectious diseases in infancy affect old-
618 age mortality? Evidence from a pre-industrial population. *Soc Sci Med* 68(9):1609–16.
619 38. Myrskylä M (2010) The effects of shocks in early life mortality on later life
620 expectancy and mortality compression: A cohort analysis. *Demogr Res* 22:289–320.
621 39. Myrskylä M (2010) The relative effects of shocks in early life mortality on later life
622 expectancy and mortality compression: a cohort analysis. *Popul Dev Rev* 36(4):803–
623 829.

624 40. Bruckner TA, Catalano RA (2009) Infant mortality and diminished entelechy in three
625 European countries. *Soc Sci Med* 68(9):1617–24.
626 41. Bengtsson T, Broström G (2009) Do conditions in early life affect old-age mortality
627 directly and indirectly? Evidence from 19th-century rural Sweden. *Soc Sci Med*
628 68(9):1583–90.
629 42. Preston SH, Hill ME, Drevenstedt GL (1998) Childhood conditions that predict
630 survival to advanced ages among African–Americans. *Soc Sci Med* 47(9):1231–1246.
631 43. Myrskylä M, Gagnon A, Bengtsson T (2014) Pathways to health and well-being. *Soc*
632 *Sci Med* 119:175–179.
633 44. Quaranta L (2014) Early life effects across the life course: the impact of individually
634 defined exogenous measures of disease exposure on mortality by sex in 19th- and
635 20th-century Southern Sweden. *Soc Sci Med* 119:266–73.
636 45. Costa DL (2000) Understanding the twentieth-century decline in chronic conditions
637 among older men. *Demography* 37(1):53–72.
638 46. Gurven M, et al. (2009) Inflammation and infection do not promote arterial aging and
639 cardiovascular disease risk factors among lean horticulturalists. *PLoS One* 4(8):e6590.
640 47. Placek CD, Quinlan RJ (2012) Adolescent fertility and risky environments: a
641 population-level perspective across the lifespan. *Proc Biol Sci* 279(1744):4003–8.
642 48. Painter RC, Roseboom TJ, Bleker OP (2005) Prenatal exposure to the Dutch famine
643 and disease in later life: An overview. *Reprod Toxicol* 20(3):345–352.
644 49. Rickard IJ, et al. (2010) Food availability at birth limited reproductive success in
645 historical humans. *Ecology* 91(12):3515–3525.
646 50. Hayward AD, Rickard IJ, Lummaa V (2013) Influence of early-life nutrition on
647 mortality and reproductive success during a subsequent famine in a preindustrial
648 population. *Proc Natl Acad Sci* 110(34):13886–13891.

51. Störmer C Sex differences in the consequences of early-life exposure to epidemiological stress--a life-history approach. *Am J Hum Biol* 23(2):201–8.
52. Ibanez L (2003) Hypergonadotrophinaemia with reduced uterine and ovarian size in women born small-for-gestational-age. *Hum Reprod* 18(8):1565–1569.
53. Sheppard P, Pearce MS, Sear R (2015) How does childhood socioeconomic hardship affect reproductive strategy? Pathways of development. *Am J Hum Biol*. doi:10.1002/ajhb.22793.
54. Helle S, Lummaa V, Jokela J (2004) Accelerated immunosenescence in preindustrial twin mothers. *Proc Natl Acad Sci USA* 101:12391–12394.

## FIGURE LEGENDS

**Figure 1:** Changes in child mortality from infectious disease in each year, 1751-1855 and calculation of our measure of early-life disease exposure. Child (aged 0-10) mortality rate from infectious disease varied substantially across the early lives of our study individuals. Grey points and lines show the proportion of children alive in a given year that died of an infectious disease; black line shows the 'trend' component of Hodrick-Prescott decomposition of the time series. Data and trends were calculated and are shown separately for (A) all parishes combined; (B) Hiittinen; (C) Ikaalinen; (D) Jaakkima; (E) Kustavi; (F) Pulkkila; (G) Rautu; (H) Rymattyla.

**Figure 2:** The deviations from the trend in child mortality from infectious disease in the early years of our study individuals' lives, represented as the 'cycle' component of the Hodrick-Prescott decomposition. The parish-specific deviations from the trend, our measure of early-life disease exposure, varied substantially across cohorts in (A) all parishes combined; (B) Hiittinen; (C) Ikaalinen; (D) Jaakkima; (E) Kustavi; (F) Pulkkila; (G) Rautu; (H) Rymattyla.