Corresponding author

Laust Hvas Mortensen

Post: Øster Farimagsgade 5, Postboks 2099, 1014 København K, Denmark

Email: lamo@sund.ku.dk

Full title

Family of origin and educational inequalities in mortality: Results from 1.7 million Swedish siblings

Short title

Education and mortality among siblings

Authors

Laust H Mortensen, Department of Public Health, University of Copenhagen, and Jenny Torssander, Swedish Institute for Social Research, Stockholm University

Funding and grant-awarding bodies

This work was supported by the Swedish Council for Working Life and Social Research under Grant number 2014-0445 for Jenny Torssander, and by the University of Copenhagen for Laust H Mortensen.

Abstract (150 words):

Circumstances in the family of origin have short- and long-term consequences for people's health. Family background also influences educational achievements – achievements that are clearly linked to various health outcomes. Utilizing population register data, we compared Swedish siblings with different levels of education (1,732,119 individuals within 662,095 sibships) born between 1934 and 1959 and followed their death records until the end of 2012 (167,932 deaths).

The educational gradient in all-cause mortality was lower within sibships than in the population as a whole, an attenuation that was strongest at younger ages (< 50 years of age) and for those with working class or farmer background. There were substantial variation between different causes of death with clear reductions in educational inequalities in lung cancer and diabetes when introducing shared family factors. In contrast, educational inequalities in suicide and, for women, other mental disorders increased when siblings where compared.

Keywords: Education, mortality, siblings, family

Introduction

Family background has consequences for individuals' health and mortality risks over the life course. Circumstances in the family of origin also influence educational achievements, which in turn are associated with various health outcomes. This understanding provides motivation for intervening at the level of the family to improve the conditions of children, for example through parenting education programs or family support services. Recent randomized studies have provided empirical support for this approach by showing that early life interventions among disadvantaged children not only affect educational and economic outcomes in adulthood, but also appears to improve health later in life (Campbell, Conti, Heckman, Moon, Pinto, Pungello et al. 2014; Gertler, Heckman, Pinto, Zanolini, Vermeersch, Walker et al. 2014).

But how important is the family of origin for the educational gradient in mortality? Existing studies have examined the influence of single, fairly well-defined circumstances in the family of origin like childhood SEP (Galobardes, Lynch & Davey, 2004), but it is clear that adequately characterizing and obtaining information on all relevant features of the family of origin is difficult. When the outcome of interest is mortality, the challenges are compounded by the fact that the outcome typically occurs many years after the exposure. Comparing siblings with different educations provide an opportunity to potentially circumvent these problems. Since siblings share their family of origin, sibling comparisons will match out any measured or unmeasured characteristics that are shared. If the family of origin has any influence on the association between education and mortality, one might expect the within-family gradient to be weaker than the association in the population as a whole (as we will discuss later, this heuristic interpretation has limitations). Although there are different views of what counts as definitive evidence for the one and the other (Deaton, 2003), there are two classes of explanations that may be applied to the association between education and health: Educational causation and confounding. For example,

education might be associated with mortality because education causally affects health by providing access to safe, well-paid job, and by increasing specific health knowledge. On the other hand, ill-health or malnutrition in early life may confound the association between education and mortality. Both types of explanations can be invoked to explain the educational gradient between siblings. To generalize, public health researchers and sociologists have primarily placed emphasis on causation, while economists call attention to confounding.

Recent studies from the US and several Nordic countries have shown that an educational gradient in morbidity and mortality exists between siblings (Krieger, Chen, Coull & Selby, 2005; Madsen, Andersen, Christensen, Andersen & Osler, 2010; Naess, Hoff, Lawlor & Mortensen, 2012; Sondergaard, Mortensen, Nybo Andersen, Andersen, Dalton, Madsen et al. 2012). In addition, some of these studies also show that the relative educational gradient in mortality is less steep between siblings than in the population as a whole, which has been interpreted as evidence of familial confounding of the association between education and mortality. However, a recent Swedish study of twins shows that the relationships between years of schooling and life expectancy is fairly similar in the population and between twins (Lundborg, Lyttkens & Nystedt, 2012).

There are limitations to the existing literature. First, sibling comparisons come at a cost as only siblings that are discordant on education and age at death can contribute, which reduces precision and induce bias. This has in general received superficial treatment in the literature. Second, the existing studies have also been limited in their ability to examine specific causes of death, particularly for women. We believe that cause-specific mortality is of particular interest because it can increase the understanding of the intermediate mechanisms as well as confounding of the association between education and mortality. In particular, this provides us the opportunity to look at outcomes where we expect the sibling comparisons to be more strongly biased by confounding than the corresponding estimate from

the total population, specifically for suicides and other causes of death where a psychiatric disease is the underlying cause of death. Third, the role of the social class background has not been examined. This is of importance because the health-related selection into education or the causal effects of education may differ between working class and service class families. According to the social class literature, the association between education and class destination – i.e., how important own educational achievements are to the subsequent position in the labor market – is weaker for individuals from more advantaged backgrounds compared to those from less advantaged families (Goldthorpe, 2007). That education is more important for future labor market opportunities for individuals from less advantaged background could imply that individuals from the service class have resources beyond education (linked to their childhood social class) that may be utilized. It is therefore possible that the relationship between education and health is also weaker for individuals from advantaged socioeconomic conditions, but it is unclear how the influence from unmeasured family factors shared between siblings influences this.

Sweden and its neighbor-countries have traditionally shown weak associations between family background and labor market outcomes, in particular in comparison to the US (Solon , 2002). Sibling correlations in income have also declined in Sweden when comparing birth cohorts between 1930 and 1950 (Björklund & Jäntti, 2009). An educational expansion has taken place during the 20th century, increasing the opportunities to continuing to secondary and tertiary education (Erikson & Jonsson, 1996). One important change was the introduction of a comprehensive school system education in the mid- 20th century that increased labor market opportunities for individuals from less advantaged backgrounds in particular (Meghir & Palme, 2005). The longevity consequences of this reform is however not significant for the population as a whole, only in certain subgroups (Lager & Torssander, 2012). Still, the comparatively weak association between family background and socioeconomic outcomes in the case of Sweden should be noticed. Because family background seems to have less

impact compared to other countries, Sweden may be a conservative test case for examining the influence of family factors on the educational gradient in mortality.

Data

The basis for including individuals in this study was the Census of 1990, covering the entire Swedish population in November that year. The Multi-generation Register with information on parents was linked to the individuals in the Census to identify full biological siblings. The Multi-generation Register contains parental information for individuals born after 1932, but the parent-child linkage is more complete for individuals who were born in 1934 or later. We therefore include birth cohorts from 1934 and onwards, with the upper limit set to 1959 so that most individuals have completed their education before the start of the follow-up. The Multi-Generation Register and the validation procedures used have been described in detail elsewhere (Ekbom, 2011). Data on education was collected from the national Educational Register in 1990. Mortality follow-up was conducted using the Cause of Death Register from January 1991 to December 2012. The study population thus consisted of full biological siblings born 1934 to 1959, who was part of a sibling group of two or more siblings that were alive and living in Sweden in January 1991 (N=1,732,119). We also select a subpopulation born between 1944 and 1959 (N=1,153,730) for whom we can track information on the social class position during childhood (ages 1-15) via the Census of 1960.

Variables

Education

Five levels of highest attained education in 1990 were distinguished: Compulsory schooling (8-9 years), short upper-secondary education (mainly vocational, less than 3 years), academic upper-secondary

education (3-4 years), some tertiary education (less than 3 years) and tertiary exam (3 years or more, including postgraduate studies).

Mortality

Besides all-cause mortality, the following specific causes of death were analyzed separately: Ischeamic heart disease (IHD, ICD10: I20-25), lung cancer (C32-34), respiratory diseases (J00-99), diabetes (E10-14), some alcohol-related deaths (F10, K70, K73-74), mental diseases excluding alcohol abuse (F00-09; F11-99), suicide (X60-84), and other external causes of death (V01-Y89, excl. X60-84).

Social background

Social background was derived from parents' occupation and employment status when the individuals born 1944 to 1959 were between 1 and 15 years old (i.e., in year 1960). Occupation and employment status was combined to determine social class according to the EGP class schema (Erikson and Goldthorpe 1992). Thereafter, family social class was assigned in agreement with the dominance principle based on both parents' individual class position (Erikson 1984). In the analyses, we use the broader classes of manual occupations (47 %); non-manual occupations (26 %); self-employed (9 %); and farmers (12%). For 6 % of the individuals born between 1944 and 1959, there was no information of parents' occupation either because of missing data or because none of the parents were active in the labor market.

Statistical analyses

In order to describe the familial aggregation of education, we calculated for each index sibling in a sibling pair the distribution of education among their co-sibling. This metric is known as the probandwise concordance rate, and is the probability of the co-sibling having a certain education, conditional on the index sibling's education. We also calculated the ratio of observed concordance

divided by the concordance expected if education was randomly distributed across families, i.e. no familial aggregation. We calculated confidence intervals of the observed and expected proportions and ratios by bootstrapping (sampling from the 1,733,282 pairs with replacement under the observed and expected distributions of education, 2,000 repeats). Because of the vast number of pairs, all confidence intervals were very tight with all proportions varying less than one percentage point.

In the survival analyses, we used age in months as the time. Individuals were counted as under risk of dying from the age they had in January 1991 until the age of emigration, age of death, or age at the end of follow-up, which was December 2012. All analyses were adjusted for sex. The sex-specific estimates of the association between education and mortality were obtained by including an interaction term between sex and education. We conducted analyses for all-cause mortality and cause-specific mortality. We further stratified the analyses by age-group and by social class background. We conducted two sets of Cox proportional hazards regression models: with and without a family fixed effect. We will refer to these as population-based analyses and family-based, respectively. In the analyses without a family fixed effect (i.e., the population-based analyses) the model is of the form $\lambda_{ij}(t) = \lambda_0(t) \exp(\beta_1 x_{ij1} + \ldots + \beta_m x_{ijm})$, where $\lambda_{ij}(t)$ is the hazard for the *i*th sibling within the *j*th family, $\lambda_0(t)$ is the baseline hazard, β_m is the *m*th element of the vector of regression coefficients, and x_{ijm} is the *m*th element of the vector of covariates for the *i*th sibling in the *j*th family (i.e. dummy variables pertaining to sex and education). In these analyses, we took interdependence of siblings into account by using a robust variance estimator. In the analyses with a family fixed effect, the model is of the form $\lambda_{ij}(t) = \lambda_{0j}(t) \exp(\beta_1 x_{ij1} + \ldots + \beta_m x_{ijm})$. Note that the model now contains a separate baseline hazard $\lambda_{0j}(t)$ for every *j*th family. The family-specific baseline hazard ensures that comparisons are only made within families. Unlike shared frailty models, this model makes no assumption of independence between the shared component (i.e. the baseline) and covariates.

To quantify the importance of the family of origin, we compare the hazard ratios between the analyses with and without sibling fixed effects. There are, however, two mechanisms that serve to inflate the sibling fixed effects estimates, and thus counteract the attenuation. First, in a Cox regression the inclusion of any variable that is associated with the outcome may cause the estimates of other variables to inflate even in the absence of confounding, which has been referred to as non-collapsibility of the Hazard Ratio or as a non-linearity effect (Martinussen & Vansteelandt, 2013). Secondly, the estimates might also inflate due to non-shared confounding from collider stratification induced by design (Frisell, Oberg, Kuja-Halkola & Sjolander, 2012). To illustrate the influence of collider stratification bias in the sibling fixed effects analyses, we calculated if we could induce non-causal associations between sex of one sibling and mortality in the other by conditioning on education in both siblings. Our analyses take sex into account, but sex serves as a good test as on average there is little familial influence on sexes of siblings (Mortensen, Nielsen, Cnattingius & Andersen, 2011). The logic behind this is explained with directed acyclic graphs in the supplemental material.

Results

The study population included is fairly similar to the entire Swedish population born 1934 to 1959 (Table 1). Those excluded from the analyses were more likely to be born at the beginning of the period and lack information on variables of interest or to be born at the end of the period and not have any siblings.

Our analysis shows that familial aggregation of education was substantial (Table 2). In any sibling pair, an individual with compulsory education was 67% more likely to have a co-sibling with compulsory education and 73% less likely to have a co-sibling with a long tertiary education than expected if paired with a randomly selected individual from the study population. At the other end of the educational distribution, in any sibling pair an individual with a long tertiary education was 182% more likely to have a co-sibling with long tertiary education.

During follow-up, 167,932 individuals died. We found a strong a strong and monotone gradient in mortality over the educational categories for both men and women in the population-based analyses (Table 3). When siblings where compared, we found the highest educated sibling to have substantially lower mortality for both men and women, but the gradient was weaker in the family-based analyses than in the population-based with attenuations in log of the Hazard Ratios (HR) when estimating the per one-level increment difference in the association between education and mortality of 15% for men and 28% for women. The importance of the family of origin as measured by the attenuation in the HRs appeared to be larger at younger ages, where the mortality was relatively low and the relative educational differences in mortality high.

When cause-specific mortality was examined (Table 4), we observed substantial attenuation for ischemic heart disease, lung cancer, respiratory, and diabetes-related deaths. For women, the reduction was also clear for alcohol-related mortality. The attenuation was limited for other mental disorders and alcohol-related mortality among men, and deaths from other external causes for both sexes. The association between education and suicide was stronger in the family-based analyses than in the population-based, and a similar pattern occurred for mental disorders among women. The population-based educational gradient in suicide is however comparatively moderate among women.

When the all-cause mortality analyses where stratified by childhood social class for the individuals born between 1944 and 1959, the differences between the population- and family-based estimates vary across childhood social class (Table 5). Within non-manual families, the reduction of the educational gradient was only very marginal when siblings were compared (8 % for men, 2 % for women). Within manual families, self-employed and farmers, the attenuations of the educational gradients when shared family factors were introduced were much clearer (men 25 % and women 29 % if originating from a family when the parent(s) worked in a manual occupation).

Discussion

In this cohort of Swedish siblings, we observed a substantial familial aggregation of education. We also observed a strong educational gradient in mortality. This was true for men and for women. The gradient was present in the population as a whole, and within families. A sibling with a higher education than their co-siblings will have lower mortality than their less educated co-siblings. The magnitude of the difference between the population-based and family-based estimates depended on the cause of death. We observed substantially weaker estimates in the family-based than in the population-based analyses for mortality from heart disease, lung cancer, respiratory, and diabetes. For alcohol-related deaths, suicide and deaths from other mental disorders the associated were similar or even stronger within families than in the population-based analyses for at least one of the sexes. Childhood social class did moderate the all-cause mortality findings. There were only very marginal changes between the population and family-based educational mortality gradient for individuals from non-manual social classes, but clear reductions in the other classes.

The cause-specific patterns of attenuation might be explained by the differences in the intermediary causal pathways involved in the difference causes of death: The perhaps most clear example is that of lung cancer, where the causal intermediary is well known. The pattern of attenuation is consistent with the explanation that the smoking (initiation and/or cessation) is affected by family of origin above and beyond the social influences on these processes that comes with education. The role of smoking in explaining educational differences in mortality illustrates the complexities of causal explanation from observational studies: Smoking exposure is often taken to be a mediator between education and mortality (Mackenbach, 2011), but it is also part of a confounding pathway running from the family of origin to mortality (Gilman, Martin, Abrams, Kawachi, Kubzansky, Loucks et al. 2008). In fact, family

influences on smoking are likely to be a contributor to a number of the causes for which we find a weakened association between siblings.

We observed that among non-manual social class families, the difference between the population- and family-based estimates were small. There might be several explanations for this. The variation in confounders might be limited non-manual social class families as compared to other social class backgrounds. It is also possible that there is a comparatively strong health selection into lower educational levels within these advantaged families where higher education of the offspring is both more likely and expected. Such a selection would increase the family-based estimates and thus work in the opposite direction of familial confounding. Childhood social class by itself is associated with mortality with low risks for people with non-manual and farmer's background, and higher for manual and self-employed (not in Tables). This association is however greatly weakened when individual education is controlled for. The educational gradient in mortality, on the other hand, is rather unaffected by control for childhood socioeconomic position (a similar result can be found in Erikson 2001).

We have in the methods section briefly outlined the statistical problems with interpreting the attenuation in education-mortality association between the population-based and family-based estimates. Because of the problem of non-collapsibility of the HR the associations within families to inflate even in the absence of confounding. The most attractive solution to this problem Aalen's additive survival model where the subgroup and population-averaged associations are the same (Martinussen & Vansteelandt, 2013), but work is needed to implement a solution in software that works with covariates with very high degrees of freedom such as family of origin in this present study.

We want to stress that the educational gradient in mortality between siblings has to be interpreted with caution. Despite earlier enthusiasm (McGue, Osler & Christensen, 2010; Rutter, 2007), it is clear that

sibling comparisons do not mimic experiments, which are characterized by exogenous variation in exposure. Sibling comparisons are a form of imperfect matching, and come at the dual loss of precision (which may be less important in the context of this study) and bias away from the causal effect of education. One source of bias might arise from the fact that families where all siblings have the same education are informative for the population as a whole, but uninformative for the between sibling analyses. For example, it is an open (but unanswerable) question if the unobservable educational gradient in siblings in families 'doomed' to low education can be inferred from the educational gradient in families where the siblings actually have different educations. Using the association between the sexes of siblings as an example, we have also demonstrated empirically that restriction to discordant pairs will induce non-causal association that will bias the estimates away from the causal effect of education. This should be taken into account when interpreting the findings for causes of death, where the underlying conditions are likely to affect educational success and mortality, for example schizophrenia in the analyses of education and risk of suicide (Agerbo, Byrne, Eaton & Mortensen, 2004). Because familial clustering of education is very strong, the potential for bias is considerable: Having an education different from one's siblings is likely to have a reason. It may be possible to get an idea of the directionality of the bias. In general, we would expect the bias induced by design to work to increase the educational gradient in mortality because we believe that the majority of factors that limit educational success are likely to be detrimental to health if they have any effect on health at all. In theory one might imagine other scenarios: Being an excellent athlete might impede the chances of getting a long education while having a positive effect on health. A sibling may have been encouraged to get a long education because it was felt that he or she was too frail for a manual job, etc. However, we believe that the net result across families is likely to result in bias away from the null, but this assumption is contingent on the context-specific distribution of all contributory factors according to education and familial background, which is obviously not something that is easily documented.

In conclusion, our findings reaffirm that family of origin explains a part of the educational gradient in mortality. For some causes of disease, it seems likely that a causal effect of education may be an important contributor to the observed association. Strong educational gradients exist within families particularly for causes of death where the sibling design itself introduces confounding. Given the limitation of the design, the findings should be interpreted with caution, particularly as it is difficult to know how to balance the competing explanations of confounding and causation.

REFERENCES

Agerbo,E., Byrne,M., Eaton,W.W., & Mortensen,P.B. (2004). Marital and labor market status in the long run in schizophrenia. *Arch.Gen.Psychiatry*, *61*(1), 28-33.

Björklund,A., & Jäntti,M. (2009). Intergenerational income mobility and the role of family background. *Oxford Handbook of Economic Inequality, Oxford University Press, Oxford* 491-521.

Campbell,F., Conti,G., Heckman,J.J., Moon,S.H., Pinto,R., Pungello,E., & Pan,Y. (2014). Early childhood investments substantially boost adult health. *Science*, *343*(6178), 1478-1485.

Deaton,A. (2003). Health, Inequality, and Economic Development. *Jorunal of Economic Literature*, *151* 113-158.

Ekbom,A. (2011). The Swedish Multi-generation Register. In J.ed. Dillner (Ed.), Methods in Biobanking (pp.215-220). New York: Humana Press.

Erikson,R., & Jonsson,J.O. (1996). Explaining class inequality in education: The Swedish test case. *Can education be equalized* 1-63.

Frisell,T., Oberg,S., Kuja-Halkola,R., & Sjolander,A. (2012). Sibling comparison designs: bias from non-shared confounders and measurement error. *Epidemiology*, *23*(5), 713-720.

Galobardes,B., Lynch,J.W., & Davey,S.G. (2004). Childhood socioeconomic circumstances and cause-specific mortality in adulthood: systematic review and interpretation. *Epidemiol.Rev.*, *26* 7-21.

Gertler,P., Heckman,J., Pinto,R., Zanolini,A., Vermeersch,C., Walker,S., Chang,S.M., & Grantham-McGregor,S. (2014). Labor market returns to an early childhood stimulation intervention in Jamaica. *Science*, *344*(6187), 998-1001.

Gilman,S.E., Martin,L.T., Abrams,D.B., Kawachi,I., Kubzansky,L., Loucks,E.B., Rende,R., Rudd,R., & Buka,S.L. (2008). Educational attainment and cigarette smoking: a causal association? *Int.J.Epidemiol.*, *37*(3), 615-624.

Goldthorpe,J. (2007). On Sociology. Stanford, California: Stanford Univeristy Press.

Krieger,N., Chen,J.T., Coull,B.A., & Selby,J.V. (2005). Lifetime socioeconomic position and twins' health: an analysis of 308 pairs of United States women twins. *PLoS.Med.*, *2*(7), e162.

Lager,A.C., & Torssander,J. (2012). Causal effect of education on mortality in a quasi-experiment on 1.2 million Swedes. *Proc.Natl.Acad.Sci.U.S.A*, *109*(22), 8461-8466.

Lundborg,P., Lyttkens,C., & Nystedt,P. (2012). Human capital and longevity. Evidence from 50,000 twins.

Mackenbach,J.P. (2011). What would happen to health inequalities if smoking were eliminated? *BMJ*, *342* d3460.

Madsen,M., Andersen,A.M., Christensen,K., Andersen,P.K., & Osler,M. (2010). Does educational status impact adult mortality in Denmark? A twin approach. *American Journal of Epidemiology*, *172*(2), 225-234.

Martinussen,T., & Vansteelandt,S. (2013). On collapsibility and confounding bias in Cox and Aalen regression models. *Lifetime.Data Anal.*, *19*(3), 279-296.

McGue,M., Osler,M., & Christensen,K. (2010). Causal Inference and Observational Research. The Utility of Twins. *Perspectives on Psychological Science*, *5*(5), 546-556.

Meghir,C., & Palme,M. (2005). Educational reform, ability, and family background. *American Economic Review* 414-424.

Mortensen,L.H., Nielsen,H.S., Cnattingius,S., & Andersen,A.M. (2011). Sex of the First-born and Risk of Preterm Birth in the Subsequent Pregnancy. *Epidemiology*.

Naess,O., Hoff,D.A., Lawlor,D., & Mortensen,L.H. (2012). Education and adult cause-specific mortality--examining the impact of family factors shared by 871 367 Norwegian siblings. *Int.J.Epidemiol.*, *41*(6), 1683-1691.

Rutter,M. (2007). Proceeding From Observed Correlation to Causal Inference. The Use of Natural Experiments. *Perspectives on Psychological Science*, *2*(4), 377-395.

Solon ,G. (2002). Cross-Country Differences in Intergenerational Earnings Mobility. *Journal of Economic Perspectives*, *16*(3), 59-66.

Sondergaard,G., Mortensen,L.H., Nybo Andersen,A.M., Andersen,P.K., Dalton,S.O., Madsen,M., & Osler,M. (2012). Does shared family background influence the impact of educational differences on early mortality? *American Journal of Epidemiology*, *176*(8), 675-683.

Table 1 Descriptives of the entire Swedish population, born 1934 to 1959, alive and living in Sweden in the end of 1990. No. and % for individuals included and excluded in the study, respectively.

	Included		Excluded, no sibling in the defined population		Excluded, no info. on parents and/or education	
	N	%	N	%	N	%
Level of education						
Compulsory	528731	31	150914	27	238804	39
Shorter secondary (< 3 y)	561673	32	185715	33	179914	29
Longer secondary (>= 3 y)	191842	11	69389	12	71676	12
Shorter tertiary (<3 y)	215007	12	74178	13	60297	10
Longer tertiary incl postgrad	234866	14	80820	14	68021	11
Total	1732119	100	561016	100	618712	100
Siblings						
One	-	-	563830	100	n/a	
Two	816658	47	-	-	n/a	
Three	484119	28	-	-	n/a	
Four	228232	13	-	-	n/a	
Five or more	203110	12	-	-	n/a	
Total	1732119	100	563830	100	n/a	
Sex						
Male	884370	51	287409	51	330763	51
Female	847749	49	276421	49	322544	49
Total	1732119	100	563830	100	653307	100
Year of birth						
1934 to 1939	236006	14	101556	18	199144	30
1940 to 1944	342383	20	87462	16	144318	22
1945 to 1949	435020	25	101786	18	130258	20
1950 to 1954	397961	23	95840	17	96799	15
1955 to 1959	320749	19	177186	31	82788	13
Total	1732119	100	563830	100	653307	100
Mortality						
Alive in the end of 2012	1564187	90	504767	90	551986	84
Died 1991-2012	167932	10	59063	10	101321	16
Total	1732119	100	563830	100	653307	100

Table 2 Probandwise concordance of education in 1,741,172 sibling pairs.

		Index sibling's education										
		1		2		3		4		5		Expected
	Education	%[a]	Ratio[b]	%[a]	Ratio[b]	%[a]	Ratio[b]	%[a]	Ratio[b]	%[a]	Ratio[b]	%
Co-sibling's educaiton	1	50.85	1.67	34.36	1.13	23.19	0.76	19.01	0.62	11.07	0.36	30.53
	2	32.6	1.01	38.96	1.20	31.89	0.98	29.76	0.92	19.33	0.60	32.43
	3	6.68	0.60	9.69	0.88	15.18	1.37	13.24	1.20	12.77	1.15	11.08
	4	6.15	0.50	10.14	0.82	14.85	1.20	18.69	1.51	18.55	1.49	12.41
	5	3.72	0.27	6.85	0.51	14.89	1.10	19.3	1.42	38.29	2.82	13.56

1=compulsory; 2=short sec; 3=long sec; 4=short tertiary; 5=long tertiary

[a] The distribution of co-sibling education for an index sibling with a given education (given in the column above)

[b] Ratio of observed proportion vs. expected proportion under assumption of no familial aggregation of education. Numbers above 1 indicates a higher observed proportion than expected, numbers below 1 indicates lower proportion observed than expected. The expected distribution is given in the last column and corresponds to the marginal distribution of education in the study population.

Table 3 All-cause mortality HR (95 % CI)* for levels of education (and for education treated as a continuous variable) for men and women born 1934-1959. All ages pooled and three age groups separately.

		Men				Women			
		Population-based		Family-based		Population-based		Family-based	
	Education	HR	ci95	HR	ci95	HR	ci95	HR	ci95
All ages	Basic	2.10	2.05,2.15	1.91	1.84,1.99	1.92	1.86,1.98	1.63	1.56,1.70
	Secondary<3y	1.83	1.79,1.88	1.70	1.64,1.78	1.50	1.45,1.54	1.37	1.31,1.43
	Secondary ≥3y	1.37	1.34,1.41	1.34	1.28,1.40	1.30	1.25,1.36	1.30	1.22,1.38
	Tertiary<3y	1.18	1.15,1.22	1.20	1.14,1.26	1.07	1.03,1.11	1.08	1.03,1.14
	Tertiary ≥3y	1 (ref)		1 (ref)		1 (ref)		1 (ref)	
	Cont. education	1.21	1.21,1.22	1.18	1.17,1.19	1.19	1.18,1.20	1.13	1.12,1.14
Age group:									
31-49	Basic	2.76	2.56,2.98	2.29	2.02,2.60	2.13	1.95,2.33	1.65	1.44,1.90
	Secondary<3y	2.27	2.10,2.45	1.99	1.76,2.25	1.44	1.32,1.57	1.24	1.08,1.41
	Secondary ≥3y	1.52	1.39,1.67	1.45	1.26,1.67	1.21	1.07,1.37	1.30	1.10,1.54
	Tertiary<3y	1.14	1.03,1.27	1.12	0.96,1.30	0.99	0.89,1.10	0.97	0.84,1.13
	Tertiary ≥3y	1 (ref)		1 (ref)		1 (ref)		1 (ref)	
	Cont. education	1.31	1.29,1.33	1.25	1.21,1.28	1.23	1.21,1.26	1.14	1.11,1.18
50-64	Basic	2.20	2.13,2.27	1.97	1.88,2.07	1.94	1.87,2.01	1.65	1.56,1.74
	Secondary<3y	1.89	1.83,1.95	1.75	1.66,1.84	1.50	1.44,1.56	1.39	1.32,1.47
	Secondary ≥3y	1.40	1.35,1.45	1.35	1.28,1.43	1.32	1.25,1.39	1.27	1.18,1.37
	Tertiary<3y	1.17	1.12,1.22	1.22	1.15,1.30	1.07	1.03,1.13	1.11	1.04,1.18
	Tertiary ≥3y	1 (ref)		1 (ref)		1 (ref)		1 (ref)	
	Cont. education	1.23	1.22,1.24	1.19	1.17,1.20	1.19	1.18,1.20	1.13	1.12,1.15
65-78	Basic	1.79	1.72,1.86	1.62	1.49,1.76	1.87	1.78,1.97	1.59	1.44,1.74
	Secondary<3y	1.62	1.55,1.69	1.48	1.35,1.61	1.52	1.45,1.61	1.38	1.25,1.51
	Secondary ≥3y	1.28	1.22,1.34	1.23	1.13,1.35	1.30	1.20,1.41	1.36	1.18,1.56
	Tertiary<3y	1.25	1.19,1.32	1.20	1.08,1.33	1.09	1.02,1.17	1.05	0.94,1.18
	Tertiary ≥3y	1 (ref)		1 (ref)		1 (ref)		1 (ref)	
	Cont. education	1.15	1.14,1.16	1.13	1.11,1.15	1.18	1.17,1.19	1.13	1.11,1.15

*based on standard errors adjusted for family clustering.

Table 4 Cause-specific mortality HR of mortality for levels of education and % reduction in ln(HR) for continuous education between the cohort and the sibling models.

		Basic	Sec<3y	Sec ≥3y	Tert <3y	Tert ≥3y	Continuous education	% reduction
MEN								
All causes	Population-based	2.10	1.83	1.37	1.18	ref	1.21	
(167,185 deaths)	Family-based	1.91	1.70	1.34	1.20	ref	1.18	15
IHD	Population-based	2.76	2.31	1.62	1.28	ref	1.29	
(24,660 deaths)	Family-based	2.14	1.88	1.41	1.28	ref	1.21	27
Lung cancer	Population-based	2.66	2.20	1.58	1.22	ref	1.28	
(13,760 deaths)	Family-based	2.13	1.80	1.42	1.11	ref	1.22	21
Respiratory	Population-based	3.42	2.66	1.68	1.26	ref	1.38	
(6,964 deaths)	Family-based	2.54	2.17	1.40	1.05	ref	1.30	19
Diabetes	Population-based	3.19	2.69	1.66	1.27	ref	1.35	
(3,327 deaths)	Family-based	2.35	2.10	1.36	0.89	Ref	1.29	15
Alcohol-related	Population-based	3.95	3.57	2.01	1.41	ref	1.40	
(6,394 deaths)	Family-based	3.48	3.16	1.77	1.48	ref	1.35	10
Other mental	Population-based	3.56	3.16	1.58	1.48	ref	1.38	
(1,297 deaths)	Family-based	3.78	4.01	2.16	1.84	ref	1.34	9
Suicide	Population-based	1.73	1.59	1.09	1.13	ref	1.16	
(6,909 deaths)	Family-based	2.03	1.71	1.25	1.36	ref	1.18	-12
Other external	Population-based	2.52	2.38	1.48	1.25	ref	1.27	
(9,866 deaths)	Family-based	2.41	2.37	1.53	1.33	ref	1.24	10
WOMEN								
All causes	Population-based	1.92	1.50	1.30	1.07	ref	1.19	28
(167,185 deaths)	Family-based	1.63	1.37	1.30	1.08	ref	1.13	
IHD	Population-based	4.57	3.05	2.03	1.33	ref	1.49	19
(24,660 deaths)	Family-based	3.34	2.44	1.86	1.16	ref	1.38	
Lung cancer	Population-based	2.71	2.24	1.78	1.27	ref	1.28	46
(13,760 deaths)	Family-based	1.79	1.79	1.47	1.21	ref	1.14	
Respiratory	Population-based	4.85	2.92	1.92	1.34	ref	1.53	16
(6,964 deaths)	Family-based	3.36	2.16	1.67	1.01	ref	1.42	
Diabetes	Population-based	5.15	2.94	1.85	1.09	ref	1.60	36
(3,327 deaths)	Family-based	3.00	2.05	1.69	1.13	ref	1.35	
Alcohol-related	Population-based	3.62	2.45	2.40	1.21	ref	1.39	24
(6,394 deaths)	Family-based	3.24	2.62	2.50	1.57	ref	1.28	
Other mental	Population-based	2.47	1.59	1.01	0.77	ref	1.35	-19
(1,297 deaths)	Family-based	3.77	2.43	1.29	1.15	ref	1.43	
Suicide	Population-based	1.16	1.13	1.06	1.05	ref	1.04	-60
(6,909 deaths)	Family-based	1.39	1.38	1.45	1.30	ref	1.06	
Other external	Population-based	1.73	1.29	1.08	0.99	ref	1.17	12

(9,866 deaths)	Family-based	1.66	1.28	0.99	1.03	ref	1.15

Table 5 All-cause mortality HR (95 % CI)* for levels of education for men and women born *1944-1959*. Separate models by childhood social class (manual, non-manual, self-employed, farmers, no information/parents' not working).

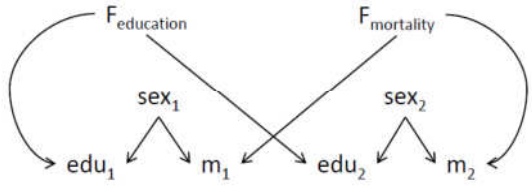
		Men				Women			
		Population-based		Family-based		Population-based		Family-based	
	Education	HR	ci95	HR	ci95	HR	ci95	HR	ci95
All born 1944-59	Basic	2.43	2.34,2.51	2.12	1.98,2.27	2.15	2.06,2.25	1.79	1.66,1.93
(N=1153730)	Sec<3y	2.04	1.96,2.11	1.87	1.74,1.99	1.54	1.48,1.61	1.41	1.31,1.52
	Sec ≥3y	1.46	1.40,1.53	1.44	1.34,1.55	1.28	1.21,1.35	1.27	1.16,1.39
	Tert<3y	1.17	1.12,1.23	1.21	1.12,1.32	1.08	1.03,1.14	1.12	1.04,1.22
	Tert ≥3y	1 (ref)		1		1		1	
	Cont. education	1.26	1.25,1.27	1.21	1.19,1.23	1.22	1.21,1.24	1.16	1.14,1.17
	% red in ln(HR) cont. education			*19%*			*29%*		
Manual families									
(N=544566)	Basic	2.38	2.22,2.54	1.94	1.72,2.18	2.10	1.94,2.27	1.81	1.58,2.06
	Sec<3y	2.02	1.89,2.16	1.77	1.57,1.99	1.52	1.40,1.65	1.41	1.24,1.61
	Sec ≥3y	1.40	1.30,1.51	1.29	1.13,1.46	1.23	1.11,1.36	1.22	1.04,1.43
	Tert<3y	1.16	1.06,1.26	1.15	1.00,1.32	1.06	0.97,1.17	1.16	1.00,1.34
	Tert ≥3y	1 (ref)		1		1		1	
	Cont. education	1.26	1.25,1.28	1.19	1.16,1.22	1.24	1.22,1.26	1.16	1.14,1.19
	% red in ln(HR) cont. education			*25%*			*29%*		
Non-manual families	Basic	2.82	2.65,2.99	2.66	2.35,3.01	2.17	2.00,2.35	2.16	1.87,2.50
(N=497218)	Sec<3y	2.05	1.93,2.17	2.03	1.81,2.27	1.57	1.46,1.68	1.60	1.41,1.81
	Sec ≥3y	1.55	1.46,1.66	1.69	1.51,1.91	1.37	1.25,1.50	1.54	1.32,1.79
	Tert<3y	1.20	1.12,1.30	1.34	1.18,1.52	1.08	1.00,1.17	1.15	1.01,1.31
	Tert ≥3y	1 (ref)		1		1		1	
	Cont. education	1.30	1.28,1.31	1.27	1.23,1.31	1.20	1.18,1.23	1.20	1.16,1.24
	% red in ln(HR) cont. education			*8%*				*2%*	
Self-employed	Basic	1.98	1.76,2.22	1.68	1.34,2.11	2.16	1.87,2.49	1.73	1.36,2.20
(N=106130)	Sec<3y	1.75	1.55,1.97	1.51	1.21,1.89	1.48	1.29,1.70	1.33	1.05,1.67
	Sec ≥3y	1.35	1.19,1.55	1.32	1.03,1.68	1.26	1.05,1.51	1.14	0.85,1.52
	Tert<3y	1.04	0.89,1.21	0.96	0.74,1.25	1.02	0.87,1.20	1.03	0.80,1.34
	Tert ≥3y	1 (ref)		1		1		1	
	Cont. education	1.20	1.18,1.23	1.16	1.10,1.22	1.23	1.19,1.27	1.15	1.09,1.22
	% red in ln(HR) cont. education			*22%*				*30%*	
Farmers	Basic	1.87	1.64,2.13	1.60	1.26,2.04	1.87	1.62,2.16	1.44	1.14,1.82
(N=132762)	Sec<3y	1.51	1.32,1.73	1.33	1.05,1.69	1.38	1.21,1.59	1.25	1.01,1.55
	Sec ≥3y	1.20	1.02,1.40	1.09	0.83,1.42	1.13	0.93,1.38	0.97	0.73,1.30
	Tert<3y	1.07	0.90,1.26	1.04	0.78,1.38	1.22	1.04,1.44	1.16	0.91,1.48
	Tert ≥3y	1 (ref)		1		1		1	
	Cont. education	1.19	1.16,1.23	1.15	1.09,1.20	1.15	1.12,1.19	1.08	1.03,1.14
	% red in ln(HR) cont. education			*23%*				*44%*	
Missing	Basic	2.73	2.34,3.19	1.88	1.38,2.57	2.24	1.86,2.69	1.30	0.90,1.88
(N=73344)	Sec<3y	2.38	2.04,2.79	1.67	1.22,2.27	1.64	1.36,1.97	1.02	0.71,1.46
	Sec ≥3y	1.54	1.29,1.85	1.34	0.94,1.89	1.20	0.93,1.54	0.95	0.60,1.51

	Tert<3y	1.15	0.94,1.42	1.09	0.75,1.60	0.98	0.78,1.24	0.66	0.43,1.01
	Tert ≥3y	1 (ref)		1		1		1	
	Cont. education	1.30	1.26,1.34	1.18	1.11,1.26	1.27	1.22,1.32	1.15	1.07,1.24
	% red in ln(HR) cont. education			*37%*				*40%*	

*based on standard errors adjusted for family clustering

SUPPLEMENT

Figure 1 DAG



This DAG suggests that the sex_1 and sex_2 are not marginally associated. It also suggests that the educations of the siblings (edu_1 and edu_2) are associated due to familial aggregation of education $F_{education}$. The mortality of the siblings (m_1 and m_2) are associated because of familial aggregation of mortality $F_{mortality}$ and $F_{education}$, the later effect mediated by a causal effect of education on mortality in both siblings. The DAG also assumes that sex has a causal influence on educational and mortality in in both siblings.

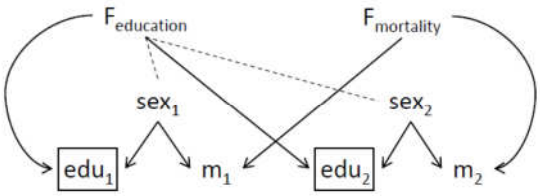
To illustrate this, we selected all individuals from two-sibling groups where both siblings were born 1945-1949, who could be followed for mortality until the age of 63 (n = 177 060). For simplicity, we dichotomized education into low (compulsory, shorter secondary) and high (longer secondary, tertiary), and mortality in to alive and dead before age 63. We then calculated risk ratios as measures of the marginal associations in the DAG in the figure above.

Since we randomly decided which sibling was sibling one and which was sibling two, we will disregard this information below where symmetrical associations exist. The sex of one sibling was also not associated with the education of the other sibling (RR = 1.00, 95% confidence interval (95% CL): 0.99-1.01, p = ns).The relative risk (RR) of one sibling being a women if the other sibling was a women was 1.01 (95% CL: 1.00-1.02, p = ns). This suggests no association as expected from the DAG above. Having a sister was weakly related with mortality before age 63 in the other sibling with a RR of 1.02 (95% CL: 1.01-1.04, p=0.01), which maybe a product of the repeated testing or might reflect some causal process.

Confirming the aggregation shown in Table X, high education in one sibling was associated with high education in the other sibling (RR = 2.49, 95% CL: 2.45-2.54). A high level education in one sibling was also strongly related to mortality in the other sibling (RR = 0.76, 95% CL: 0.72-0.80), and if one sibling died before age 63, then the other sibling was also at increased risk (RR = 1.59, 95% CL: 1.46-1.73).

To examine if we can induce an association between the sexes of the siblings by conditioning on the co-siblings education, we revised the DAG.

Figure 2 Revised DAG



The sibling fixed effect involves using only siblings that are discordant with regards to education. This is equivalent to conditioning on the education of the other sibling. The revised DAG now includes square boxes around education to indicate conditioning, and dotted lines between sex and $F_{education}$ to indicate that an association will be induced as a result of conditioning on education. The logic is that because education is a collider on the pathways between sex and $F_{education}$, conditioning on it opens a pathway between sex of one sibling and the sex of the other sibling: $sex_1 \rightarrow edu_1 \leftarrow F_{education} \rightarrow edu_2 \leftarrow sex_2$. This induces a non-causal association, which will bias associations because any association with the sex of the other sibling will also induce non-causal association with the mortality of the other sibling.

If the siblings had the same education (long-long or short-short), the relative risk of the other sibling having the same sex as the index sibling was 1.05 (95% CL: 1.04-1.07). This situation is, however, not of problem per se as such sibling pairs would not be informative for the fixed effects estimator. If the siblings had different lengths of education, the corresponding relative risk was 0.91 (95% CL: 0.89-0.94). If we did not have information on sex, we would inadvertently get confounding from sex in the sibling fixed effects analyses. This means that the conditioning that occurs by design can cause non-causal associations, which will result in bias of the causal effect of education.

A different way of phrasing the association induced by adjustment is to think of sex as a cause of the difference in exposure. Even if sex is assigned completely at random to each sibling, the difference in education between the siblings will be associated with the difference in sex as sex is a cause of education. This is trivial for sex, which we do take account of in the analyses, but by extension this applies to all unmeasured factors that act common causes even if they are randomly assigned, so it will apply to all conceivable confounders that affect education and mortality (school environment, IQ, gestational age, early life morbidity, genes, peer influences, etc.).

Appendix tables

Table A Education differences between siblings according to sibship size.

Educational difference	Two siblings		Three siblings		Four siblings		Five or more siblings		All		Expected	Observed vs. expected
	N	%	N	%	N	%	N	%	N	%	%	Ratio
1 1 - 1	46967	12	68417	14	62103	18	128178	25	305665	18	9	1.88
2 1 - 2	72114	18	97957	20	79709	23	142078	28	391858	23	20	1.14
3 1 - 3	18609	5	22472	5	16515	5	22767	4	80363	5	7	0.68
4 1 - 4	16459	4	20612	4	15006	4	21819	4	73896	4	8	0.56
5 1 - 5	11120	3	13311	3	9103	3	11218	2	44752	3	8	0.31
6 2 - 2	52556	13	60541	13	43640	13	65436	13	222173	13	11	1.21
7 2 - 3	30888	8	33170	7	21111	6	25381	5	110550	6	7	0.88
8 2 - 4	31680	8	34428	7	22135	6	27420	5	115663	7	8	0.83
9 2 - 5	22706	6	24831	5	14927	4	15698	3	78162	4	9	0.51
10 3 - 3	8895	2	8560	2	4609	1	4251	1	26315	2	1	1.23
11 3 - 4	16602	4	16951	4	9091	3	8811	2	51455	3	3	1.07
12 3 - 5	17503	4	17601	4	9294	3	7215	1	51613	3	3	0.99
13 4 - 4	11998	3	11643	2	6531	2	6141	1	36313	2	2	1.35
14 4 - 5	24578	6	25749	5	13951	4	10716	2	74994	4	3	1.28
15 5 - 5	25654	6	27876	6	14623	4	9247	2	77400	4	2	2.42
Total no. pairs	408329	100	484119	100	342348	100	506376	100	1741172	100	100	-

1=compulsory; 2=short sec; 3=long sec; 4=short tert; 5=long tert

Table B Mortality differences between siblings according to sibship size.

Mortality difference	Two siblings		Three siblings		Four siblings		Five or more siblings		All	
	N	%	N	%	N	%	N	%	N	%
0 - 0	339986	83.26	398557	82.33	276241	80.69	395508	78.11	1410292	81.00
0 - 1	62810	15.38	78648	16.25	60616	17.71	101016	19.95	303090	17.41
1 - 1	5533	1.36	6914	1.43	5491	1.60	9852	1.95	27790	1.60
Total no. pairs	408329	100	484119	100	342348	100	506376	100	1741172	100

0=alive December 2012; 1=died before December 2012

Table C: HR (95% CI) for cause-specific mortality by educational level (reference: tertiary>=3 years).

		MEN				WOMEN			
		Cohort		Sibling		Cohort		Sibling	
		HR	ci95	HR	ci95	HR	ci95	HR	ci95
IHD	Basic	2.76	2.60,2.93	2.14	1.91,2.39	4.57	4.00,5.23	3.34	2.76,4.05
(24,660 deaths)	Sec <3	2.31	2.17,2.46	1.88	1.68,2.11	3.05	2.66,3.50	2.44	2.01,2.95
	Sec>=3	1.62	1.51,1.73	1.41	1.25,1.59	2.03	1.69,2.44	1.86	1.45,2.39
	Tert<3	1.28	1.18,1.39	1.28	1.12,1.47	1.33	1.12,1.57	1.16	0.93,1.45
Cerebrovascular	Basic	2.26	2.03,2.52	1.99	1.64,2.41	2.67	2.31,3.09	2.14	1.71,2.68
(8,115)	Sec <3	1.88	1.68,2.10	1.68	1.38,2.05	1.96	1.69,2.27	1.64	1.31,2.05
	Sec>=3	1.40	1.23,1.59	1.32	1.07,1.62	1.46	1.18,1.80	1.62	1.20,2.21
	Tert<3	1.19	1.03,1.39	1.27	1.01,1.61	1.25	1.04,1.50	1.25	0.97,1.62
Lung cancer	Basic	2.66	2.42,2.92	2.13	1.81,2.50	2.71	2.45,3.00	1.79	1.52,2.11
(13,760)	Sec <3	2.20	1.99,2.44	1.80	1.53,2.11	2.24	2.02,2.48	1.79	1.53,2.11
	Sec>=3	1.58	1.42,1.77	1.42	1.20,1.69	1.78	1.55,2.05	1.47	1.19,1.81
	Tert<3	1.22	1.07,1.40	1.11	0.91,1.35	1.27	1.12,1.45	1.21	1.01,1.45
Breast cancer	Basic					0.97	0.90,1.05	1.29	1.07,1.57
(6,858)	Sec <3					0.88	0.81,0.94	1.15	0.96,1.37
	Sec>=3					1.02	0.92,1.14	1.19	0.96,1.48
	Tert<3					0.95	0.87,1.04	1.10	0.91,1.31
Pancreas cancer	Basic	1.46	1.29,1.66	1.19	0.95,1.49	1.74	1.50,2.03	1.72	1.36,2.19
(5,211)	Sec <3	1.38	1.20,1.57	1.22	0.97,1.53	1.56	1.34,1.82	1.59	1.26,2.00
	Sec>=3	1.19	1.02,1.38	1.00	0.78,1.27	1.34	1.08,1.66	1.46	1.06,2.01
	Tert<3	1.05	0.88,1.25	0.99	0.74,1.30	1.22	1.01,1.46	1.27	0.97,1.65
Colon cancer	Basic	1.18	1.04,1.33	1.20	0.97,1.50	1.29	1.12,1.48	1.09	0.86,1.39
(4,941)	Sec <3	1.02	0.89,1.17	1.10	0.88,1.38	1.24	1.08,1.43	1.12	0.89,1.41
	Sec>=3	0.96	0.83,1.12	1.08	0.85,1.38	1.16	0.94,1.42	1.16	0.84,1.60
	Tert<3	1.01	0.86,1.20	1.07	0.81,1.40	1.00	0.84,1.19	0.93	0.72,1.21
Prostate cancer	Basic	1.32	1.18,1.47	1.27	0.95,1.71				
(3,474)	Sec <3	1.15	1.02,1.30	1.12	0.83,1.52				
	Sec>=3	1.17	1.02,1.33	1.15	0.85,1.55				
	Tert<3	1.18	1.01,1.37	1.02	0.72,1.46				
Respiratory	Basic	3.42	2.96,3.95	2.54	1.98,3.25	4.85	4.07,5.78	3.36	2.55,4.42
(6,964)	Sec <3	2.66	2.29,3.09	2.17	1.68,2.80	2.92	2.44,3.50	2.16	1.64,2.84
	Sec>=3	1.68	1.42,1.99	1.40	1.06,1.84	1.92	1.51,2.45	1.67	1.15,2.43
	Tert<3	1.26	1.03,1.54	1.05	0.76,1.45	1.34	1.07,1.67	1.01	0.72,1.40
Diabetes	Basic	3.19	2.66,3.81	2.35	1.71,3.24	5.15	3.79,7.00	3.00	1.97,4.57
(3,327)	Sec <3	2.69	2.24,3.24	2.10	1.52,2.89	2.94	2.15,4.02	2.05	1.34,3.12
	Sec>=3	1.66	1.35,2.05	1.36	0.96,1.92	1.85	1.21,2.83	1.69	0.96,2.97
	Tert<3	1.27	0.99,1.63	0.89	0.59,1.34	1.09	0.72,1.63	1.13	0.67,1.90
Alcohol-related	Basic	3.95	3.46,4.50	3.48	2.75,4.41	3.62	2.86,4.58	3.24	2.27,4.63
(6,394)	Sec <3	3.57	3.12,4.08	3.16	2.49,4.00	2.45	1.93,3.11	2.62	1.84,3.74
	Sec>=3	2.01	1.73,2.34	1.77	1.39,2.26	2.40	1.79,3.23	2.50	1.65,3.81
	Tert<3	1.41	1.18,1.68	1.48	1.12,1.96	1.21	0.90,1.63	1.57	1.04,2.37
Mental	Basic	3.56	2.57,4.93	3.78	2.15,6.64	2.47	1.74,3.50	3.77	1.89,7.52
(1,297)	Sec <3	3.16	2.26,4.42	4.01	2.24,7.16	1.59	1.11,2.28	2.43	1.20,4.91
	Sec>=3	1.58	1.07,2.32	2.16	1.08,4.35	1.01	0.57,1.78	1.29	0.48,3.45
	Tert<3	1.48	0.96,2.29	1.84	0.92,3.68	0.77	0.47,1.25	1.15	0.51,2.59
Suicide	Basic	1.73	1.57,1.91	2.03	1.70,2.41	1.16	0.99,1.35	1.39	1.11,1.75
(6,909)	Sec <3	1.59	1.44,1.76	1.71	1.44,2.03	1.13	0.97,1.31	1.38	1.12,1.72
	Sec>=3	1.09	0.97,1.23	1.25	1.03,1.51	1.06	0.86,1.31	1.45	1.09,1.92
	Tert<3	1.13	0.99,1.29	1.36	1.11,1.68	1.05	0.88,1.25	1.30	1.02,1.65
Other external	Basic	2.52	2.30,2.77	2.41	2.05,2.83	1.73	1.51,1.98	1.66	1.35,2.04
(9,866)	Sec <3	2.38	2.17,2.62	2.37	2.02,2.79	1.29	1.12,1.48	1.28	1.05,1.56
	Sec>=3	1.48	1.33,1.66	1.53	1.29,1.83	1.08	0.89,1.32	0.99	0.75,1.30
	Tert<3	1.25	1.10,1.41	1.33	1.09,1.62	0.99	0.84,1.17	1.03	0.82,1.30

Note: Cancer types with more than 3,000 deaths in the study population are included in the above table.

Table D: The effect of social background. Population-based HR of mortality for individuals born 1944-1959 (No sibling design).

	Men						Women					
	Model 1		**Model 2**		**Model 3**		**Model 1**		**Model 2**		**Model 3**	
	HR	ci95	HR	ci95	HR	ci95	HR	ci95	HR	ci95	HR	ci95
Childhood SEP												
Manuals	1.32	1.29,1.36			1.04	1.01,1.07	1.23	1.20,1.27			1.02	0.99,1.05
Non-manuals (ref)	1				1		1				1	
Self-employed	1.12	1.08,1.17			0.94	0.90,0.97	1.06	1.01,1.11			0.95	0.91,1.00
Farmers	0.97	0.94,1.01			0.76	0.73,0.79	0.92	0.87,0.96			0.81	0.77,0.84
Not in labor mark	1.60	1.54,1.66			1.27	1.22,1.32	1.41	1.34,1.48			1.17	1.11,1.23
Education												
Compulsory			2.43	2.34,2.51	2.43	2.34,2.52			2.15	2.06,2.25	2.12	2.03,2.22
Short secondary			2.03	1.96,2.11	2.02	1.95,2.10			1.54	1.48,1.61	1.54	1.47,1.60
Long secondary			1.46	1.40,1.53	1.46	1.40,1.52			1.28	1.21,1.36	1.27	1.20,1.35
Short tertiary			1.17	1.12,1.23	1.17	1.12,1.23			1.08	1.03,1.14	1.08	1.03,1.14
Long tertiary			1		1				1		1	
N	590714		590714		590714		563016		563016		563016	

Table E: HR for education level and per one-step increase, respectively (individuals born 1944-1959). Stratified by childhood SEP (manual/non-manual) and sex.

	Men				Women			
	Manual		**Non-manual**		**Manual**		**Non-manual**	
	HR	ci95	HR	ci95	HR	ci95	HR	ci95
Education								
Compulsory	2.38	2.22,2.54	2.82	2.65,3.00	2.10	1.94,2.27	2.17	2.00,2.35
Short secondary	2.02	1.88,2.16	2.05	1.93,2.18	1.52	1.41,1.65	1.57	1.46,1.68
Long secondary	1.40	1.30,1.51	1.55	1.46,1.66	1.23	1.11,1.36	1.36	1.24,1.50
Short tertiary	1.15	1.06,1.26	1.21	1.12,1.30	1.06	0.97,1.17	1.08	1.00,1.17
Long tertiary	1		1		1		1	
Per one-step increase	1.26	1.25,1.28	1.30	1.28,1.31	1.20	1.18,1.23	1.24	1.22,1.26
N	277591		153364		266975		143564	