Premature Death of Adult Adoptees: Analyses of a Case-Cohort Sample

Liselotte Petersen,1,* Per Kragh Andersen,2,3 and Thorkild I. A. Sørensen1

1Danish Epidemiology Science Centre at Institute of Preventive Medicine, Copenhagen University Hospital, Copenhagen K, Denmark
2Danish Epidemiology Science Centre, Statens Serum Institut, Copenhagen S., Denmark
3Department of Biostatistics, University of Copenhagen, Copenhagen N, Denmark

Genetic and environmental influence on risk of premature death in adulthood was investigated by estimating the associations in total and cause-specific mortality of adult Danish adoptees and their biological and adoptive parents. Among all 14,425 non-familial adoptions formally granted in Denmark during the period 1924 through 1947, we selected the study population according to a case-cohort sampling design. As the case-control design, the case-cohort design has the advantage of economic data collection and little loss in statistical efficiency, but the case-cohort sample has the additional advantages that rate ratio estimates may be obtained, and re-use of the cohort sample in future studies of other outcomes is possible. Analyses were performed using Kalbfleisch and Lawless's estimator for hazard ratio, and robust estimation for variances. In the main analyses the sample was restricted to birth years of the adoptees 1924 and after, and age of transfer to the adoptive parents before 7 years, and age at death was restricted to 16 to 70 years. The results showed a higher mortality among adoptees, whose biological parents died in the age range of 16 to 70 years; this was significant for deaths from natural causes, vascular causes and all causes. No influence was seen from early death of adoptive parents, regardless of cause of death. Genet. Epidemiol. 28:376–382, 2005. © 2005 Wiley-Liss, Inc.

Key words: survival analysis; Cox proportional hazard; Kalbfleisch and Lawless estimator; cause-specific deaths

INTRODUCTION

The heredity of longevity has been of interest for centuries. It has been investigated in different settings and the presence of familial influence on the length of human lifetime is well established [Cohen, 1964; Glasser, 1981; Christensen and Vaupel, 1996]. Twin or adoption studies may be used to separate genetic and environmental effects and to investigate the strength of these effects. During the recent decades, registers mainly in the Nordic countries have made large twin studies possible [Hayakawa et al., 1992; McGue et al., 1993; Herskind et al., 1996; Iachine et al., 1998]. A twin study, based on registers from Denmark, Sweden, and Finland in which the resemblance in mortality among pairs of monozygotic twins and pairs of dizygotic twins was compared, supports the fact that there is a genetic influence on mortality [Iachine et al., 1998]. Twin studies have showed a genetic effect of death caused by vascular causes [Marenberg et al., 1994]. One critical assumption in the interpretation of results from such twin studies is that monozygotic twin pairs do not have a more similar environment than dizygotic twin pairs [Neale and Cardon, 1992].

Non-familial adoption studies, based on information on biological and adoptive relatives where the adoptee is transferred early in life to an unrelated adoptive family, provide an alternative approach to separate assessment of genetic and environmental influences. This is achieved by comparing the associations of the (cause-specific) mortality of adult adoptees with (1) the (cause-specific) mortality of their biological parents and (2) the (cause-specific) mortality of their adoptive parents [Borch-Johnsen and Sørensen, 1993].

The unique Danish register of 14,425 non-familial adoptions, carried out during 1924–1947, contains information on biological and adoptive relatives. Earlier smaller studies based on this
register showed a moderate genetic influence on risk of premature death in adulthood, which most likely stems mainly from a genetic influence on the risk of dying from vascular causes [Sørensen et al., 1988; Petersen et al., 2002]. A genetic effect on death due to infections was found only in the cohort study; one hypothesis is the time trends in pattern of death from infections and type of infections. The present case-cohort study of Danish adoptees was carried out as a continuation of the adoption studies in order to investigate the genetic and environmental influences on early cause-specific death and death in total in further detail and with increased statistical power.

MATERIALS AND METHODS

THE ADOPTION REGISTER

The study was based on the Danish Adoption Register, which contains records on all non-familial adoptions formally granted in Denmark during the period 1924 through 1947 [Kety et al., 1968; Jacobsen and Schulsinger, 1981]. The register was established in 1963–64, with schizophrenia as a primary interest. The decision not to let it cover a wider period was due to local civil registers being initiated in 1924, making it easier to trace the families; by not going further than 1947, all adoptees had reached the age of 17, before which schizophrenia is unlikely to occur. Also, the number of adoptions decreased after World War II. The records, compiled by the Psykologisk Institut (former name of the Institute of Preventive Medicine) at Kommunehospitalet in Copenhagen, provided the name and date of birth of each adoptee and his or her biological and adoptive parents, date of transfer to adoptive parents, and date of formal adoption, as well as addresses at the time of adoption. Among the identified biological fathers, paternity was usually acknowledged by the fathers themselves; if not, it was confirmed by biological tests, as required by the law. From these initial studies, all adoptees were followed in the local civil registers through all changes in address and marital status, to death, emigration, loss of follow-up, or end of study.

CASE-COHORT SAMPLING

The case-cohort sampling, proposed by Prentice [1986], consists of a random sample from the cohort, called a sub-cohort, and, in addition, all cases in the entire cohort occurring outside the sub-cohort. Covariate information is then collected in this case-cohort sample, rather than in the entire cohort. As the case-control design, the case-cohort design has the advantage of saving time in data collection, and the loss in statistical efficiency is of the same or of smaller magnitude than for an equally sized case-control sample [Prentice, 1986]. The case-cohort design has the advantage that rate ratio estimates may be obtained, and it allows different time variables to be used in the analysis. Moreover, re-use of the random sub-cohort for further analyses of different outcomes of interest is possible.

Assuming that failures occur in the full cohort according to the Cox proportional hazards model, several estimators of the covariate effect have been proposed. Based on a simulation study, designed to cover the situation of the adoption data [Petersen et al., 2003], we chose to use Kalbfleisch and Lawless’s [Kalbfleisch and Lawless, 1988] estimator for hazard ratios. The idea for Kalbfleisch and Lawless’s estimator is inverse probability weighting. Cases, whether or not they occur in the random sub-cohort, are given weight 1, and non-cases in the sub-cohort are given a weight equal to the inverse sampling probability for a non-case in the cohort to be in the sub-cohort sample. Due to the over-sampling of cases, the usual estimator of variance will overestimate the precision of the regression parameter estimate. From simulations, the robust estimator of variances was found to perform well [Barlow, 1994; Petersen et al., 2003].

STUDY POPULATION

We included those adoptees born in 1924 or later, who were traceable, who reached at least age 16, and who were transferred to the adoptive family before the age of 7 years (see flowchart in Fig. 1 for details). This gives a sample of 12,301 adoptees of which 1,403 had died before April 1, 1993. For these 1,403 cases and for a random sub-cohort of 1,683 adoptees, the adoptive and biological parents were traced through the regionally organised civil registers. When necessary, this information was supplemented from censuses, parish registers, and the National Archives. The overlap between the 1,403 cases and the random sub-cohort was 203 adoptees dying within the sub-cohort, i.e., the study populations comprise 2,883 adoptees. In planning the study, power was evaluated based on simulations of Gompertz distributed lifetimes and Prentice’s estimator of hazard ratios. The study was designed with size of
the sub-cohort approximately equal to the number of deaths such that the power to find a hazard ratio of 1.3 regarding death in total exceeded 80%. Correspondingly, the power exceeds 80% to find hazard ratios of 1.3 regarding natural deaths, of ratios 1.5 regarding death with cancer or vascular causes, and ratios of 2.0 regarding death with infections.

FOLLOW-UP

Biological and adoptive parents were followed up till they died or reached 70 years of age. In the final sample of families, unidentified parents and parents lost to follow-up due to emigration reduced the number of lifetimes of adoptees for analysis to 2,426 in the analysis of biological families, and to 2,787 in the analysis of adoptive families. The number of adoptees included in both analyses of biological and adoptive parents was 2,365.

ASCERTAINMENT OF THE CAUSES OF DEATH

As previously reported [Sørensen et al., 1988], we considered death in total, or death from natural causes, infections, vascular causes, or cancer (“death from natural causes” comprises all deaths except death due to accidents, suicides, and homicides). The official death certificates hold information on cause of death. These are obtained in the Death Cause Register for deaths occurring after 1969, at the National Health Service for deaths during 1943–67, and in paper files at the National Archives for deaths before 1943. The death certificates may list up to three causes of death determined on the basis of information available at the time of death, possibly including autopsy. For all deaths, each cause was coded according to the international classification system used by Danish health authorities from 1942. Deaths occurring before 1942 were recoded according to that system. The cause of interest was identified irrespective of its position on the death certificate, which may not reflect an unambiguous relation between causes of death; thus, some case subjects were included in more than one cause-specific analysis. Among adoptees, 125 were not registered with a death cause and could not be included in the cause-specific analyses. Numbers of non-cases and cases for death from all causes and cause-specific deaths are listed for adoptees and parents in Table I. Note that the numbers of non-cases are higher for the cause-specific analyses; these further adoptees are those selected for the random sub-cohort who died before April 1 1993, but without the specific cause.

In contrast to earlier studies [Sørensen et al., 1988; Petersen et al., 2002], gastric ulcer is now coded as an infectious disease in our cause-specific analyses. To explore the differences in earlier findings on genetic effects of death with infections, time trends of infections were analysed and possible sub-divisions of infections were studied. The sub-divisions considered included also gastrointestinal, genito-urinary, nerve system infections, and tuberculosis. Only for respiratory infections was the number of deaths among adoptees adequately large to be analysed.

STATISTICAL METHODS

In a Cox model, lifetimes of adoptees were analysed from age 16, stratified by gender and in 7 bands according to birth cohort of adoptees. The proportional hazards assumptions were tested.
Based on Schoenfeld’s residuals [Schoenfeld, 1982; Grambsch and Therneau, 1994] and could not be rejected. Biological and adoptive parents were analysed separately. The risk of dying, among adoptees with a parent that died before age 70 years, was compared with the risk among adoptees with both parents alive. The variable constructed to estimate this hazard ratio took the value 1 if a parent had died from the cause of interest before age 70 years (irrespective of what happened to the other parent) or the value 0 if both parents were known to be alive at age 70 years. Consequently, an adoptee was excluded from a particular analysis if the information on the parents was insufficient for a value to be assigned the variable, i.e., if one parent was untraceable or had died of causes not of interest before age 70 years and, therefore, censored, and the other parent was still alive at age 70 years. For the estimation, we used the Kalbfleisch and Lawless’s hazard ratio estimator [Kalbfleisch and Lawless, 1988], which is implemented by ascribing cases weight 1 and non-cases weight equal to the inverse probability of being in the sub-cohort (=10,898/1,480 for death from all causes) and the robust variance estimator [Barlow, 1994].

In addition, the case-cohort was viewed as a cohort study with a sample of “exposed” biological parents sharing genes with a case adoptee, and a sample of “non-exposed” biological parents sharing genes with a living non-case adoptee. Similarly, we considered exposed adoptive parents sharing environment with a case adoptee and non-exposed adoptive parents sharing environment with a living non-case adoptee. Having this view of the data, the exact lifetime of adoptees is not taken into account, and, therefore, we need to restrict the non-exposed to non-cases known to be alive at time of selection of the cases. A Cox proportional hazards model with age as time scale [Andersen et al., 1993] was used to compare the survival of biological parents to a case adoptee with the survival of biological parents to a non-case adoptee, and similarly for adoptive parents. Parents of the same adoptee made up a cluster, allowing for possible mutual dependence, and therefore robust standard error estimates [Lin and Wei, 1989] were used. In these analyses, we expect a gain in efficiency, firstly due to nearly twice as many survival times in each analysis, and, secondly, because every traceable parent can be included, in contrast to the above described analyses, irrespectively of exclusion of the other parent before age 70 years.

Analyses were carried out using the computer package Stata version 8, in which weighting procedures in the Cox models and the robust estimator of variance are directly available.

## RESULTS

As seen in Table II, we found a significant hazard ratio of 1.27 (95% CL 1.06,1.52) when at least one biological parent had died before age 70 years, compared to those having both parents known to be alive at age 70 years. The equivalent hazard ratio of 0.93 (95% CL 0.79,1.09) for adoptive parents was not significantly different from 1. The cause-specific analyses showed significantly increased hazard ratios if biological parents died of the same cause before age 70 years regarding natural and vascular causes. Death with infections or respiratory infections also showed an increased hazard ratio, though not significant. None of the analyses of adoptive parents showed any relation to lifetimes of adoptees. Changing the covariate to measure death between 16 and 60 or 16 and 80 years, the same effects turn out to be significant, with a tendency that hazard ratios for the adoptees dying with vascular causes or from all causes combined increase with advancing age.

<table>
<thead>
<tr>
<th>TABLE I. Number of non-case and case adoptees and deaths among parents</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Non-cases in the</strong></td>
</tr>
<tr>
<td><strong>sub-cohort</strong></td>
</tr>
<tr>
<td><strong>adoptees</strong></td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>Natural</td>
</tr>
<tr>
<td>Infections</td>
</tr>
<tr>
<td>Respiratory</td>
</tr>
<tr>
<td>Vascular</td>
</tr>
<tr>
<td>Cancer</td>
</tr>
</tbody>
</table>
at death of the biological parents. Analysing parents' lifetimes, using adoptees' death as covariate, gave the same overall picture, and with more narrow confidence limits as expected (Table III). Compared to Table II, the only difference is that death with infections did not show any effect on the risk of dying of the biological parents. The time trends of infections were analysed by allowing the effect to be different for the 7 different groups of the adoptees based on their birth year, which showed significantly higher hazard ratios in the birth cohort of 1930 and earlier (data not shown).

Testing whether the effect of having a case child is different for fathers and mothers, showed no significance difference for either biological or adoptive parents (data not shown). The effect of early death did not significantly depend on the gender of the adoptee (data not shown). Age at death of the adoptee, dichotomised at age 50 years, was found not to significantly modify the effect of early death, though with a tendency to increased hazard ratios among biological parents for death from vascular causes and all causes when the adoptee died after age 50 years (data not shown).

One reason to have a child adopted is that either one, or both, biological parents have died. In our case-cohort data, the time of transfer was later than death of a parent for 49 non-case children in the sub-cohort, 11 cases in the sub-cohort, and a further 59 case children not already in the sub-cohort. Data were re-analysed excluding these adoptions, which are probably due to the early death of a biological parent; this did not change results (data not shown).

**DISCUSSION**

The overall impression is that the risks of premature death in adulthood are affected by the genetic background, whereas the familial environmental influences are weaker or absent. Significant associations were found for biological parents and their adopted children for deaths between age 16 and 70 years from natural causes, vascular causes, and from all causes combined.

The finding of a moderate genetic influence on the risk of dying prematurely in adulthood agrees with results from twin research [Iachine et al., 1998]. The effect of family history of heart disease

<table>
<thead>
<tr>
<th>TABLE II. Effects of parents' death between age 16 and 70 years on risk of death of adoptees using Kalbfleisch &amp; Lawless's estimator and robust variance estimates, stratified by gender, birth cohort of adoptees (7 groups), and parents' birth cohorts (3 groups)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biological parents (at least one)</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>Natural</td>
</tr>
<tr>
<td>Infections</td>
</tr>
<tr>
<td>Respiratory</td>
</tr>
<tr>
<td>Vascular</td>
</tr>
<tr>
<td>Cancer</td>
</tr>
</tbody>
</table>

*Hazard ratios (HR), 95% confidence limits in parentheses, and P values.

<table>
<thead>
<tr>
<th>TABLE III. Effect of adoptees' death after age 16 years, and occurring before April 1, 1993, on risk of death of parents between age 16 and 70 years using Cox's model with clusters and robust variance estimates, stratified by gender of parents and adoptees, birth cohort of adoptees (7 groups), and parents' birth cohorts (3 groups)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biological parents (16–70 years)</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>Natural</td>
</tr>
<tr>
<td>Infections</td>
</tr>
<tr>
<td>Respiratory</td>
</tr>
<tr>
<td>Vascular</td>
</tr>
<tr>
<td>Cancer</td>
</tr>
</tbody>
</table>

*Hazard ratios (HR), 95% confidence limits in parentheses, and P values.
is seen in several cohorts. A Swedish twin study strongly suggests that the familial aggregation is due to genetic effect, which our findings support [Marenberg et al., 1994]. It is particularly noteworthy that the twin study revealed the genetic influence on risk of heart disease in the same age range as in the present study, whereas there was no such effect among the older twins, but with no sign of an increasing part being genetic with increasing age. The finding of no genetic or environmental effects on death with cancer is in contrast to genetic and environmental effects for some types of cancer, as found in twin studies [Lichtenstein et al., 2000], but this could not be further investigated in the present study due to too few cases of death with cancer. We found the genetic influence on death with infections in cohorts of adoptees born in 1930 or earlier, but no effect for later cohorts. This is in agreement with differences in findings from the 1924–26 cohort study [Sørensen et al., 1988] and the case-control study [Petersen et al., 2002]. This could be caused by the type of infections, e.g., a genetic component for tuberculosis [Comstock, 1993], malaria and/or other infection, as found in twin studies [Kwaitkowski, 2000], but with the limited number of deaths with infections in our data this could not be further explored.

Valid assessment to separate of genetic and environmental influences using the adoption method requires the adoptees to be separated from the biological parents shortly after birth and transferred without delay to adoptive parents unrelated to the child. This condition is reasonably met in our study: all were non-familial adoptions, more than 90% were transferred to the adoptive family before the age of 2 years, and from a validation study it is known that most of the children not transferred immediately after birth did not remain with their biological parents, but were transferred to foster homes [Eldred et al., 1976].

For the Kalbfleisch & Lawless estimator, censored observations are weighted with the inverse probability of being in the sub-cohort. This probability may change with time, which could be taken into account by using information on lifetimes of all members in the full cohort. Some gain in efficiency might be obtained by using such time-dependent weights, although the gain is likely to be slight [Borgan et al., 2000].

In the cause-specific death analyses, case adoptees are compared to the same sub-cohort, as this is the point in case-cohort sampling; this creates a correlation between estimated effects [Sørensen and Andersen, 2000]. With the size of our cohort and the sample size used, this issue is unlikely to be of importance, as seen in the simulation study [Sørensen and Andersen, 2000].

It would be more satisfactory to avoid dichotomisation of both parental and adoptees’ lifetimes in one analysis, i.e., handling bivariate censored lifetimes. Therefore, we plan to explore the field of shared frailty models in connection with a case-cohort sample to further optimize the use of the exceptional dataset.

Although we found evidence of genetic influences by studying parent-offspring relationships, these effects may have been underestimated if they are of a non-additive nature, such as in intra-locus allele-allele interaction (recessive effects) or inter-locus gene-gene interaction (epistatic effects). These effects may be addressed by studying the biological full- and half-siblings of the adoptees [Borch-Johnsen and Sørensen, 1993; Sørensen, 1995]. A plausible reason for lack of support of familial environmental influences on early death could be that such effects require shared environment at the same time in life, which may be weak for the parent-child relationship. This could be better investigated by studying adoptive siblings of the adoptees.

In conclusion, this study supports that there is a moderate genetic influences on premature death in adulthood, which most likely originates from the genetic influence on the risk of dying from vascular causes.

ACKNOWLEDGMENTS

We are indebted to Marianne Zimmermann Sørensen for her effort in data collection; to Claus Holst for building up the database; to Knud Juel for his search of the death register at the Danish Institute of Clinical Epidemiology; to Jenny Bua for sub-divisions of infections. The Danish Epidemiology Science Centre is supported by The Danish National Research Foundation.

REFERENCES


