

ESTIMABILITY AND ESTIMATION OF EXCESS AND ETIOLOGIC FRACTIONS

JAMES M. ROBINS

Occupational Health Program, Harvard School of Public Health, 665 Huntington Avenue Boston, MA 02115, U.S.A.

AND

SANDER GREENLAND

Division of Epidemiology, UCLA School of Public Health, Los Angeles, CA 90024, U.S.A.

SUMMARY

This paper describes conditions under which epidemiologic data can provide estimates of the excess fraction (proportionate increase in caseload due to an exposure) and the etiologic fraction (fraction of cases caused by exposure). The excess fraction can be estimated under essentially the same conditions often cited for general study validity. In contrast, estimation of the etiologic fraction will usually require very specific non-identifiable assumptions about exposure action and interactions, although one can derive simple lower and upper bounds for the fraction from survival comparisons. Since the etiologic fraction is equivalent to the probability of causation, our results have implications for injury compensation in lawsuits involving the probability of causation.

KEY WORDS Attributable fraction Attributable risk Probability of causation Assigned share
Epidemiologic methods

INTRODUCTION

Estimation of the impact of an exposure on disease incidence and life expectancy is an important aspect of public health and health services research. The analogous process of estimating the impact of exposure on individual disease and survival experience is an important aspect of determining compensation for adverse exposure effects on health. In both situations there is a need for clarification of concepts and the conditions under which desired quantities can be estimated.

We have recently pointed out¹ that the most commonly used measure of public-health impact, the attributable fraction (also called attributable proportion,² etiologic fraction^{3,4} and attributable risk⁵) must actually be separated into three distinct measures in order to make any sense of current usage. One measure, which we call the excess fraction, is the proportionate increase in caseload produced by exposure over a specified time interval. Another measure, which we call the etiologic fraction, is the fraction of cases whose disease originated from a pathologic process in which exposure had an effect. A third measure, which we call the hazard fraction or incidence-density fraction, is $[h_1(t) - h_0(t)]/h_1(t) = [R(t) - 1]/R(t)$, where $h_1(t)$ and $h_0(t)$ are the hazard rates that the study population would experience at time t under exposure and non-exposure, and $R(t) = h_1(t)/h_0(t)$. We have documented¹ that these measures can be very far from one another in epidemiologic settings, and yet textbooks often equate the underlying concepts. The present paper describes conditions under which excess and etiologic fractions can be

0277-6715/89/070845-15\$07.50

© 1989 by John Wiley & Sons, Ltd.

Received March 1988

Revised December 1988

estimated. To obtain further background for the present results, we recommend the reader to Reference 1.

A full appreciation of the distinction and conditions for estimability is important because of the growing application of attributable fractions to compensation for exposure-induced disease (see, for example, References 6 and 7). For example, the probability that a randomly sampled exposed case occurring at time t was caused by exposure is equivalent to the etiologic fraction among the exposed cases occurring at that time.³ It has been suggested that when uncertainty exists as to the cause of a particular individual's disease, award levels in toxic tort suits be based on this 'probability of causation'. It is commonly held that, in the absence of bias, one can equate the probability of causation for a particular exposed case occurring at time t with the hazard fraction $[R(t) - 1]/R(t)$ (see, for example, Reference 3). Nevertheless, it is not difficult to show that the hazard fraction at t is not in general equal to the probability of causation.¹ Here we characterize those special cases in which the hazard fraction does equal the probability of causation, show that under certain plausible biological assumptions the hazard fraction is a lower bound for the probability of causation, and derive some general mathematical relationships between the two quantities. Elsewhere we examine the estimability of measures of exposure impact on life expectancy, as such measures may often provide a more rational basis for evaluating the impact of exposure on health.⁸

NOTATION AND BASIC DEFINITIONS

Consider a population of individuals indexed by i , each subject to a binary point exposure (for example, exposure to an atomic blast). Suppose that each subject has a disease occurrence time d_{1i} when exposed, and another (unobserved) disease occurrence time d_{0i} when unexposed, with time measured from time of exposure (we consider a stochastic version of this model in Reference 9). For convenience, let $d_{1i} = \infty$ if disease would never occur for individual i when exposure is present, and similarly define $d_{0i} = \infty$. Finally, suppose each individual is followed up until either disease occurs or a maximum follow-up time t_i is reached.

The t_i may or may not be the same across individuals. For example, in a study of intrapartum deaths after oxytocin administration, t_i is end of labour and so varies across individuals; whereas in a study of neonatal deaths (without loss-to-follow-up), t_i is by definition 28 days postpartum for everyone. Define.

- $A_0 = B_0 =$ number of individuals who experience the outcome by t_i and for whom exposure had no impact on occurrence time, that is, $d_{1i} = d_{0i} < t_i$;
- $A_1 = B_1 =$ number of individuals for whom exposure shortened time to occurrence, but would have experienced the outcome by t_i even if exposure was absent, that is, $d_{1i} < d_{0i} \leq t_i$;
- $A_2 =$ number of individuals for whom exposure resulted in the outcome occurring by t_i , that is, $d_{1i} \leq t_i < d_{0i}$;
- $A_3 = B_3 =$ number of individuals for whom exposure lengthened time to occurrence, but the outcome still occurred by t_i , that is, $d_{0i} < d_{1i} \leq t_i$;
- $B_2 =$ number of individuals for whom exposure resulted in the outcome *not* occurring by t_i , that is, $d_{0i} \leq t_i < d_{1i}$; and
- $A_+ = \sum A_k$ and $B_+ = \sum B_k$.

A_+ is simply the observed number of cases that occur in the exposed (study) population over the follow-up period. B_+ is the number of cases that would have occurred in the same (exposed) population had exposure been absent (or its effect been blocked).

ESTIMABILITY AND ESTIMATION OF THE EXCESS FRACTION

The *excess fraction* over the follow-up period is simply the proportionate excess caseload $(A_+ - B_+)/A_+ = (A_2 - B_2)/A_+$ if this quantity is non-negative (otherwise we employ the prevented fraction $(B_2 - A_2)/B_+$).

Note that the excess fraction's numerator does *not* involve A_1 or A_3 , despite exposure's role in the etiology of cases contributing to A_1 and A_3 . Even so, the excess fraction can be a useful measure of exposure effect. For example, suppose the exposure under study is oxytocin administration, and the outcome is intrapartum death. Then, from a public-health standpoint, the primary issue in evaluating the safety of the exposure is whether treatment elevated or reduced the final number of deaths, not when in the course of labour the deaths occurred.

Alone among the measures discussed here, the criteria for estimability of the excess fraction in an exposed fixed cohort of initial size N_1 are identical to the usual criteria for estimating IPR, the incidence-proportion ratio (that is, the risk ratio, or proportionate increase in average risk), since the IPR is simply $(A_+/N_1)/(B_+/N_1) = A_+/B_+$, and so $(IPR - 1)/IPR = (A_+ - B_+)/A_+$. For non-randomized comparisons of exposed and unexposed cohorts, these criteria are simply lack of misclassification, differential censoring, or confounding at time of exposure.

By no confounding at time of exposure, we mean that the proportion who develop disease by any time t would have been the same in both the exposed and unexposed cohorts had neither cohort been exposed,¹⁰ that is, the proportion who develop disease by t_i in the unexposed cohort is B_+/N_1 . This condition is likely to be approximated in a large randomized trial. In a non-randomized study, one tries to select the unexposed group such that this condition is approximated overall or within levels of covariates measured at $t=0$. Henceforth, we shall suppose that we always have access to an unexposed control group for which there exists no confounding at time of exposure.

We may relax the requirement of non-differential censoring to one of independent differential censoring if we employ a life-table method (such as Kaplan-Meier) to estimate the incidence proportions. Henceforth, we shall suppose that, conditional on exposure, any right censoring is independent of risk, that is, for any times s and t , the exposed and unexposed incidence proportions over (s, t) among individuals who survive to s do not depend on censoring status.

The excess fraction is not a useful measure if *when* the outcome occurs during the follow-up period is important. To take an extreme example, suppose one wishes to evaluate the impact of smoking on total mortality in the cohort of U.S. males born in 1860, and $t_i = 120$ years, that is, individuals are followed throughout their lives. Then the excess fraction of deaths among smokers is zero, because by the end of follow-up everyone (smoker and non-smoker) is dead, leaving no room for any excess. More generally, the excess fraction is always zero (and is thus a vacuous measure) whenever the outcome inevitably occurs before t_i , regardless of exposure status.

ESTIMABILITY AND ESTIMATION OF THE ETIOLOGIC FRACTION

A non-vacuous measure in the last example is the fraction of all deaths among smokers in which smoking was a contributory cause. We call this quantity the causal etiologic fraction, or simply the *etiologic fraction*. For the remainder of the paper we will assume that exposure is never preventive so that d_{1i} is always less than or equal to d_{0i} and $B_2 = B_3 = 0$. Under this assumption, the etiologic fraction equals $(A_1 + A_2)/A_+$.

One major problem with the etiologic fraction is that it is not identifiable (estimable) in most settings without the use of strong biological assumptions.¹ In other words, even with random error and biases (for example, confounding at time of exposure) entirely absent, a set of

observations can be compatible with more than one value for the etiologic fraction. Suppose, for example, we observe three exposed individuals with occurrence times of 1, 2, and 3 years, and three unexposed individuals with occurrence times of 2, 3, and 4 years. Even if the unexposed group's distribution of occurrence times is exactly what the exposed group would have shown had exposure been absent, there is still no way to tell from these data whether exposure advanced all three exposed occurrence times by one year, so that the etiologic fraction is 1.0, or exposure advanced one person's occurrence time by 3 years and had no effect on the other two, so that the etiologic fraction is 1/3.

We will examine this point in a more general setting. Suppose that each pair of death times (d_{0i}, d_{1i}) for each subject under study is a realization from a bivariate distribution with absolutely continuous marginal distributions for d_{0i} and d_{1i} . We denote the marginal survival distributions for the d_{0i} and d_{1i} by $S_0(u)$ and $S_1(u)$. $S_1(u)$ is directly estimable from observations on the exposed cohort; in the absence of confounding at time of exposure, $S_0(u)$ is estimable from observations on the unexposed comparison cohort. If the observations are subject to left or right censoring, $S_1(u)$ and $S_0(u)$ remain estimable provided that the censoring processes are independent of disease occurrence (conditional on exposure). As the preceding example showed, however, the etiologic fraction depends on the joint distribution of (d_{0i}, d_{1i}) , and this joint distribution is not identifiable without further assumptions. We now consider estimation of the etiologic fraction under various assumptions about this joint distribution (which correspond to assumptions about the mechanism of exposure action). The discussion focuses on comparisons of exposed and unexposed closed cohorts, but the results extend immediately to situations that involve censoring (although in such cases one must estimate survival distributions and derived quantities by censored-data methods).

Suppose the survival distribution the exposed cohort would have had in the absence of exposure is equal to the (actual) survival distribution the exposed cohort actually had up to a time t_0 (the minimum induction time) and is strictly greater thereafter, as one would expect of a purely causal exposure. Without loss of generality, we can restrict attention to the subcohort that survived to t_0 and translate the time axis so that time zero now corresponds to t_0 ; the preceding assumption then becomes

$$S_0(t) = Pr(d_{0i} > t) > Pr(d_{1i} > t) = S_1(t) \text{ if } t > 0 \text{ and } S_0(t) > 0. \quad (1)$$

Given this condition, without further assumptions one cannot reject the hypothesis that the underlying process preserves ranks of occurrence times, that is, for every pair of individuals i and j ,

$$d_{0i} < d_{0j} \text{ if and only if } d_{1i} < d_{1j}. \quad (2)$$

Conditions 1 and 2 taken together imply $d_{1i} < d_{0i}$ for all i , so that the etiologic fraction over any interval $(0, t)$ is 1. Thus any pair of exposed and unexposed distributions obeying condition 1 is compatible with the hypothesis that exposure caused every case occurring after the minimum induction time (in that exposure affected the occurrence time of every exposed case occurring after this time).

Now let $h_1(u)$ and $h_0(u)$ denote the hazard functions¹¹ that correspond to $S_1(t)$ and $S_0(t)$, and let $R(u) = h_1(u)/h_0(u)$. We will call $R(u)$ the *empirical hazard ratio* because it is estimable from a comparison of an exposed and unexposed group when there is no confounding at time of exposure. Suppose that, for all $u > 0$, $R(u) > 1$ (a condition which implies but is not implied by condition (1)). Then without further assumptions one cannot reject the independent competing-cause model, under which: (a) each person has a second potential occurrence time d_{ei} , defined as the time at which disease would occur from a cause that involved exposure if all causes that did *not* involve exposure (competing causes) were absent; and (b) the distribution of d_{ei} is independent of d_{0i} and has hazard $h_e(u)$. Under this model, $d_{1i} = \min(d_{0i}, d_{ei})$ and the etiologic fraction over an

interval (s, t) is simply the proportion of exposed cases that occur over the interval for whom $d_{ei} < d_{oi}$.

The conditional probability that exposure caused an exposed case that occurred at u (the 'instantaneous' etiologic fraction) is

$$p_e(u) = \Pr(d_{oi} > u | d_{1i} = u). \quad (3)$$

This quantity is often called the *probability of causation* for an exposed case that occurs at t . We saw above that, under a rank-preserving model, $p_e(u) = 1$ for $u > 0$. Under an independent competing-cause model,

$$\begin{aligned} p_e(u) &= \Pr(d_{ei} = u | d_{1i} = u) = h_e(u) / (h_e(u) + h_o(u)) \\ &= (h_1(u) - h_0(u)) / h_1(u) \\ &= (R(u) - 1) / R(u). \end{aligned} \quad (4)$$

Thus the hazard fraction $[R(u) - 1] / R(u)$, which is commonly thought to be equivalent to the probability of causation $p_e(t)$, is in fact smaller than $p_e(t)$ if a rank-preserving model holds, but is equivalent to $p_e(t)$ if an independent competing-cause model holds.

Now suppose that the expected number of exposed individuals under observation at time u is $N_1(u)$. Let $a_+(u) = h_1(u)N_1(u)$ be the expected number of exposed cases occurring per unit time at u , $A_+(s, t) = \int_s^t a_+(u)du$ the expected number of exposed cases that occur over the interval (s, t) when exposure is present, and $A_0(s, t) = \int_s^t h_0(u)N_1(u)du$. Regardless of the biological model, the expected number of exposed cases caused by exposure over (s, t) (that is, for which $d_{1i} < d_{oi}$ and $s < d_{1i} < t$) is

$$A_1(s, t) = \int_s^t p_e(u) a_+(u) du. \quad (5)$$

Under the independent competing-cause model,

$$A_1(s, t) = \int_s^t (h_1(u) - h_0(u)) N_1(u) du = A_+(s, t) - A_0(s, t). \quad (6)$$

The etiologic fraction over the interval is then $A_1(s, t) / A_+(s, t) = (\text{SMR} - 1) / \text{SMR}$, where we define SMR as $A_+(s, t) / A_0(s, t)$.

The preceding result extends to models other than the competing-causes model by noting that the only property of the model necessary for the above derivation is that $p_e(u) = [R(u) - 1] / R(u)$. The etiologic fraction equals $(\text{SMR} - 1) / \text{SMR}$ under any model that satisfies this assumption.

One may obtain a consistent estimate of the SMR parameter if one can obtain consistent estimates of $A_+(s, t)$ and $A_0(s, t)$. A consistent (and unbiased) estimate of $A_+(s, t)$ is O , the observed number of exposed cases occurring over (s, t) . One can obtain a consistent estimate of $A_0(s, t)$ by the traditional method of computing the 'expected number' of cases, E , over the person-time experience of the exposed. If $p_e(u) = (R(u) - 1) / R(u)$, then $A_0(s, t)$ is the number of cases expected to occur among the exposed due to causes other than exposure, SMR can be consistently estimated by O/E , and the corresponding etiologic fraction estimate is simply $(O - E) / O$.

We wish to make two cautionary points regarding the preceding results. First, $p_e(u)$ may fail to equal $(R(u) - 1) / R(u)$ and thus the SMR transform $(\text{SMR} - 1) / \text{SMR}$ can be far from the true etiologic fraction over (s, t) if the independent competing-cause model does not hold. Suppose for

example the effect of exposure (that is, the joint distribution of d_{0i} and d_{1i}) follows a strict accelerated-life model¹¹

$$\log d_{xi} = b_0 + b_1x + \varepsilon_i, \quad (7)$$

where the ε_i are independent identically distributed with mean zero, with $b_1 < 0$. Then $d_{1i} < d_{0i}$ for all i and so the etiologic fraction is 1 under this model (in particular, this model preserves ranks, as in condition (2)), but under this model if d_{0i} is exponentially distributed with mean T_0 , d_{1i} will be exponentially distributed with mean $T_0 \exp(b_1)$, the hazards under exposure and non-exposure will be the constants $1/T_0$ and $\exp(-b_1)/T_0$, and thus $R(u) = \exp(-b_1)$ and $(R(u) - 1)/R(u) = 1 - \exp(b_1) < 1$. In particular, if $b_1 = -\log 2$, $(R(u) - 1)/R(u) = 0.5$, which is clearly much less than the true etiologic fraction.

Second, even if equation (4) holds, only in special cases will $A_0(s, t)$ equal the true expected number of cases that would occur if exposure was absent,¹² despite the fact that it is often interpreted as such. To see this, note that if exposure affects survival, exposure will also affect the number of persons under observation. For example, suppose $d_{1i} < u < d_{0i}$. Then under exposure, individual i will not survive until u and so cannot be counted in $N_1(u)$, the exposed population-at-risk at u ; but if exposure is absent, individual i will survive past u and may contribute to the population-at-risk at u . Let $N_0(u)$ denote the expected size of the exposed population-at-risk at u if exposure effects were absent. The true expected number of cases among the exposed if exposure effects were absent is

$$E_0(s, t) = \int_s^t h_0(u)N_0(u)du. \quad (8)$$

If equation (1) holds, $N_1(t)$ will be strictly less than $N_0(t)$ (due to exposure-induced attrition of the population-at-risk), and consequently $E_0(s, t)$ will be strictly greater than $A_0(s, t)$. The excess fraction over the interval is $[A_+(s, t) - E_0(s, t)]/A_+(s, t)$, and this will be strictly smaller than $(SMR - 1)/SMR$ over the interval; thus, if equation (4) holds, the excess fraction will be strictly smaller than the etiologic fraction. Note, however, that, under equation (1), the excess fraction and $(SMR - 1)/SMR$ will approximate one another as the interval (s, t) is made smaller or if the exposure's effects are infrequent enough so that $N_1(u)$ approximates $N_0(u)$.

A lower bound for the etiologic fraction

We have seen that, if exposure results in a survival distribution $S_1(u)$ strictly less than what would have held in the absence of exposure, $S_0(u)$ (condition (1)), one cannot rule out that condition (2) holds and thus that the etiologic fraction is 1. We now derive the minimum possible etiologic fraction over the interval $(0, t)$ consistent with the marginal distributions $S_1(u)$ and $S_0(u)$ when these distributions satisfy condition (1). This minimum may be most easily computed using the unconditional occurrence-time densities $f_j(u) = h_j(u)S_j(u) = -dS_j(u)/du$, $j = 0, 1$. If exposure is never preventive, the minimum possible value for the etiologic fraction in a closed cohort under condition (1) is

$$\int_G [f_1(u) - f_0(u)]du / (1 - S_1(t)) \quad (9)$$

where G is the set of all $u < t$ such that $f_1(u) > f_0(u)$. To see this, note that the minimum etiologic fraction for subjects who fail at u when $f_1(u) \leq f_0(u)$ is zero, whereas the minimum when $f_1(u) > f_0(u)$ is $[f_1(u) - f_0(u)]/f_1(u)$. The etiologic fraction over $(0, t)$ must then be bounded below by $\int_0^t [\max(0, [f_1(u) - f_0(u)]/f_1(u))]a_+(u)du / A_+(0, t)$. This reduces to expression (9), since

$a_+(u) = h_1(u)N_1(u) = f_1(u)N_1(u)/S_1(u) = f_1(u)N_1(0)$ and $A_+(0, t) = [1 - S_1(t)]N_1(0)$. In Appendix I we show that there always exists a joint distribution for (d_{0i}, d_{1i}) that has marginal distributions $S_1(u)$ and $S_0(u)$ and for which the etiologic fraction attains the minimum just derived. (It is easy to show that the hazard fraction at u is always a possible value for the instantaneous etiologic fraction whenever the former is positive.)

Note that the excess fraction over $(0, t)$ is given by equation (9) modified so that the integral is evaluated over $(0, t)$ rather than G . It follows that, if c is the smallest value of t that satisfies $f_1(t) < f_0(t)$, the minimum etiologic fraction over $(0, t)$ equals the excess fraction over $(0, t)$ for all $t < c$ and exceeds the excess fraction for all $t > c$. (Here and throughout we have assumed that the densities are smooth functions of t .)

As an example, suppose that d_{0i} and d_{1i} are exponentially distributed with hazards h_0 and $h_1 = Rh_0$. Then $f_1(u) > f_0(u)$ if and only if $u < c = (\log R)/[(R - 1)h_0]$. Expression (9) then becomes

$$\int_0^{\min(c, t)} h_0 [R \exp(-Rh_0u) - \exp(-h_0u)] du / [1 - \exp(-Rh_0t)],$$

which is approximately 0.25 when $R = 2$, $h_0 = 0.05/\text{yr}$, and $t > 40$ years. One should compare this to the value of 1.0 that one would obtain under an accelerated-life model, and $0.50 = (2 - 1)/2$ that one would obtain under a competing-cause model.

The lower bound just derived is disappointingly low, and yet without further assumptions one cannot rule out values as low as the bound, even when there is no confounding at time of exposure.

Suppose now that the identifiable quantity $R(t)$ is a constant $R > 1$ over time. We have shown that, without making any non-identifiable assumptions, we can place no better upper bound on the etiologic fraction than 1 (since a rank-preserving model may hold). This upper bound is essentially vacuous, since the etiologic fraction by definition cannot exceed one. The situation for the lower bound for the lifetime etiologic fraction (that is, the etiologic fraction over $(0, \infty)$) is a little better: Although expression (9) approaches the logical lower bound of zero as R approaches 1, it approaches 1 as R goes to infinity. To see this, note that when $R(t) = R$ we have

$$\begin{aligned} f_1(t) &= Rh_0(t)S_0(t)^R \\ &= Rf_0(t)S_0(t)^{R-1}. \end{aligned} \tag{10}$$

Substitution of expression (10) into expression (9) and integration yield a lower bound of $S_0(c) - S_0(c)^R$, where c is the unique positive solution of $f_1(t) = f_0(t)$. Substitution of expression (10) into the latter equation and solving yields $c = S_0^{-1}(R^{-1/(R-1)})$, so that the lower bound becomes

$$R^{-1/(R-1)} - R^{-R/(R-1)} = (R - 1)/R^{R/(R-1)}. \tag{11}$$

Using l'Hopital's rule, we see that the lower bound given by expression (11) approaches zero as R approaches 1 and approaches 1 as R goes to infinity.

The following table compares the lower bound given by expression (11) to the hazard fraction $(R - 1)/R$:

R	1	2	4	8	16	∞
$(R - 1)/R$	0	0.50	0.75	0.88	0.94	1
Expression (11)	0	0.25	0.47	0.65	0.78	1

As can be seen, the (constant) hazard fraction is not close to either the lower bound (expression (11)) or the upper bound (of 1) of the lifetime etiologic fraction unless R is quite large. Later, we will

discuss conditions under which the hazard fraction may serve as a plausible lower bound for the etiologic fraction.

The etiologic hazard ratio and exposure-induced confounding

Let $N_1(u)$ now denote all persons in the exposed population at risk at u (in a closed cohort $N_1(u)$ is the set of all exposed persons such that $d_{1i} \geq u$), and let $h_0[u|i \in N_1(u)]$ denote what the hazard would be among persons in $N_1(u)$ if exposure (or its effect) had been absent. Also, let $N_0(u)$ denote all persons who were exposed and would have been at risk at u had they not been exposed (in a closed cohort, $N_0(u)$ is the set of all exposed persons such that $d_{0i} \geq u$); note that $h_0(u) = h_0[u|i \in N_0(u)]$ is the hazard this group would have had if exposure was absent. If $R_e(u) = h_1(u)/h_0[u|i \in N_1(u)]$, then under any model, the instantaneous etiologic fraction is

$$\begin{aligned} p_e(u) &= (h_1(u) - h_0[u|i \in N_1(u)])/h_1(u) \\ &= (R_e(u) - 1)/R_e(u). \end{aligned} \tag{12}$$

We will call $R_e(u)$ the *etiologic hazard ratio* among the exposed at time u . $R_e(u)$, although not identifiable, compares the actual hazard among the exposed at u , $h_1(u) = h_1[u|i \in N_1(u)]$, to what the hazard would have been among the *same* persons had they not been exposed, $h_0[u|i \in N_1(u)]$. In contrast, $R(u)$ (which is estimable if there is no confounding at time of exposure) compares $h_1(u)$, the rate of disease in $N_1(u)$ (that is, the rate among exposed subjects with $d_{1i} \geq u$) to $h_0(u)$, the rate of disease in $N_0(u)$ (that is, the rate among exposed subjects with $d_{0i} \geq u$). (Note that under the accelerated failure time model, $d_{1i} < d_{0i}$ for all individuals, and so

$$h_0[u|i \in N_1(u)] = \lim_{\delta \rightarrow 0} \Pr(d_{0i} \leq u + \delta | d_{0i} > d_{1i} \geq u) / \delta = 0$$

and $R_e(u) = \infty$.)

A difference between $R_e(u)$ and $R(u)$ can only come about if $h_0[u|i \in N_1(u)] \neq h_0(u)$; another way of saying this is that $R_e(u) \neq R(u)$ if and only if the hazards in $N_1(u)$ and $N_0(u)$ would differ. Since $R_e(u)$ represents the actual effect of exposure at time u , we will say that there is confounding at time u if $R_e(u) \neq R(u)$. If there was no confounding at the time of exposure (time zero) but there is at time $u > 0$, we will say that exposure has induced confounding at time u (due to exposure's effect on survival). The earlier examples show that exposure will inevitably induce confounding under an accelerated failure time model, whereas under the independent competing-cause model, exposure can induce confounding only if survival and censoring are not independent conditional on exposure.

COMPETING RISKS

We now extend the earlier results to explicitly consider death as a potential censoring event for the study disease. Assume each individual has four occurrence times: d_{1i} and d_{0i} as before, and c_{1i} and c_{0i} , the times of death from causes other than the study disease ('other causes') in the presence and absence of exposure. Note that, among the exposed, d_{0i} and c_{0i} are not observed, and d_{1i} is observed if and only if $c_{1i} > d_{1i}$. For the remainder of this discussion, we will assume that $d_{1i} \leq d_{0i}$ and $c_{1i} \leq c_{0i}$ for all subjects, that is, exposure is never preventive of disease or death. We will say that there is no confounding, at time of exposure, if the joint distribution of c_0 and d_0 is the same in the exposed and unexposed groups.

Excess fractions

We will consider three possible generalizations of the excess fraction in the presence of competing risks. One, the unconditional excess fraction, is defined as

$$\frac{\Pr(d_1 < c_1, d_1 \leq t) - \Pr(d_0 < c_0, d_0 \leq t)}{\Pr(d_1 < c_1, d_1 \leq t)}, \quad (13)$$

which is the excess unconditional probability of the study disease, expressed as a fraction of the total unconditional probability of the study disease. The conditional excess fraction is defined as

$$\frac{\Pr(d_1 \leq t) - \Pr(d_0 \leq t)}{\Pr(d_1 \leq t)}, \quad (14)$$

which is the excess probability of the study disease with competing risks removed, expressed as a fraction of the total probability of the study disease with competing risks removed. The third generalization, which we will call the 'semi-conditional' excess fraction, is defined as

$$\frac{\Pr(d_1 < c_0, d_1 \leq t) - \Pr(d_0 < c_0, d_0 \leq t)}{\Pr(d_1 < c_0, d_1 \leq t)}, \quad (15)$$

which is the excess probability of the study disease with exposure effects on competing risks blocked, expressed as a fraction of the total probability of the study disease with exposure effects on competing risks blocked.

Consider first the unconditional excess fraction over an interval $(0, t)$ following exposure at time zero. In the absence of confounding, at time of exposure, or loss to follow-up, one can estimate this fraction from the observed excess of study disease in the exposed cohort compared to a second, unexposed cohort. If A_+ and B_+ now represent the number of exposed and unexposed cases observed out of N_1 and N_0 initial exposed and unexposed cohort members, one may estimate the excess fraction as simply $(A_+/N_1 - B_+/N_0)/(A_+/N_1)$ (a more complicated estimate would be required to adjust for loss to follow-up). The value of this unconditional measure would have doubtful utility, however, if exposure affected risk of other causes of death. To take an extreme example, consider the small cohort of workers and firefighters at the Chernobyl accident who received so much radiation exposure that they died of acute effects in the ensuing six months. Consider the unconditional excess fraction of leukaemia over the next 50 years in this cohort compared to, say, the cohort of Soviet workers unexposed to excess radiation: This measure would be negative (indicative of a protective effect of massive radiation exposure), since none of the exposed got leukaemia, yet a few of the unexposed would get leukaemia (from background causes). This protective effect would of course have come about by exposure causing death before leukaemia could occur ($c_{1i} < d_{1i}$) rather than the usual sense of disease prevention ($d_{0i} < d_{1i}$). The same problem afflicts the unconditional incidence-proportion (risk) ratio (estimated by $(A_+/N_1)/(B_+/N_0)$), which would be zero in this example.

Because of the problems just described, some investigators suggest estimating the effect of an exposure on occurrence of the study disease by the conditional excess fraction. Unfortunately, the conditional excess fraction is not identifiable without further assumptions. One such assumption is independence of competing risks and the study disease, that is, $f_{cd}(c_1, c_0, d_1, d_0) = f_c(c_1, c_0)f_d(d_1, d_0)$, where f_{cd} is the joint density and f_c and f_d the marginal densities of the competing risks and the study disease. Given this assumption and the absence of confounding at time of exposure, the conditional excess fraction equals the identifiable parameter

$$\left[\int_0^t h_{1d}(u)S_{1d}(u)du - \int_0^t h_{0d}(u)S_{0d}(u)du \right] / \int_0^t h_{1d}(u)S_{1d}(u)du \quad (16)$$

where $h_{jd}(t)$ is the rate of the study disease in the presence ($j = 1$) or absence ($j = 0$) of exposure, and $S_{jd}(t) = \exp[-\int_0^t h_{jd}(u)du]$.

The $h_{jd}(t)$ and $S_{jd}(t)$ are estimable quantities. Under the independence assumption the $h_{jd}(t)$ are the study disease rates with competing risks removed, and $S_{jd}(t) = \Pr(d_{ji} > t)$. If, however, the disease and competing risks are dependent, the $h_{jd}(t)$ may give no indication of the rates when competing risks are removed. For example, if for all subjects i at exposure level j we have $c_{ji} < d_{ji}$, then $h_{jd}(t) = 0$ and $S_{jd}(t) = 1$ at all times.

Many people have difficulty with the concept of the effect of exposure on study disease with competing risks removed. An alternative measure of exposure effect that avoids the difficulties associated with expression (13) but does not require us to remove deaths from competing risks is given by expression (15). Given no confounding at time of exposure and independence of the study disease and competing risks, expression (15) becomes

$$\left[\int_0^t h_{1d}(u)S_{1d}(u)S_{0c}(u)du - \int_0^t h_{0d}(u)S_{0d}(u)S_{0c}(u)du \right] / \int_0^t h_{1d}(u)S_{1d}(u)S_{0c}(u)du \quad (17)$$

where $S_{0c}(t) = \exp[-\int_0^t h_{0c}(u)du]$ and $h_{0c}(t)$ is the rate of competing risks at time t .

The etiologic fraction

We next consider the instantaneous etiologic fraction in the presence of competing risks. We continue to assume that $S_1(t) < S_0(t)$ whenever $S_0(t) > 0$. In this setting, it is always possible that $d_{1i} < d_{0i}$ for all i , so that the etiologic fraction is 1 (as in our rank-preserving model). Note that if the study disease and competing risks are not independent, it is possible to have $S_{1d}(t) > S_{0d}(t)$ for all $t > 0$, even if exposure is always harmful; this would happen, for example, if $c_{1i} < d_{1i} < d_{0i} < c_{0i}$ for all subjects.¹³

In general, the greatest lower bound for the etiologic fraction among all exposed cases that occur at u is

$$\frac{h_{1d}S_1(u) - \min[h_{1d}(u)S_1(u), h_{0d}(u)S_0(u)]}{h_{1d}(u)S_1(u)} \quad (18)$$

(see Theorem 6, Reference 9). Define $f_{jd}(u) = h_{jd}(u)S_{jd}(u)$. When the competing risks are independent of the study disease, the greatest lower bound for the etiologic fraction among all exposed cases occurring at u is zero if $f_{1d}(u) \leq f_{0d}(u)$, and is $[f_{1d}(u) - f_{0d}(u)]/f_{1d}(u)$ if $f_{1d}(u) > f_{0d}(u)$.

Define $R_d(u) = h_{1d}(u)/h_{0d}(u)$. In Appendix II we show that if the competing risks are independent of the study disease and an independent competing-cause model holds for the study disease, the etiologic fraction among cases that occur at u will equal the disease-hazard fraction $[R_d(u) - 1]/R_d(u)$. We also show that these fractions will be equal if the competing risks and study disease are dependent but satisfy a *joint* independent competing-cause model, that is, if:

- each person has a second potential disease occurrence time d_{ei} , defined as the time at which the study disease would occur from a cause that involves exposure if all causes that do not involve exposure *and* all competing risks were absent;
- each person has second potential time of death from competing risks, c_{ei} , defined as the time at which such a death would occur from a cause that involves exposure if both the study disease *and* all competing risks that do *not* involve exposure were absent;
- the joint distribution of d_{ei} and c_{ei} does not depend on d_{0i} and c_{0i} , that is,

$$f(c_{ei}, d_{ei} | c_{0i}, d_{0i}) = f(c_{ei}, d_{ei}).$$

Note that under this model, $c_{1i} = \min(c_{ei}, c_{0i})$ and $d_{1i} = \min(d_{ei}, d_{0i})$, and some dependency may exist between the competing risks and the study diseases.

DISCUSSION

Case-control studies

In the above development, we have assumed that one can construct estimates of the survival distributions $S_1(u)$ and $S_0(u)$, or equivalently the hazard functions $h_1(u)$ and $h_0(u)$. It is not possible to construct such estimates from case-control data (unless one has some additional population data). Nevertheless, both excess and etiologic fractions are estimable from case-control data under certain conditions.

The incidence-proportion ratio IPR in a fixed cohort is estimable from a case-cohort or case-control study nested within the cohort.⁴ As shown earlier, the excess fraction in the cohort is equal to $(IPR - 1)/IPR$, and so this fraction is directly estimable from case-cohort or nested case-control designs.

In more general settings, one may estimate the SMR parameter $A_+(s, t)/A_0(s, t)$ from case-control studies that employ random sampling of cases and matched or unmatched 'density' sampling¹⁴ of controls over (s, t) . As shown earlier, under an independent competing cause model the etiologic fraction will equal $(SMR - 1)/SMR$, and so will be directly estimable from such studies; if in addition the hazard ratio $R(u)$ is a constant R over (s, t) , the etiologic fraction over (s, t) will reduce to $(R - 1)/R$, and R is estimable using any of the usual common rate-ratio estimators.³⁻⁵ If an independent competing-cause model does not hold but the hazard ratio is a constant R for all u , the lower bound for the lifetime etiologic fraction given in expression (11) will still be estimable from density study designs.

Plausibility of models and bounds

The hypothesis that the etiologic fraction is 1 is not reasonable if there exists more than one independent causal pathway or mechanism that can produce the outcome, and exposure does not affect one or more of the pathways: subjects who contract the outcome from a mechanism unaffected by exposure do not have their occurrence time changed by exposure. In particular, the rank-preserving hypothesis (condition (2)) and the accelerated-life model will not be reasonable under such conditions.

We think that in most situations one can, on biological grounds, rule out values approaching the lower bound in expression (9). The 'best case' assumption (in terms of minimizing estimated exposure effects) used to derive the bound is often highly implausible: it says in effect that all exposure-induced cases occur as early as possible and always before the density $f_1(u)$ finally falls below $f_0(u)$ (as it eventually must, since both densities integrate to 1). Of course, the range of plausible values left after all biological considerations is still likely to be broad. Elsewhere⁹ we show that, under certain biologically plausible assumptions, the etiologic fraction lies between that calculated under the independent competing-cause model and that computed under the rank-preserving model.

We would argue that the assumption of independence of the study disease and competing risks, often invoked to estimate conditional measures,¹⁵ is usually unrealistic: if there exists an unmeasured extraneous factor that influences the distributions of both the d_{ji} and c_{ji} , and if the d_{ji} and c_{ji} are independent within levels of the factor, the d_{ji} and c_{ji} will not be independent without controlling the factor.

As noted earlier, if we observe $R(u) > 1$ for all u , we cannot refute the hypothesis that the empirical ratio $R(u)$ equals the etiologic ratio $R_e(u)$ from the data alone. Nevertheless, we think it is often unreasonable to assume that $R(u) = R_e(u)$. Suppose that, as is almost always the case, there

exist strong unmeasured risk factors (for example, genetic factors). Then, under mechanisms leading to $R(u) = R_e(u)$, this equation would most plausibly hold within levels of such factors. If, however, $R(u) = R_e(u)$ holds within levels of the factors, it will not hold for the crude ratios except in special cases. This problem is of importance, since only the crude hazards are estimable. ('Crude' here only means calculated ignoring unmeasured factors; we still assume that there is no confounding at time of exposure, both conditional on measured factors, and unconditionally, as in a large randomized trial.)

To explore the problem further, let z be a vector of unmeasured fixed covariates, let $h_j(u)$ denote the crude hazard in the presence ($j=1$) or absence ($j=0$) of exposure, with $R(u) = h_1(u)/h_0(u)$ as before; let $h_j(u|z)$ denote the covariate-specific hazards; and let $R(u|z) = h_1(u|z)/h_0(u|z)$ and $D(u|z) = h_1(u|z) - h_0(u|z)$. We will say the joint effects of exposure and the covariates are *additive* if, at all times u , $D(u|z)$ is constant across z . If the joint effects are not additive but

$$D(u|z_1) > D(u|z_2) \quad (19)$$

whenever $h_0(u|z_1) > h_0(u|z_2)$, we will say the joint effects are *superadditive*. If the joint effects are not additive but (19) holds whenever $h_0(u|z_1) < h_0(u|z_2)$, we will say the joint effects are *subadditive*.

Assume that an independent competing-cause model holds within levels of z , so that $[R(u|z) - 1]/R(u|z)$ is the etiologic fraction within covariate level z . It then follows that the crude hazard fraction $[R(u) - 1]/R(u)$ will equal the etiologic fraction if the joint effects of exposure and the covariates are additive. More generally, if at all times u $D(u|z)$ and $h_0(u|z)$ are independently distributed in the population under study, the crude hazard fraction will equal the etiologic fraction. If, however, we additionally assume that the unexposed covariate-specific hazards never cross one another (that is, for all z_1, z_2 either $h_0(u|z_1) \geq h_0(u|z_2)$ or $h_0(u|z_1) \leq h_0(u|z_2)$ at all times) with strict inequality for some u , then the crude hazard fraction will fall below the etiologic fraction if the joint effects are superadditive, and fall above it if the joint effects are subadditive.

The preceding results are a direct corollary of results given in Reference 9. They imply, in particular, that unmeasured risk factors can bias the observed hazard fraction as an estimate of the etiologic fraction, *even if an independent competing-cause model holds within levels of the factors and the factors produce no confounding at time of exposure* (so that, for example, the incidence-proportion ratio and excess fraction can be unbiasedly estimated). In particular, many biological models are superadditive, and under such models the observed hazard fraction will be a downwardly biased estimate of the etiologic fraction. Of course, not even the covariate-specific hazard fractions need equal the etiologic fractions if the competing-cause model fails to hold within levels of the covariates.

The etiologic fraction as a measure of exposure impact

Aside from estimability problems, the etiologic fraction possesses another limitation if one interprets it as a measure of exposure impact: it is insensitive to how much exposure shortens time to occurrence. Consider the impact on life expectancy of two genetic exposures, both of which affect death time according to the accelerated-life model of example 1, but with $R=20$ for the first exposure (for example, Tay-Sachs) alone and $R=2$ for the second (for example, Huntington's disease). The etiologic fraction is essentially 1 for both exposures, yet the impacts of the exposures are dramatically different: the first guarantees death in childhood, while the second can allow

survival to middle age. In another paper,⁸ we analyse summary measures for the impact of exposure on time of occurrence that explicitly take account of how much exposure reduces life expectancy. In particular, we show that the expected years of life lost due to exposure must approach zero as R approaches 1, and that the expected years of life lost is identifiable in the absence of confounding (in contrast to the etiologic fraction).

The results of this paper have strong implications for compensation schemes with awards based solely on the 'probability of causation'. In particular, they imply that awards based on such schemes will be sensitive to assumptions about the mechanism by which exposure and other covariates produce disease. More robust compensation schemes can be formulated using expected years of life lost.^{1, 8} Robins^{16, 17} discusses related issues of causal inference that arise when the exposure and covariates are time dependent.

APPENDIX I

In this Appendix, we show that there always exists a joint distribution (d_{0i}, d_{1i}) with given marginal distributions $S_0(u)$ and $S_1(u)$ for which the etiologic fraction attains the minimum given in (9).

Theorem

Given $S_0(t)$ and $S_1(t)$ such that equation (1) holds, there exists a joint distribution F^* for (d_{0i}, d_{1i}) with $d_{1i} \leq d_{0i}$ for all subjects i such that the etiologic fraction over $(0, t)$ is given by (9).

To avoid notational complexity, we shall only show how to construct F^* in the special case in which there exists some time t_m such that $G = (0, t_m)$. That is, the densities $f_1(t)$ and $f_0(t)$ cross but once. (An algorithm for construction of F^* when $f_1(t)$ and $f_0(t)$ cross either a finite or a countable number of times is available from the authors upon request.)

Proof

Consider a joint distribution F^* for (d_{0i}, d_{1i}) such that the marginal distribution $F^*(d_{1i})$ has density $f_1(t)$ and the conditional distribution of d_{0i} give d_{1i} , that is, $Pr^*[d_0 < u | d_1 = t]$ is

$$\begin{aligned} & 0 \quad \text{if } u \leq t, \\ & 1 \quad \text{if } u > t \quad \text{and} \quad t \geq t_m, \\ & f_0(t)/f_1(t) \quad \text{if } t < u < t_m, \quad \text{and} \\ & \frac{f_0(t)}{f_1(t)} + \left[\frac{f_1(t) - f_0(t)}{f_1(t)} \right] \frac{\int_{t_m}^u [f_0(x) - f_1(x)] dx}{\int_{t_m}^{\infty} [f_0(x) - f_1(x)] dx} \quad \text{if } t < t_m < u. \end{aligned}$$

It is clear that by its construction F^* has etiologic fraction over the interval $(0, t)$ as given by (9). Therefore, it only remains to prove the following.

Lemma

$Pr^*(d_0 < t)$ equals $Pr(d_0 < t)$ where Pr without an $*$ is evaluated under $S_0(t)$ (that is, $Pr(d_0 < t) = 1 - S_0(t)$) and Pr^* is evaluated under F^* .

Proof

Suppose $u > t_m$. Then

$$\begin{aligned}
 Pr^*[d_0 < u] &= \int_0^\infty Pr^*[d_0 < u | d_1 = t] f_1(t) dt \\
 &= \int_0^{t_m} \frac{f_0(t)}{f_1(t)} f_1(t) dt + \left[\int_0^{t_m} \left[\frac{f_1(t) - f_0(t)}{f_1(t)} \right] f_1(t) dt \right] \\
 &\quad \times \frac{\int_{t_m}^u [f_0(t) - f_1(t)] dt}{\int_{t_m}^\infty [f_0(t) - f_1(t)] dt} + \int_{t_m}^u f_1(t) dt \\
 &= Pr(d_0 < t_m) + [Pr(d_1 < t_m) - Pr(d_0 < t_m)] \\
 &\quad \times \frac{Pr(t_m < d_0 < u) - Pr(t_m < d_1 < u)}{Pr(t_m < d_0) - Pr(t_m < d_1)} + Pr(t_m < d_1 < u) \\
 &= Pr(d_0 < t_m) + [Pr(t_m < d_0 < u) - Pr(t_m < d_1 < u)] + Pr(t_m < d_1 < u) \\
 &= Pr(d_0 < u)
 \end{aligned}$$

The proof for $u \leq t_m$ is similar.

APPENDIX II

In this Appendix, we prove the following.

Theorem

For a subject observed to die of cause d at time t the instantaneous etiologic fraction is $[R_d(t) - 1]/R_d(t)$ where $R_d(t) = h_{1d}(t)/h_{0d}(t)$, if either (i) competing risks are independent of the study disease and the independent competing cause model holds for the study disease d or (ii) the independent competing cause model holds jointly for disease d and competing risks c .

Proof

By definition the instantaneous etiologic fraction for a subject observed to die of cause d at time t is

$$Pr[d_0 > t | d_1 = t < c_1] = [R_{de}(t) - 1]/R_{de}(t) \text{ where } R_{de}(t) = \frac{h_{1d}[t | i \in N_1(t)]}{h_{0d}[t | i \in N_1(t)]} = \frac{h_{1d}(t)}{h_{0d}[t | i \in N_1(t)]},$$

$N_1(t)$ is the set $\{i : d_{1i} > t, c_{1i} > t\}$,

and

$$h_{0d}[t | i \in N_1(t)] = \lim_{\Delta t \rightarrow 0} \frac{Pr[t < d_0 < t + \Delta t | d_1 > t, c_1 > t]}{\Delta t}.$$

Thus it is sufficient to show that $h_{0d}[t|eN_1(t)] = h_{0d}(t)$. Now under the assumptions of independent competing risks and independent competing causes for disease d , as $\Delta t \rightarrow 0$,

$$\begin{aligned} Pr[t < d_0 < t + \Delta t | d_1 > t, c_1 > t] / \Delta t &= Pr[t < d_0 < t + \Delta t | d_1 > t] / \Delta t \\ &= Pr[t < d_0 < t + \Delta t | d_0 > t] / \Delta t \\ &= Pr[t < d_0 < t + \Delta t | d_0 > t, c_0 > t] / \Delta t \\ &= h_{0d}(t) \end{aligned}$$

where the first and third equality hold by independence of competing risks, the second by independence of competing causes for cause d , and the last equality is definitional.

Next suppose that the independent competing cause model holds jointly for causes c and d , that is

$$Pr(c_{ei}, d_{ei} | c_{0i}, d_{0i}) = Pr(c_{ei}, d_{ei}) \text{ where, for example, } c_{1i} = \min(c_{0i}, c_{ei}).$$

Then, as $\Delta t \rightarrow 0$,

$$\begin{aligned} Pr[t < d_0 < t + \Delta t | d_1 > t, c_1 > t] / \Delta t &= Pr[t < d_0 < t + \Delta t | d_0 > t, c_0 > t, d_e > t, c_e > t] / \Delta t \\ &= Pr[t < d_0 < t + \Delta t | d_0 > t, c_0 > t] / \Delta t = h_{0d}(t), \end{aligned}$$

where the first and third equalities are definitional and the second follows from the joint independent competing cause model.

REFERENCES

1. Greenland, S. and Robins, J. M. 'Conceptual problems in the definition and interpretation of attributable fractions', *American Journal of Epidemiology*, **128**, 1185-1197 (1988).
2. Rothman, K. J. *Modern Epidemiology*, Little-Brown, Boston, 1986.
3. Kleinbaum, D. G., Kupper, L. L. and Morgenstern, H. *Epidemiologic Research: Principles and Quantitative Methods*, Lifetime Learning Publications, Belmont, 1982.
4. Miettinen, O. S. *Theoretical Epidemiology*, Wiley, New York, 1985.
5. Breslow, N. E. and Day, N. E. *Statistical Methods in Cancer Research: I. The Analysis of Case-Control Studies*, IARC, Lyon, 1980.
6. Hatch, O. G. 'Medical/legal aspects of radiation-induced cancer', *Health Physics Society Newsletter*, December (1984).
7. Lagakos, S. W. and Mosteller, F. 'Assigned shares in compensation for radiation-related cancers', *Risk Analysis*, **6**, 345-347 (1986).
8. Robins, J. M. and Greenland, S. 'Estimability and estimation of years of life lost due to a hazardous exposure', Technical Report No. 8, Occupational Health Program, Harvard School of Public Health, 1988.
9. Robins, J. M. and Greenland, S. 'The probability of causation under a stochastic model for individual risk', SIMS Technical Report, No. 36, Harvard School of Public Health, 1988.
10. Greenland, S. and Robins, J. M. 'Identifiability, exchangeability, and epidemiologic confounding', *International Journal of Epidemiology*, **15**, 412-418 (1986).
11. Cox, D. R. and Oakes, D. *Analysis of Survival Data*, Chapman and Hall, New York, 1984.
12. Keiding, N. and Vaeth, M. 'Calculating expected mortality', *Statistics in Medicine*, **5**, 327-334 (1986).
13. Slud, E. and Byar, D. 'How dependent causes of death can make risk factors appear protective', *Biometrics*, **44**, 265-269 (1988).
14. Greenland, S. and Thomas, D. C. 'On the need for the rare disease assumption in case-control studies', *American Journal of Epidemiology*, **116**, 402-406 (1982).
15. Flanders, W. D. and Louv, W. C. 'The exposure odds ratio in nested case-control studies with competing risks', *American Journal of Epidemiology*, **124**, 684-692 (1986).
16. Robins, J. M. 'A new approach to causal inference in mortality studies with a sustained exposure period—application to the control of the healthy-worker survivor effect', *Mathematical Modelling*, **7**, 1393-1512 (1986).
17. Robins, J. M. 'The probability of causation and years of life lost when exposures and covariates are time-varying', SIMS Technical Report, No. 37, Harvard School of Public Health, 1988.