

**Premature Deaths, Statistical Lives, and Years of Life Lost:
Identification, Quantification, and Valuation of Mortality Risks**

James K. Hammitt
Harvard University (Center for Risk Analysis & Center for Health Decision Science)
Toulouse School of Economics, Université Toulouse Capitole
jkh@harvard.edu

Peter Morfeld
Ruhr-Universität Bochum

Jouni Tuomisto
National Institute for Health and Welfare, Kuopio

Thomas C. Erren
University of Cologne

29 August 2019

For presentation at the Risk Assessment, Economic Evaluation, and Decisions Workshop, Harvard Center for Risk Analysis, 26-27 September 2019.

Abstract

Mortality effects of exposure to air pollution and other environmental hazards are often described by the estimated number of 'premature' or 'attributable' deaths and the economic value of a reduction in exposure is quantified as the product of an estimate of 'statistical lives saved' and a 'value per statistical life'. These terms can be misleading because the number of deaths advanced by exposure cannot be determined from epidemiological information alone (it is not statistically identified). The fraction of deaths 'attributed' to exposure is conventionally derived as equal to the hazard fraction $(R - 1) / R$, where R is the relative risk of mortality between high and low exposure levels. The fraction of deaths advanced by exposure (the 'etiologic' fraction) can be substantially larger or smaller: it can be as large as one and as small as $1/e$ (≈ 0.37) times the hazard fraction (if the association is causal and zero otherwise). Recent literature reveals conceptual misunderstanding and confusion about premature, etiologic, excess, and attributable deaths. Total life years lost in a population due to exposure can be estimated but cannot be disaggregated by age or cause of death. Economic valuation of a change in exposure-related mortality risk is independent of the fraction of deaths that are etiologic. When individuals facing larger or smaller changes in mortality risk cannot be identified, the mean change in population hazard is sufficient for valuation; otherwise, the population value of a change in exposure can depend on the distribution of risk reductions in the population.

Keywords: premature death, attributable death, hazard fraction, value per statistical life (VSL), years of live lost (YLL), disability-adjusted life year (DALY), environmental burden of disease (EBD)

1. Introduction

Health effects of environmental exposures are quantified for use in impact studies and environmental burden of disease studies. Impact studies evaluate the effects on health of government regulations and other policies to reduce exposures, such as requiring emission-control equipment or imposing fees on emissions. The estimated health effects are often used as an input to a benefit-cost analysis that quantifies the positive and negative effects of the policy in monetary terms to help determine whether it is on balance desirable. Environmental burden of disease studies evaluate the harms to health from exposure compared with some counterfactual situation, to provide perspective on the importance of different exposures and the possible benefits of reducing them.

Both impact and environmental burden of disease studies present results described using terms such as 'premature deaths' and 'statistical lives saved'. The economic value of reducing mortality risk is calculated using the concept of 'value per statistical life'. The literature suggests there is much confusion about how these and similar terms are defined and what can be known about the mortality effects of environmental exposures.

In particular, the number or fraction of deaths among individuals exposed to some hazard that occur earlier than they would have if the individuals had not been exposed ('etiologic deaths') cannot be determined from mortality data alone, whether from epidemiological studies or randomized controlled trials. The problem is that etiologic deaths are not statistically identified: the time at which an exposed individual dies can be observed, but the time at which she would have died had she not been exposed is counterfactual and cannot be observed. From mortality data alone, one cannot distinguish between situations in which a few people die much earlier than they would have if unexposed and situations in which many people die a little earlier than they would have if unexposed. The fraction of deaths in an exposed population that are etiologic can be bounded but these bounds are much farther apart than the ends of typical confidence or uncertainty intervals presented in the literature. The total number of life years lost due to exposure is statistically identified but cannot be disaggregated into life years lost conditional on death at specified ages or from specified diseases, as these quantities are not statistically identified. Nevertheless, uncertainty about the fraction of deaths that are etiologic and the number of life years lost conditional on death at a specified age does not affect the economic value of reducing exposure as conventionally defined, as the economic value depends on the change in mortality risk, which is identified.

In Section 1.1, we describe some of the problems with how mortality risks are described in the literature that motivate this paper. The terms we discuss are defined in Table 1.

In Section 1.2, we present a simple example that illustrates the key points of our paper: the fraction of deaths in an exposed population that are etiologic cannot be determined from mortality data alone nor can the life years lost to fatality at specified ages or from specified causes. In contrast, total life years lost in the exposed population and the economic value of reducing exposure-related mortality risk can be determined. The economic value of reducing exposure-related mortality risk is unaffected by uncertainty about the fraction of deaths that are etiologic, though it could be affected by knowledge of how the fatality risk differs among identifiable individuals. Readers who are not interested in the technical details may wish to read only the introduction (Section 1) and conclusions (Section 4).

In the following sections, we describe in principle what aspects of the mortality effects of environmental exposures can be estimated using mortality data (in Section 2) and how the effects on mortality are valued using economics (in Section 3). We focus on concepts and statistical identifiability rather than questions of sample size, data quality, and empirical methods that affect the precision with which quantities can be estimated. These sections present the basic concepts and equations underlying quantification and valuation of exposure-related mortality effects and can serve as an introduction to these topics. Conclusions and implications are in Section 4.

For concreteness, we focus our discussion on the mortality effects of (fine particulate) air pollution. But our analysis is more general and is applicable to other cases in which one cannot determine whether a specific death was the result of exposure to a particular hazard. As noted, our discussion is not limited to epidemiological studies; the same issues arise in randomized controlled trials. The critical issue is that the age at which an individual dies is observable but the (counterfactual) age at which she would have died had her exposure been different is not.

1.1. Terminology commonly used in the literature

Environmental burden of disease (EBD) and impact studies¹ that quantify the harms to human health from air pollution or other exposures, or the health benefits of reducing exposure, often use wording

¹ See Ostro (2004) for a definition and example of an EBD study and COMEAP (2010) for a description of EBD studies and a distinction between burden and impact analyses. EBD studies focus on populations during a period, usually a specified calendar year.

such as ‘premature deaths’ (Anenberg et al. 2017, Landrigan et al. 2017), ‘excess deaths’ (Burnett et al. 2018), ‘attributable deaths’ (Cohen et al. 2017), ‘deaths attributable to AAP [ambient air pollution]’ (WHO 2016), ‘PM_{2.5}-related premature mortality’ (Wang et al. 2017), and ‘premature mortality attributable to air pollution’ (Lelieveld et al. 2015). These terms are often used interchangeably and in a causal sense: e.g., Cohen et al. (2017) described numbers of deaths ‘attributed’ to exposure as numbers of deaths ‘caused’ by exposure; Anenberg et al. (2019) used ‘premature deaths’ in the same sense as ‘attributable deaths’ and described emissions as ‘responsible’ for pollutant-attributable deaths; Burnett et al. (2018) equated ‘excess’ and ‘avoided’ deaths. Goodkind et al. (2019) conclude that “anthropogenic PM_{2.5} was responsible for 107,000 premature deaths in 2011, at a cost to society of \$886 billion.”

The estimates of ‘premature’, ‘excess’ or ‘attributable’ deaths in these papers are computed as products of mortality counts and hazard fractions² (e.g., EPA 1996, Lim et al. 2012, GBD 2017 Risk Factors Collaborators 2018). Some papers (e.g., WHO 2016, Cohen et al. 2017, Landrigan et al. 2017) also present results in terms of disability-adjusted life years (DALYs), calculated as the sum of years of life lost to death and years lived with disability (Murray 1994, Prüss-Üstün et al. 2003, GBD 2017 Risk Factors Collaborators 2018, WHO 2016).

These terms can be misleading. It is not possible based on current science to know the number of etiologic deaths, i.e., the number of individuals who die earlier than they would have because of air pollution or other exposures lacking a distinctive relationship between exposure and disease.³ As is well known (Rothman et al. 2008), epidemiology can identify associations between exposure and mortality, but determining whether an exposure causes mortality requires other forms of reasoning and evidence, for example to rule out the possibility that both exposure and mortality are caused by some other (confounding) factor. Well-conducted randomized controlled trials with large samples have the advantage that any association between exposure and mortality is unlikely to be caused by a third factor, since the probabilities of associations between exposure and possible confounding factors should be eliminated by randomization. The question of causality is not the focus of this paper; from this point

² Hazard fractions are defined in Section 2. They are also called population attributable fractions if based on the exposure pattern of a population (Rothman et al. 2008).

³ The relationship between mesothelioma and asbestos exposure is often cited as an example of a distinctive relationship in which the disease can be linked to an exposure, but even in this case, there can be uncertainty about the cause. Mesothelioma can be caused by exposure to erionite (IARC 1987, Dikensoy 2008, Jasani and Gibbs 2012) and may arise even without significant exposure to asbestos or erionite in susceptible individuals (Testa et al. 2011).

onward, we assume that any observed relationship between exposure and mortality is causal. The limitations we describe apply even in an ideal situation of true causal effects and epidemiological studies or randomized controlled trials free of bias and random error. If the association between exposure and mortality is not causal, then no deaths are caused by exposure.

The standard calculation used to quantify the mortality effects of exposure estimates the number of ‘deaths attributable to exposure’ (or ‘attributable deaths’), which is the difference between the number of deaths expected to occur in an exposed population (during a defined period) and the number that would be expected to occur in that population if mortality risk in the period were the same as in an unexposed population. Attributable⁴ deaths are calculated as the product of total deaths in the exposed population and the hazard fraction. The number of etiologic deaths can be smaller or much larger than the number of attributable deaths. Attributable deaths can also differ from ‘excess deaths’, which is the difference between the steady-state numbers of deaths that occur in a period if the population has been continuously exposed and the number that would occur if the population was never exposed.

In contrast to etiologic deaths, the total number of life years lost in the exposed population is statistically identified and can be estimated using mortality data. However, for the same reason that the number of etiologic deaths is not identifiable, the life years lost among individuals who die at specified ages or from specified causes (which are components of DALYs) are not identified and cannot be known using only mortality information. Similarly, estimates of the monetary value of a change in mortality risk based on the change in life expectancy (e.g., using a constant ‘value per statistical life year’) can be calculated at the population level but cannot be disaggregated to categories such as deaths at particular ages or from specific diseases.

1.2. A tale of two cities

Suppose there are two cities that are identical, except in air pollution and mortality. In one (Clearville), the air is clean and one-quarter of the population dies at ages 70, 80, 90, and 100 years; in the other (Smokeville), the air is polluted and one-quarter of the population dies at ages 60, 70, 80, and 90.

Assume that individuals live their entire lives in their city of birth, that the difference in mortality is

⁴ While excess and etiologic deaths are well-defined (Greenland 2015, Suzuki et al. 2012), the term ‘attributable’ is used inconsistently. A leading epidemiology textbook (Rothman et al. 2008) reports that attributable deaths sometimes refers to excess deaths and sometimes to etiologic deaths; Poole (2015) describes the history and multiple applications of the term.

caused by air pollution, and that if pollution in Smokeville were eliminated the mortality rates would instantaneously become identical to those in Clearville. Consider two questions:

1. What fraction of the deaths in Smokeville are etiologic, i.e., occur earlier than they would have if there were no pollution? Equivalently, what fraction of the people in Smokeville would live longer if pollution were eliminated?
2. What is the economic value of eliminating the mortality risk due to air pollution in Smokeville? Does it depend on the answer to question 1?

The fraction of deaths that are etiologic is not statistically identified using mortality data. It could be one, if all deaths are advanced by 10 years. It could be one-quarter, if people who would die at 100 absent air pollution die at 60 in its presence and no one else is affected. It can be any number between these extremes; for example, if some people are of type A (who die 10 years earlier if exposed to air pollution) and the remainder are of type B (who are unaffected, except the quarter who would live to 100 in the absence of air pollution die at 60 in its presence).

Figure 1 illustrates the probability distributions by age of death in Smokeville and Clearville for three possible fractions of deaths that are etiologic. Panel A shows the distribution in Smokeville, where the colors correspond to people who would die at each age. In Panel B, each individual's age of death in Clearville is 10 years older than in Smokeville (i.e., all of the individuals in each age-of-death bin are shifted to the next higher bin). In Panel C, individuals who would die at 60 in Smokeville die at 100 in Clearville, and all other individuals die at the same age in Smokeville and in Clearville. Panel D presents the case in which the fraction of deaths at each age that are etiologic equals the fraction that are attributable to air pollution, as described in Section 2. In this case, all the people who die at 60 would have lived longer in Clearville: some would have lived to 70, some to 80, some to 90, and some to 100. The fractions of people who die at ages 70, 80, and 90 who would have lived longer in Clearville are one-quarter, one-third, and one-half, respectively.

Although the probability distribution of age at death is the same in panels B, C, and D, the fractions of people whose age at death in Smokeville is less than their age at death in Clearville equals 1, 0.25, and 0.52 in Panels B, C, and D, respectively. Knowing only the probability distributions of age at death does not allow one to distinguish between these three situations (or many others that could be constructed).

Some important concepts are statistically identified and can be estimated using mortality data for the two cities. The excess deaths by age (the difference between the expected number of deaths by age in

Smokeville and Clearville) is identified; as a fraction of each birth cohort it is one-quarter at 60, minus one-quarter at 100, and zero at all other ages. The number of life years lost due to exposure is also identified: it is 10 per capita. However, the number of life years lost due to exposure and conditional on death at a particular age is not identified: for deaths at 60, it is between 10 and 40; for deaths at 70, 80, and 90 it is between 0 and 30, 20, and 10, respectively.⁵ Similarly, the number of life years lost due to exposure and conditional on cause of death is not identified. Assume that some of the deaths at age 60 are caused by heart attack and that all the deaths at older ages are from other causes. Then the life years lost due to exposure and conditional on death from heart attack is between 10 and 40.⁶

The economic value of reducing mortality risk in Smokeville by eliminating air pollution is equal to the sum over the population of each individual's value of her risk reduction. For an individual of age $a < 60$, the value is equal to the expected present value of the risk reductions at future ages,

$$V = \sum_{t=60}^{90} v(t) r(t) S(t) \rho^{t-a} \quad (1.1)$$

where V is her total monetary value, $v(t)$ is her average value per unit risk reduction of reducing her hazard at age t by $r(t)$, $S(t)$ is her probability of surviving to t , and ρ is her discount factor (equal to $1/(1 + \text{her discount rate})$).⁷ If there is no additional information about the age at which each individual will die, the value V does not depend on the etiologic fraction: whatever its value, each individual substitutes a lottery with equal chances of dying at ages 70, 80, 90, or 100 for a lottery with equal chances of dying at ages 60, 70, 80, or 90.

⁵ It is important to distinguish between the number of life years lost conditional on death at age t , and the number of life years lost conditional on death at age t that are due to exposure. The former equals life expectancy at age t in the exposed population (i.e., 0, 10, 15, and 20 years at ages 90, 80, 70, and 60, respectively). The latter is not statistically identified because it is not known how many deaths at each age are due to exposure. The bounds on life years lost due to exposure reflect the possibilities that air pollution has no effect on people who would have lived to 70, 80, or 90 in Clearville, and that deaths at these ages in Smokeville are only among individuals who would have lived to 100 in Clearville.

⁶ Again, it is important to distinguish between the number of life years lost conditional on death from a specified disease and the number lost conditional on death from a specified disease and due to exposure. In the example, the expected number of life years lost due to heart attack equals the life expectancy in Smokeville conditional on death at 60 (20 years).

⁷ The model leading to this equation is explained in Section 3.

To confirm that the value of eliminating exposure-related mortality risk is independent of the etiologic fraction, assume that an individual's total value depends on whether she is of type A or B; the values V^A and V^B are the results of equation (1.1) conditional on being of each type.

The elements of equation (1.1) are presented in Table 2. Conditional on being of type A, eliminating pollution eliminates the mortality hazard at ages 60, 70, 80, and 90. The individual will benefit from exactly one of these (though which one is unknown) and the total value is the sum of the four terms, each weighted by its probability. Conditional on being of type B, eliminating pollution eliminates the mortality hazard at age 60 and the total value is the value of this change multiplied by the probability the individual would die at 60. Note that the value of reducing risk at each age depends on type. For type A, $v^A(t)$ is the value of delaying death from t to $t + 10$; for type B, $v^B(t)$ is the value of delaying death from t to $t + 40$. But $v^B(60)\rho^{60-a} = \sum_{t=60}^{90} v^A(t)\rho^{t-a}$ because both sides of this equation are the value at age a of delaying death from 60 to 100.⁸ Hence the value of eliminating exposure-related mortality risk is the same for types A and B; knowledge of the individual's type has no effect on her value.

Knowledge about an individual's type could affect valuation if the age at which she would die in Smokeville is known. If she would die at 70, 80, or 90 in Smokeville, the value of eliminating exposure is zero if she is of type B and is positive if of type A. If she would die at 60 in Smokeville, the value of eliminating pollution is greater if she is of type B rather than type A.⁹

The following sections generalize the results of this example, describe the calculation of attributable deaths, and provide additional details. To summarize, the difference in the probability of death by age between populations that differ in exposure is identified and can be estimated, as can the total life years lost due to exposure. However, the fraction of deaths that occur earlier and the life years lost conditional on age or cause of death in the exposed population are not identified; their values differ between alternative models of the relationship between exposure and mortality that cannot be distinguished using only mortality data. These problems affect estimates of the burden of disease but not of the economic value of a change in exposure, as valuation depends only on changes in individuals'

⁸ Note that the summation includes only four terms, $t = 60, 70, 80, 90$.

⁹ Information about type would also affect valuation in the less extreme situation where the individual has information showing that her probabilities of dying at ages 60, 70, 80, or 90 are not equal.

risks. Knowledge of the etiologic fraction does not change the economic value, unless there is also information about how an individual's risk of death at each age differs from the population average risk.

2. Epidemiology: estimating mortality effects

Epidemiological studies collect data that can be used to estimate population-level mortality-hazard functions and survival functions for different exposures. For example, studies of the effects of fine-particulate air pollution on mortality compare mortality rates between populations living in different locations (long-term cohort studies, e.g., Dockery et al. 1993, Hoek et al. 2013) or within the same population at different times (time-series studies, e.g., Dominici et al. 2005, Atkinson et al. 2014, and intervention studies, e.g., Pope et al. 1992, 2007).¹⁰ Hazard and survival functions may be estimated for subpopulations defined by individuals' age, sex, ethnicity, smoking status, and other observable factors within the limits of epidemiology (e.g., sample size, difficulties in measuring individual characteristics, inability to control for unobserved individual characteristics, uncertainty about appropriate functional form such as constant proportional hazard).

For simplicity, we assume stable populations. A population is stable if the birth rate (number of births per year) is constant, all birth cohorts face the same hazard function, and steady state has been reached. Given these assumptions the age distribution of a population in every calendar year coincides with the survival function of the birth cohorts. We will make use of this correspondence in the following, often implicitly. Thus, our results apply to EBD population studies (that assume stable populations) as well as to epidemiological cohort studies.

Assume there are two exposure levels of interest, 'high' and 'low' (we will sometimes describe individuals exposed to these levels over their lifetimes¹¹ as 'exposed' and 'unexposed', respectively). We assume there are no other differences between the populations that affect mortality. Consider the hazard functions for total mortality (i.e., for all causes of death) corresponding to these two exposure

¹⁰ Long-term cohort studies yield estimates of the effects of continuous exposure to alternative (time-varying) pollution levels; time-series and intervention studies yield estimates of the effect of changing exposure.

¹¹ This assumption of a long-term exposure difference is typical for EBD studies: e.g., GBD 2017 Risk Factors Collaborators (2018, p. 1926) estimates attributable burden by multiplying total cases by the population attributable fraction (i.e., hazard fraction), which "is the proportion by which the outcome would be reduced in a given population and in a given year if the exposure to a risk factor in the past were reduced to the counterfactual level."

levels for a birth cohort (or a subcohort defined by observable characteristics). The hazards are functions of individual age t ($t = 0$ at birth). Let $h_0(t)$ and $h_1(t)$ be the hazard functions corresponding to low and high exposure over a lifetime, respectively. Then, for $j = 0, 1$, the corresponding survival curves are

$$S_j(t) = \exp\left(-\int_0^t h_j(\tau) d\tau\right) \quad (2.1)$$

and the marginal probability distributions for age at death are

$$f_j(t) = h_j(t)S_j(t) = -\frac{d}{dt}S_j(t). \quad (2.2)$$

For stable populations, the age distribution of the population is also described by equations (2.1) and (2.2). To simplify notation, let population be measured as a proportion of the birth cohort, so annual births and (at steady state) annual deaths equal one.

If exposure increases mortality risk at all ages (i.e., $h_1(t) > h_0(t)$ for all $t > 0$), then the number of people of any age $t > 0$ is larger in the unexposed population ($S_0(t) > S_1(t)$). Figure 2 provides an illustration of these functions where $h_1(t)$ is the hazard function for U.S. males (the exposed population) and $h_0(t) = (2/3) h_1(t)$ is the hazard function for the unexposed population (the factor 2/3 is chosen for illustration and does not represent typical hazard ratios found in environmental studies). In the figure, $h_1(t) > h_0(t)$ and hence $S_1(t) < S_0(t)$ for all t greater than zero.

The standard calculation of attributable deaths at age t , $D(t)$, is

$$D(t) = \frac{R(t)-1}{R(t)} f_1(t) \quad (2.3)$$

where $f_1(t)$ is the total number of deaths at age t in the exposed population, $[R(t) - 1] / R(t)$ is the hazard fraction, and $R(t) = h_1(t) / h_0(t)$ is the relative risk or rate ratio, which is typically treated as constant across t (the proportional hazard assumption).¹²

¹² Alternatively, the deaths attributed to each air-pollution-related disease can be calculated using disease-specific values of both relative risk R and total deaths, then summed over diseases to calculate total attributable deaths. Note that although the sum over diseases of hazard fractions can exceed one (Rowe et al. 2004), the sum of deaths attributable to each air-pollution-related disease cannot exceed

Substituting for $R(t)$ and $f_1(t)$, equation (2.3) can be expressed as

$$D(t) = [h_1(t) - h_0(t)] S_1(t). \quad (2.4)$$

At age t , attributable deaths $D(t)$ is the difference between observed deaths in the exposed population $h_1(t) S_1(t)$ and expected deaths in that population in the absence of exposure $h_0(t) S_1(t)$. It describes the result of a hypothetical and extreme situation: the instantaneous reduction in the expected number of deaths occurring in a population with age distribution $S_1(t)$ when exposure is suddenly eliminated, if the effects of exposure on mortality depend on current but not on past exposure (no latency).¹³

An alternative measure of the mortality effect of exposure is the ‘excess’ deaths at t , which is the expected steady-state difference in the number of deaths by age between exposed and unexposed populations,

$$\begin{aligned} \Delta(t) &= f_1(t) - f_0(t) = h_1(t) S_1(t) - h_0(t) S_0(t) \\ &= [h_1(t) - h_0(t)] S_1(t) + h_0(t) [S_1(t) - S_0(t)] \\ &= D(t) - h_0(t) [S_0(t) - S_1(t)]. \end{aligned} \quad (2.5)$$

Excess deaths $\Delta(t)$ equals attributable deaths $D(t)$ minus a term that accounts for the difference in the population at risk, which depends on previous exposure. The standard calculation of attributable deaths $D(t)$ from (2.4) overestimates the number of excess deaths $\Delta(t)$ that could be delayed by a permanent reduction in exposure at every age t for which $S_0(t) > S_1(t)$ and $h_0(t) > 0$. Unless exposure decreases the hazard at some ages, equation (2.4) overestimates the steady-state reduction in deaths at all ages. The second term in equation (2.5) is the additional number of deaths at age t that would have occurred in the absence of exposure, equal to the increase in the population surviving to age t multiplied by the

total deaths. This follows because the sum of disease-specific deaths $\sum_{k=1}^K f_{1k}$ (where k indexes disease)

cannot exceed total deaths f_1 and hence the sum of the attributable disease-specific deaths

$$\sum_{k=1}^K \frac{R_k - 1}{R_k} f_{1k} < \sum_{k=1}^K f_{1k} \leq f_1.$$

¹³ These assumptions are unrealistic for air pollution (e.g., COMEAP 2010) and conflict with assumptions about long-term exposure contrasts regularly made in EBD studies (e.g., GBD 2017 Risk Factors Collaborators 2018).

hazard if unexposed. Hence the number of deaths attributable to exposure does not correspond to the number of excess deaths due to exposure except in special situations (e.g., for a rare disease, $h_0(t) \approx 0$, or one that has negligible effect on mortality, $S_1(t) \approx S_0(t)$). Moreover, attributable deaths in different periods cannot be meaningfully summed, since these depend on the exposed population living in each period.

The total number of deaths that occur in a year is the integral of $f_j(t)$ over all ages and the difference in total annual deaths between exposed and unexposed populations is the integral of $\Delta(t)$ over all ages. In steady state, total deaths equal total births and the difference in total annual deaths is zero.¹⁴ However, the age distribution of deaths is different: for any age t , the total number of deaths occurring at ages younger [older] than t is smaller [larger] in the unexposed than in the exposed population (on the assumption that $S_0(t) - S_1(t) > 0$ for all $t > 0$).

For the case illustrated in Figure 2, as shown in the bottom panel, fewer [more] deaths occur in the unexposed than in the exposed population at every age younger [older] than 86 years. In contrast, attributable deaths $D(t)$ is strictly greater than zero for all t , because the instantaneous hazard is larger for the exposed than the unexposed population ($h_1(t) > h_0(t)$) and some members of the exposed population are living at all ages shown in the figure ($S_1(t) > 0$).

The hazard functions $h_i(t)$ and hence the survival functions $S_i(t)$ and hazard fraction $[R(t) - 1] / R(t)$ can be estimated under “essentially the same conditions often cited for general study validity,” i.e., absence of confounding, misclassification, and differential censoring (Robins and Greenland 1989a). From these, the expected number of deaths that would occur at each age given continuous high or low exposure can be estimated. However, these estimates are not sufficient to determine how many individuals would die earlier if exposed rather than unexposed, how much earlier each would die, or how much longer each exposed individual would have lived if unexposed. While the marginal distributions of deaths by age in the exposed and unexposed birth cohorts are observable, mortality data provide no information about the joint distribution of the ages at which an individual would die conditional on high or low exposure (one is observable, the other is counterfactual).¹⁵ As illustrated in Section 1.2 and Figure 1, many

¹⁴ That is, the integral of $\Delta(t)$ over all ages equals 0 and so if there are fewer deaths at some ages ($\Delta(t) < 0$ for some values of t) there must be more deaths at other ages ($\Delta(t) > 0$ for some other values of t).

¹⁵ Both epidemiological studies and randomized controlled trials compare the mortality experience of different subpopulations under the assumption that the members of each can be modeled as being

patterns of association between the age of death conditional on exposure to high and to low air pollution are consistent with the two marginal distributions. Recognition of this point is not new; Tsiatis (1975) proved that any set of marginal probability distributions for age at death conditional on alternative exposure levels is consistent with an infinity of joint distributions across exposure levels, and hence estimates of concepts that depend on the joint distribution require additional information. Rabl (2003) provided an excellent graphical illustration of why the number of deaths that are advanced and the extent to which each is advanced cannot be identified from epidemiological data.

The inability to identify the joint distribution of the ages at which an individual would die if she were exposed or unexposed can be illustrated using Figure 2. Consider the probability distributions for age at death in the bottom panel. To transform the probability distribution if exposed $f_1(t)$ to the distribution if unexposed $f_0(t)$, one can imagine shifting a small amount of probability of death at a specific age t' to some older age t'' . This corresponds to shifting the age at death for some individual from t' to t'' . The effect will be to decrease the value of $f_1(t')$ and increase the value of $f_1(t'')$. One can continue in this way to shift probability density to the right (increasing individuals' ages of death) until the modified version of $f_1(t)$ falls to the level of $f_0(t)$ for all ages where $f_1(t) > f_0(t)$ (i.e., all ages less than 86 years) and rises to the level of $f_0(t)$ for all ages where $f_1(t) < f_0(t)$ (i.e., all ages older than 86 years). The number of different possibilities for shifting probability from one age to another to achieve this transformation is infinite; Figure 1 illustrates three possibilities for the example described in Section 1.2. Hence the number of joint distributions associating age at death if exposed and if unexposed that are consistent with the marginal probability distributions is infinite (Tsiatis 1975).¹⁶

randomly drawn from a common source population (perhaps controlling for observable characteristics). (For randomized controlled trials, this assumption is valid by construction.) Hypothetically, one could know the age of death for each person conditional on being selected into each subpopulation, in which case one would know the population distribution of years of life lost. But it is not possible to associate a member of one subpopulation with a single counterpart in the other: her counterpart is equally likely to be any member of the other subpopulation. Hence her counterfactual time of death if she had been selected into the other subpopulation is equally likely to be the time of death of each member of that subpopulation.

¹⁶ For the example illustrated in Figure 1, this procedure amounts to distributing all of the probability for death at age 60 in Smokeville to deaths at ages 70, 80, 90, and 100 in Clearville, and either shifting to older ages or leaving in place amounts of probability corresponding to deaths at ages 70, 80, and 90, subject to the constraint that the sum of these changes yields the probability distribution for age of death in Clearville. Assuming that exposure never decreases mortality hazard, no probability increments

Equivalently, one can think of representing the relationship between individuals' ages of death if exposed and unexposed by shifting the survival curve in the middle panel from $S_1(t)$ to $S_0(t)$. Any shift of probability density from a younger to an older age at death (e.g., from t' to t'') in the density function $f_1(t)$ has two effects on the survival curve $S_1(t)$. First, it shifts the point $(t', S_1(t'))$ to the right (to t'') and downward (to $S_1(t'')$). The downward shift is because the affected individual moves from dying before to dying after all those who die between ages t' and t'' . Second, all the points on $S_1(t)$ for values of t between t' and t'' shift upward (because these individuals are moved from dying after to dying before the affected individual). One can continue shifting the survival curve $S_1(t)$ in this way until it corresponds to the survival curve if unexposed $S_0(t)$. Again, there are an infinite set of these small probability shifts that transform the survival function if exposed to the survival function if unexposed.

2.1. Bounds on the etiologic fraction

In this section we investigate what can be inferred about etiologic deaths from mortality data. We change our focus from the exposed and unexposed populations to the birth cohorts that generate the populations. As noted above, the exposed and unexposed birth cohorts have the same survival functions as the exposed and unexposed populations.

Robins and Greenland (1989a) derive bounds on the 'etiologic fraction', defined as the fraction of deaths occurring in the exposed birth cohorts that occur earlier than they would have if unexposed (equivalently, the probability that a death randomly drawn from all exposed deaths was advanced by air pollution). If $S_0(t) > S_1(t)$ for all $t > 0$, the upper bound on the etiologic fraction is 1. This occurs when air pollution shifts all deaths earlier in time. On the stronger (and common) assumption that $R(t) = R > 1$ is constant across t (constant proportional hazard), the lower bound on the etiologic fraction is $(R - 1) / R^{R/(R-1)}$. This value is smaller than the hazard fraction $(R - 1) / R$.

The upper bound of 1 is independent of the value of R . In the limit as R decreases toward 1, the hazard fraction and the lower bound decrease toward 0 and the ratio of the lower bound to the hazard fraction decreases toward its limiting value $1/e \approx 0.37$. As R increases toward infinity, the hazard fraction, the lower bound, and the ratio of the lower bound to the hazard fraction all increase toward the limiting

can be shifted to younger ages of death. Hence, any probability distribution for age of death in Clearville for which the bar at age 70 contains no green or blue and the bar at age 80 contains no blue is possible.

value 1.¹⁷ In contrast, the value of the ratio of the upper bound to the hazard fraction, $R / (R - 1)$, depends on R . In the limit as R increases to infinity, this ratio decreases toward its limiting value, 1. But as R decreases toward 1 the ratio of the upper bound to the hazard fraction increases without limit; for $R = 1.01$ the ratio is approximately 100. Hence the bounds on the ratio of the etiologic fraction to the hazard fraction are asymmetric; the number of etiologic deaths cannot be smaller than about 37 percent of the total number of attributable deaths (i.e., of the sum over all ages of $D(t)$), but it can be much larger than the number of attributable deaths when R is close to one.¹⁸

The maximum number of etiologic deaths at age t equals the total number of deaths at age t , $f_1(t)$. This occurs when all deaths are advanced by exposure. The minimum number can be derived using the number of excess deaths at age t , $\Delta(t) = f_1(t) - f_0(t)$. If $\Delta(t) \leq 0$ it is possible that none of the deaths at t are caused by exposure and so the minimum number of etiologic deaths at t is zero. If $\Delta(t) > 0$, at minimum these excess deaths must have been caused by exposure. Combining these results, the minimum number of etiologic deaths at t is $\max(0, \Delta(t))$. Assuming constant proportional hazard with $R > 1$, the integral of $\max(0, \Delta(t))$ over t leads to the lower bound described above. Robins and Greenland (1989a) showed that the range defined by these upper and lower bounds cannot be narrowed using only mortality data. There always exist consistent causal models that generate the lower and upper bounds of the etiologic fraction.

Under the assumption that exposure never increases the chance of survival (i.e., that $S_1(t) < S_0(t)$ for all $t > 0$), then $\Delta(t)$ must be less than zero for very old ages. Clearly, if some members of the unexposed population live to be older than anyone in the exposed population, excess deaths will be less than zero when these individuals die.¹⁹ At ages for which excess deaths are less than or equal to zero, the lower bound on the fraction of deaths that are etiologic is zero; it is possible that none of the deaths at these ages are caused by exposure. An alternative (and often more plausible) explanation for excess deaths being less than zero at very old ages is that, although some of the deaths at every age are caused by exposure, the number of people surviving to very old ages is larger when exposure is low; for these ages,

¹⁷ Limiting values of the ratio of the lower bound to the hazard fraction are derived in the appendix.

¹⁸ Greenland, Robins, and coauthors describe other aspects of the relationship between the etiologic and hazard fractions in a series of papers (Robins and Greenland 1989a, 1989b, 2000, Beyea and Greenland 1999, Greenland 1999, 2000, 2012, 2015).

¹⁹ In the example of Section 1.2, excess deaths at age 100 equal minus one-quarter of the number in the birth cohort.

both the total number of deaths and the number due to exposure are larger when exposure is low rather than high.

As noted before, these results about identifiability and bounds on numbers of etiologic deaths apply to populations and EBD studies as well as to cohorts. For cases of interest, the difference between the numbers of etiologic and attributable deaths can be quite large. We illustrate with calculations for a few recent examples from the literature.

Lelieveld et al. (2015) calculate that “outdoor air pollution, mostly by PM_{2.5}, leads to 3.3 (95 per cent confidence interval 1.61 – 4.81) million premature deaths per year worldwide.” For values of R typical of the parts of the world that dominate global air-pollution-impact estimates, e.g., 1.2 for ischemic heart disease (Lelieveld et al. 2015), the hazard fraction equals 0.17. The lower and upper bounds on the etiologic fraction are 0.067 and 1. Hence the number of premature deaths due to exposure could be between 0.4 and 6 times as large as the reported value. This possible bias is large compared with the reported confidence interval ($1.61/3.3 = 0.5$, $4.81/3.3 = 1.5$). As noted by Lelieveld et al., the reported confidence interval reflects imprecision in the estimation of R , which is distinct from the inability to identify the etiologic fraction, even if R is known precisely.²⁰

Héroux et al. (2015) reported $R = 1.064$ per 10 $\mu\text{g}/\text{m}^3$ PM_{2.5} (annual mean) for all-cause natural mortality, ages 30 years and older. The bounds on the etiologic fraction are 0.023 and 1, whereas the hazard fraction is 0.060. The corresponding bounds on the ratio of the etiologic to the hazard fraction are approximately 0.4 and 17. In their reply to a letter to the editor, the authors acknowledge the difference between the hazard and etiologic fractions and cite passages from HRAPIE (2013) and COMEAP (2010) that describe the difference qualitatively (Héroux et al. 2017).

Jonson et al. (2017) estimate the number of premature deaths in Europe that may be attributed to excess NO_x emissions from diesel cars. They estimate that diesel cars contribute 0.19 $\mu\text{g}/\text{m}^3$ to ambient PM_{2.5} and use a relative-risk to concentration slope of 1.062 per 10 $\mu\text{g}/\text{m}^3$ for all-cause natural mortality in the adult population (HRAPIE 2013). A 50 percent reduction in diesel-car contribution to PM_{2.5} (0.095 $\mu\text{g}/\text{m}^3$) implies the relative risk $R = 1 + (1.062 - 1) 0.095/10 = 1.00059$ and the hazard fraction $(R - 1)/R$

²⁰ Note that the upper and lower bounds on the etiologic fraction calculated using the limits of a confidence interval or other range of values for R are farther apart than the bounds calculated for a single value of R in that range.

= 0.00059. The lower and upper bounds on the etiologic fraction are 0.00022 and 1, less than 0.4 and almost 1700 times as large as the hazard fraction, respectively.

The Committee on the Medical Effects of Air Pollutants (COMEAP 2010) analyzed EBD studies and stated that “attributable deaths are not an estimate of the number of individuals in whose earlier death air pollution has played some part” (p. 61). Moreover, “it is not plausible to think of the figure of ‘attributable’ deaths as enumerating an actual group of individuals ... the group is in reality a fiction, and ... the number of people in whose deaths air pollution has played a part might be much larger We therefore also consider it inappropriate to use the term ‘premature’ deaths to express the outcome of the burden calculation” (p. 71).

With biological or other non-epidemiological information about the relationship between exposure and mortality, the bounds may be narrowed. Robins and Greenland (1989a) note that the hypothesis that the etiologic fraction equals its upper bound (all deaths are advanced by exposure) is implausible if there are multiple causes of mortality and some are unaffected by exposure. Even when (as is common) air-pollution studies limit their outcomes to non-accidental mortality or to cardiovascular mortality, the existence of additional pathways that are not affected by air-pollution exposure seems evident; hence the possibility that the etiologic fraction equals one is implausible.

Robins and Greenland (1989a) also suggest the lower bound is implausible as it assumes all exposure-related deaths occur as early as possible and before the number of deaths in the exposed population falls to the number in the unexposed population (i.e., before $f_1(t) \leq f_0(t)$; they assume $h_1(t) > h_0(t)$ for all t). However, they present an example of a mechanism that generates an etiologic fraction equal to the theoretical minimum (see their Appendix I). The lower bound $(R - 1) / R^{R/(R-1)}$ is the result of following the cohort forever (evaluating the integral on their p. 851 over t from 0 to infinity), which may differ from the value calculated over reasonable spans. To investigate the effect of restricting the integration to realistic follow-up periods (50 years, from age 30 to 80) for small values of the hazard ratio relevant to air pollution ($R=1.1$, HRAPIE 2013), we use the baseline hazard $h_0 = 0.05/\text{year}$ suggested by Robins and Greenland (1989a).²¹ The fraction of deaths that are attributable (the hazard fraction) = 0.0909. The fraction that are etiologic calculated using the integral is 0.0374, close to the lower bound $(R - 1) / R^{R/(R-1)} = 0.0350$. Using short follow-up periods (10 and 20 years), the ratio is 0.070 and 0.053. Hence, the value

²¹ Using this constant hazard, the probability a 30 year old survives to 100 is $\exp(-70 \cdot 0.05) = 0.03$.

of the integral exceeds the lower bound but can approach it even if R is small and realistic values are chosen for the integration span.

When the population is heterogeneous (so that an individual's hazard may differ from the population hazard), Robins and Greenland (1989b) show that the etiologic fraction equals the hazard fraction $[R(t) - 1] / R(t)$ when individual increases in hazard $h_{1i}(t) - h_{0i}(t)$ are distributed independently of individual baseline hazards $h_{0i}(t)$ (where i indexes individuals). Sufficient conditions for such independence are (a) when the increase is additive to the baseline or (b) when there is no heterogeneity in the baseline hazard. If the population is well-ordered by hazard,²² the hazard fraction is a lower bound on the etiologic fraction if the effect of exposure is superadditive to baseline (people at higher baseline risk face a larger increase in risk) and is an upper bound if the effect is subadditive. Note that none of the conditions required for these results (well-ordered by hazard, additivity, superadditivity, and subadditivity) can be tested using epidemiological data (Robins and Greenland 1989b); any claim that one or another is satisfied must be supported by other sources of information.²³

2.2. Life years lost

The inability to statistically identify the number of etiologic deaths has implications for estimating the number of years of life lost and DALYs caused by air pollution and other exposures. The expected number of years of life lost due to exposure in the population (YLL) is the steady-state reduction in life expectancy, i.e., the difference between the areas under the two survival curves,

$$YLL = \int_0^{\infty} S_0(t) dt - \int_0^{\infty} S_1(t) dt. \quad (2.6)$$

²² A population is well-ordered by hazard if individuals' rank by instantaneous baseline hazard $h_{0i}(t)$ does not depend on t . More formally, if and only if, for all individuals i and k and for all t , either $h_{0i}(t) \leq h_{0k}(t)$ or $h_{0i}(t) \geq h_{0k}(t)$.

²³ Additivity and superadditivity appear to be more plausible in the context of mortality effects of air pollution. However, there is some evidence suggesting subadditivity. Silverman et al. (2012) report a case-control study on diesel-engine emissions in coal miners and observe "an attenuation of the effect of cigarette smoking among study subjects who were exposed to high levels of diesel exhaust." They identify possible mechanistic explanations for a potential subadditivity, including (a) that PAHs, nitro-PAHs, and related compounds in diesel exhaust could compete with the metabolic activation of PAHs in tobacco smoke, leading to enzyme saturation and (b) that diesel-exhaust constituents may suppress enzymes that activate or induce enzymes that detoxify carcinogens in tobacco smoke.

Because the two survival curves are identified, YLL due to exposure can be estimated using epidemiological data. However, because the number of etiologic deaths at age t is not identified, years of life lost due to exposure and conditional on death at age t , $YLL(t)$, cannot be estimated. As noted by Robins and Greenland (1991), YLL can be expressed alternatively as a function of total or of attributable deaths by age. Using total deaths, YLL can be expressed as

$$YLL = \int_0^{\infty} f_1(t) \left\{ S_0^{-1} [S_1(t)] - t \right\} dt \quad (2.7)$$

where $f_1(t)$ is the number of exposed deaths at age t and the expression in curly brackets is the horizontal distance between the survival curves at $S_1(t)$ (illustrated for $t = 75$ years by the short horizontal line segment in the middle panel of Figure 2). This horizontal distance corresponds to the years of life lost by people dying at age t if all exposed deaths are etiologic and the population is well-ordered by hazard (i.e., the order in which people die if all are exposed is the same as the order in which they would die if all are unexposed). Alternatively, using attributable deaths YLL can be expressed as

$$YLL = \int_0^{\infty} D(t) \left[\frac{1}{S_0(t)} \int_t^{\infty} S_0(u) du \right] dt, \quad (2.8)$$

where $D(t)$ is the number of attributable deaths at age t and the term in square brackets is life expectancy at age t in the unexposed population. Expression (3.8) corresponds to the situation in which etiologic deaths are equal to attributable deaths and hazard depends only on current exposure. In this case, if exposure were suddenly eliminated, individual life expectancy would immediately increase to its value for an unexposed individual of age t .

Because total exposed deaths at age t ($f_1(t)$) and attributable deaths at age t ($D(t)$) are not equal but expressions (3.7) and (3.8) yield the same total YLL, the formulas for years of life lost at age t embedded in the two expressions cannot be equal. This illustrates the fact that $YLL(t)$ is not identified by mortality data. Returning to the example in Section 1.2, although exposure may reduce life expectancy by 10 years, mortality data do not allow one to determine which of many possible relationships between age at death if exposed or unexposed is accurate. These include the possibilities that: (a) people who die at 60 if exposed would have lived 40 years longer if unexposed and no one else is affected; (b) everyone dies 10 years earlier if exposed than unexposed (as modeled by equation (2.7)); and (c) etiologic deaths are equal to attributable deaths $D(t)$, i.e., the fractions of the exposed population whose deaths are

advanced by air pollution are $1/4$, $1/16$, $1/12$, and $1/8$, and $YLL(t)$ are 25, 20, 15, and 10, at ages 60, 70, 80, and 90, respectively (as modeled by equation (2.8) and illustrated by panel D of Figure 1).²⁴

DALYs are defined as the sum of two components: years of life lost due to mortality and years lived with disability (Prüss-Üstün et al. 2003, GBD 2017 Risk Factors Collaborators 2018, WHO 2016). The standard calculation of years of life lost due to exposure multiplies estimates of attributable deaths at each age $D(t)$ by age-specific reference life expectancies (equation (2.8)). This total is correct for the population but cannot be disaggregated by age or cause of death. Estimates of years lived with disability (YLD), the other component of DALYs, are subject to a similar identification problem. These are based on estimates of the attributable cases of disease at each age, which are not necessarily equal to the etiologic cases.²⁵ As noted by Murray et al. (2003, p. 12 of 20), “If the outcome variable used in causal attribution of disease and mortality to a risk factor only involves counting of adverse events (such as disease incidence or death), it is not possible to characterize those cases whose occurrence would have been delayed in the absence of the risk factor.”

3. Economics: valuing mortality effects

The monetary value to a population of a reduction in mortality risk (as defined for benefit-cost analysis) does not depend on the number of etiologic deaths. The monetary value is defined as the sum over the affected population of each person’s private value for her risk reduction. For a small change in current mortality risk (over a short period), an individual’s private value is approximately²⁶ equal to the product

²⁴ The hazard fraction $[R(t) - 1] / R(t)$ equals 1, $1/4$, $1/3$, and $1/2$ at ages 60, 70, 80, and 90, respectively. The probability distributions for age of death conditional on surviving to age t if unexposed are: age 60, $1/4$ chance to die at ages 70, 80, 90, 100; age 70, $1/3$ chance to die at ages 80, 90, 100; age 80, $1/2$ chance to die at ages 90, 100; 90, sure to die at age 100. The number of attributable deaths (as a fraction of all exposed deaths) is $1/4 \times (1 + 1/4 + 1/3 + 1/2) = 0.5208$. Residual life expectancy (if unexposed) at age 90 is 10; at age 80 is $(20 + 10)/2 = 15$; at age 70 is $(30 + 20 + 10)/3 = 20$; and at age 60 is $(40 + 30 + 20 + 10)/4 = 25$. Using expression (2.8), life years lost due to exposure (per capita) are $1/4 \times 25 + 1/16 \times 20 + 1/12 \times 15 + 1/8 \times 10 = 10$. Using expression (2.7), YLL conditional on death at ages 60, 70, 80, and 90 = 10 and the fraction of deaths occurring at each of these ages = $1/4$; life years lost due to exposure (per capita) equals $4 \times (1/4 \times 10) = 10$.

²⁵ Moreover, YLD is calculated assuming that if an individual who develops the disease from exposure were not exposed, she would not suffer any disability from that disease (ignoring the possibility she would have developed the disease later and suffered the identical number of years lived with disability, albeit at an older age).

²⁶ The approximation is exact for an infinitesimal risk reduction and its accuracy decreases as the magnitude of the risk reduction increases. Alolayan et al. (2017) and Hammitt et al. (2019) provide bounds on the approximation error that are described below.

of the change in risk and her marginal rate of substitution between wealth and mortality risk in that period, often called her ‘value per statistical life’ (VSL). An individual’s VSL depends on her preferences and circumstances. As illustrated in Figure 3, it can be represented as the slope of her indifference curve between wealth and probability of surviving the current period at her current wealth and survival probability. VSL may also depend, inter alia, on the individual’s age, life expectancy, and anticipated future health, income, and expenses.

The term ‘value per statistical life’²⁷ is often misinterpreted as the monetary value assigned to an individual’s life, or the amount an individual would pay to avoid, or perhaps accept as compensation for, immediate death (Cameron 2010). Although estimates of VSL are used to calculate the monetary value to a population of a change in mortality risk, the VSL that is used is conceived as an average of the individual rates of substitution in the affected population. As illustrated by Figure 3, VSL is a slope conventionally measured in units of dollars (or other currency) per unit change in survival probability (i.e., from 0 to 1). Alternatively, it can be measured in units of dollars per ‘micromort’ (a survival-probability increment of 10^{-6} ; Howard 1984), dollars per ‘standardized mortality unit’ (a survival-probability increment of 10^{-4} ; Jamison et al. 2013), or other units.

For a small increase in current survival probability at age t , the individual’s monetary value is approximately equal to $v_i(t) r_i(t)$, where $v_i(t)$ is her VSL and $r_i(t)$ is the increase in survival probability; if mortality risk depends only on current exposure, $r_i(t) = h_{1i}(t) - h_{0i}(t)$ (i indexes individuals). The total value V to a population of N individuals of a reduction in their current hazards is

$$V = \sum_{i=1}^N v_i r_i \quad (3.1)$$

(to simplify notation, the dependence on age is suppressed). The population value is often approximated as

$$V \approx \bar{v} \sum_{i=1}^N r_i \quad (3.2)$$

²⁷ The similar term ‘value of a statistical life’ is perhaps more common. ‘Value per statistical life’ seems preferable as it emphasizes the concept is a rate. In the United Kingdom, VSL is described as the ‘value of a prevented fatality’ (HM Treasury 2003).

where \bar{v} is the mean of v_i and the sum of risk reductions is the decrease in the expected number of deaths occurring in the period, often called the number of ‘statistical lives saved’ or ‘premature fatalities averted’. The number of statistical lives saved (or lost) depends on the population and the change in hazard; e.g., when the hazard depends on current but not on past exposure and a population with age distribution $S_1(t)$ is suddenly unexposed, the instantaneous number of lives saved is the sum over all ages of attributable deaths $D(t)$ (equation (2.4)).²⁸

The approximation (3.2) is exact if v_i and r_i are uncorrelated in the population. One case in which they are not is when the change in hazard is correlated with wealth (e.g., because the geographic patterns of a change in air pollution and in wealth are similar). As described below, VSL is positively correlated with wealth.

The value to an individual at age t of a small continuing perturbation to her mortality hazard (e.g., due to a permanent change in air pollution) can be represented as the expected present value of the associated stream of risk reductions,

$$V_i = \int_{\tau=t}^{\infty} v_i(\tau) r_i(\tau) \sigma_i(\tau) e^{-\rho(\tau-t)} d\tau \quad (3.3)$$

where $v_i(\tau)$ is her VSL and $r_i(\tau)$ is the reduction in hazard at age τ , $\sigma_i(\tau)$ is her probability of survival to τ conditional on being alive at age t , and ρ is the rate at which she discounts future monetary flows (e.g., the interest rate on savings).²⁹ VSL $v_i(t)$ may depend on age t and on the survival function $\sigma_i(t)$, which itself depends on her new hazard function, $h_{11}(t) - r_i(t)$. Future VSL may also depend on the conditions under which the individual can save or borrow money (and hence reallocate her consumption of market goods and services over time). The value to a population is the sum over individuals of the private values V_i , which depend on the individuals’ time paths of VSL and of hazard reductions.

²⁸ If exposure is suddenly reduced from high to low and remains low thereafter (and annual births are constant), the annual number of deaths will initially fall then gradually rise to the initial level. The change in the age distribution of deaths from the initial distribution $f_1(t)$ will initially approximate the distribution of attributable deaths $D(t)$ given by equation (2.4) (the approximation will be close if the hazard depends on current but not on previous exposure) and will eventually converge to the steady-state distribution of excess deaths $\Delta(t)$ given by equation (2.5). An epidemiological study could estimate excess deaths by age as they change over time; the sum of excess deaths over all ages need not equal zero if the population is not in steady-state.

²⁹ Equation (3.3) is an approximation that is exact for infinitesimal changes in survival probability. Like equation (3.1), it neglects the change in the rate of substitution as the change in risk increases.

The economic model underlying VSL assumes an individual seeks to maximize her expected utility, which depends on her mortality hazard h in the current period and her wealth w ,

$$EU = (1 - h) u_a(w) + h u_d(w) \quad (3.4)$$

where $u_a(w)$ and $u_d(w)$ are her utility conditional on surviving and not surviving the period, respectively (Drèze 1962, Jones-Lee 1974, Weinstein et al. 1980). The function $u_d(w)$ is often described as a bequest function: it represents the utility the individual gains from knowing her wealth will benefit her heirs or others after her death. The function $u_a(w)$ is the individual's expected utility conditional on surviving the current period and depends inter alia on her future income, expenses, and risks to life, health, and wealth (Eeckhoudt and Hammitt 2001).

The individual's marginal rate of substitution of wealth for a reduction in current hazard holding expected utility constant is derived by totally differentiating equation (3.4), yielding

$$VSL = \frac{dw}{dh} = \frac{u_a(w) - u_d(w)}{(1-h)u'_a(w) + hu'_d(w)} = \frac{\Delta U(w)}{EU'(w)}, \quad (3.5)$$

where ' denotes first derivative with respect to w . The numerator of equation (3.5) equals the gain in utility from surviving the period rather than dying and the denominator represents the expected marginal utility of wealth (i.e., the utility forgone by spending a small amount to increase survival probability, which cannot be spent on other goods and services). It is conventional and reasonable to assume that $u_a(w) > u_d(w)$ (survival is preferred to death), $u'_a(w) > u'_d(w) \geq 0$ (the marginal utility of wealth is strictly larger conditional on survival than on death and the marginal utility of a bequest is non-negative), and $u''_a(w) \leq 0$ and $u''_d(w) \leq 0$ (the two utility functions for wealth are weakly concave, which implies the individual is averse to financial risk; '' denotes second derivative with respect to w). Under these conditions, VSL is strictly greater than zero and is increasing in w and in h , which implies the indifference curves are downward sloping and convex, as illustrated in Figure 3.

The economic value of a small decrease in mortality risk is well-approximated by the product of the individual's VSL and reduction in current hazard, $v_i r_i$. For larger risk reductions, the value depends on the concept of economic valuation that is adopted, compensating or equivalent surplus. Compensating surplus is the amount by which the individual's wealth would have to be reduced so that she is equally well off with the risk reduction and smaller wealth as she would be without the change (the reduction in wealth compensates for the risk reduction, leaving her no better off). Equivalent surplus is the amount

by which her wealth would need to be increased so that she is equally well off with her initial hazard and the increased wealth, or with the reduction in hazard and initial wealth (the increase in wealth has the equivalent effect on her wellbeing as the risk reduction).³⁰ For any reduction in current hazard, the compensating surplus is less than or equal to the equivalent surplus. Moreover, compensating surplus is a concave function of risk reduction (it increases at a decreasing rate) and equivalent surplus is a convex function (it increases at an increasing rate), which implies the two values diverge as the hazard reduction increases.

The magnitude of the difference between the economic value and the approximation vr can be evaluated using equation (3.5) supplemented by empirical estimates of the sensitivity of VSL to wealth (Alolayan et al. 2017, Hammitt et al. 2019). The compensating surplus c for a reduction in hazard r equals the product of r and a VSL that is between the initial VSL (at wealth w and hazard h) and the smaller VSL at hazard $h - r$ and wealth $w - c$ (illustrated in Figure 3). The effect of the change in hazard on VSL can be derived from equation (3.5); holding wealth constant, the ratio of VSL with hazard h to VSL with hazard $h - r$ is bounded above by $\frac{1 - (h - r)}{1 - h} = 1 + \frac{r}{1 - h}$. When h is much smaller than one, this ratio is close to one (r cannot exceed h). The effect of the reduction in wealth after paying c , holding risk constant, depends on the magnitude of c relative to wealth and on the sensitivity of VSL to wealth, generally measured by the wealth or income elasticity. Theory provides little guidance about the magnitude of the income elasticity; most empirical estimates are on the order of one, with some estimates as high as two or three (Viscusi and Masterman 2017, Masterman and Viscusi 2018, Hammitt and Robinson 2011, Hammitt et al. 2019). For the US, typical estimates of VSL are roughly \$9 million and annual household income averages roughly \$60,000 (Robinson and Hammitt 2016). This implies the value of a risk reduction of 1/10,000 is less than \$900 or 1.5 percent of income. Compared with VSL at the initial point, VSL at the smaller income would be 1.5 percent smaller with an income elasticity of 1 and 4.4 percent smaller with an income elasticity of 3 (i.e., VSL after paying \$900 would be \$8.6 million = \$9 million $(1 - 0.015)^3$). Combining the effects of changes in risk and income implies the value of the risk reduction of 1/10,000 is between \$900 and \$860. The equivalent surplus m for a reduction in hazard r equals the product of r and a VSL that is bounded by the initial VSL (at wealth w and hazard h) and the

³⁰ Compensating surplus is often described as the individual's 'willingness to pay for the risk reduction' and equivalent surplus as the individual's 'willingness to accept compensation to forgo the risk reduction'.

larger VSL with hazard h and wealth $w + m$. For a risk reduction of 1/10,000 and an income elasticity of 3, VSL after an increase in wealth equal to $m = \$900$ is about \$9.4 million (= \$9 million $(1 + 0.015)^3$); hence the equivalent surplus for a risk reduction of 1/10,000 is between \$900 and about \$940. For larger risk reductions, the difference between the equivalent and compensating surpluses can be arbitrarily large; compensating surplus is bounded by ability to pay (i.e., by wealth or income) but equivalent surplus can increase without limit (Hanemann 1991).

As suggested by the notation $v_i(t)$, an individual's VSL may depend on age. Yet it is common practice in applications to use the same value for all ages, in part because the evidence on how VSL varies with age is weak. The effect of age on VSL in the theoretical model is ambiguous: an increase in life expectancy conditional on surviving the current period (negatively associated with age) presumably increases the utility of survival $u_o(w)$, which increases the numerator in equation (3.5). But it may also increase the opportunity cost of spending $u_o'(w)$, hence increasing the denominator and making the effect on the ratio indeterminate. To illustrate, an individual who must support herself from a fixed quantity of wealth must husband her resources carefully (and has a high opportunity cost of spending) if her life expectancy is long but may spend profligately (with low opportunity cost) if it is short. Empirical estimates of how VSL varies with age conflict. There is some evidence that VSL rises then falls with age, but estimates of the amplitude of the change, the age at which VSL peaks, and the rates at which it rises and falls differ substantially among studies (Aldy and Viscusi 2008, Krupnick 2007).

3.1. Valuing life years lost

An alternative to the conventional approach to valuing mortality risk using VSL is to value the corresponding change in life expectancy using the concept of 'value per statistical life year' (VSLY), defined as the marginal rate of substitution between wealth and life expectancy. Any change to an individual's hazard function implies a unique change to her life expectancy. Hence the value V_i of a continuing perturbation $r(\tau)$ to her mortality hazard described in equation (3.3) can be characterized as the product of a VSLY and the corresponding increase in life expectancy (which depends on the initial hazard function and the perturbation; Hammitt 2007). A special case of equation (3.3) is an instantaneous decrease r in mortality hazard $h(t)$ for a short duration δ , which increases life expectancy by the product of the increase in probability of surviving beyond age t and life expectancy at age t , i.e.,

by $r\delta \frac{\int_t^\infty S(\tau)d\tau}{S(t)}$. The value of this risk reduction can be described as the product of VSL and the risk

reduction $r\delta$, or as the product of VSLY and the gain in life expectancy. As described above, the appropriate VSL depends on the individual's age, hazard function, and other characteristics; hence the appropriate VSLY depends on the same characteristics. VSLY also depends on the specific perturbation of the hazard function; many perturbations yield the same increase in life expectancy (e.g., a small hazard reduction over a long period and a larger hazard reduction over a shorter period) and the values of these perturbations to an individual need not be the same.³¹ If VSLY for an individual is independent of age, VSL must be proportional to age-specific life expectancy, which is generally decreasing with age. The common finding that VSL rises then falls with age implies VSLY must increase with age over some period, after which it may remain constant, increase, or decrease depending on how rapidly VSL decreases with age. For example, Aldy and Viscusi (2008) estimate that VSLY increases from age 18 to 54 then decreases to age 62 (cohort-adjusted VSLY as illustrated in their Figure 2).³²

Both the population change in hazard function and in life expectancy are identified in epidemiological studies; hence valuation using VSL or VSLY are both feasible. The inability to identify etiologic deaths does not preclude economic valuation, which is based on the monetary value of a reduction in hazard. Given uncertainty about how VSL and VSLY vary with age, life expectancy, and other factors, valuation in practice applies a constant VSL to all risk reductions or a constant VSLY to all life-expectancy gains. Which of these approaches better approximates the true value (e.g., equation (3.3)) is unknown and merits investigation.

3.2 Valuation and identification of individual risks

An individual's valuation for a reduction in exposure-related mortality risk depends on her hazard function and the change in hazard due to the change in exposure. Neither her baseline hazard nor the change can be observed. Absent additional information, an individual may perceive the change to her hazard as a random draw from the population distribution of changes in individual hazards with mean equal to the change in the population hazard; in this case, the change in population hazard is sufficient for valuation. An individual with information about her exposure, health, or other factors might perceive her change in hazard to be larger or smaller than the population average, but an analyst will typically not have access to such information and cannot do better than to estimate the individual's risk reduction as

³¹ Hammitt and Tunçel (2015) show that individuals have heterogeneous preferences over different time patterns of reductions in their hazard functions that yield the same increase in life expectancy.

³² Their analysis is limited to employed individuals aged 18 to 62 years.

equal to the population average. (In some circumstances, such as when the change in hazard differs with spatial patterns of air pollution, it may be possible to estimate the correlation with individuals' wealth or other characteristics and hence part of the correlation with VSL.)

Even if it were known whether exposure advances everyone's death by a small amount or advances the deaths of a few people by a larger amount and has no effect on other people's mortality, valuation would not be affected unless individuals had information about how their own hazard functions differ from the population average. More-refined information about an individual's prospects can affect her valuation. In the extreme case where there is a subset of the population who know their lives would be extended by a reduction in exposure and a complementary subset who know their lifespans would be unaffected, those in the first subset would value a reduction in exposure and those in the second subset would not (setting aside altruism or other ways in which their wellbeing may depend on others' survival). Whether the population value of a change in exposure is larger or smaller when the subpopulations can be identified depends on the sensitivity of individuals' values to the magnitude of the change in hazard and on the direction of the change. Hammitt and Treich (2007) show that the economic value to a population of a reduction in the population hazard is larger when the individuals whose risk reductions differ from the average cannot be identified; when individuals are identified, the increase in value of those whose risk reduction is larger than average is insufficient to offset the decrease in value of those whose risk reduction is smaller. In contrast, the total amount a population would demand as compensation for an increase in the population hazard is larger when the subsets who face larger and smaller changes than average are identified.³³ If the risk changes to all individuals are small, individual valuations are well-approximated by the product of individual VSL and risk reduction and so whether individuals facing different risk changes are identified and whether the value is for a risk reduction or risk increase should have little effect on the population value.

4. Conclusion

Table 1 summarizes some of the key concepts related to quantifying mortality and economic valuation we have addressed. Our analysis yields several conclusions:

³³ The cases described in the text are of compensating surplus. For the equivalent surplus measures, information about differences in individuals' risks increases the amount of money a population would demand to forgo a reduction in population hazard and decreases the amount it would pay to prevent an increase in population hazard. See Hammitt and Treich (2007).

Mortality data, whether from epidemiological studies or randomized controlled trials, can be used to estimate marginal probability distributions such as population hazard and survival functions. Hence concepts that can be calculated from differences in marginal distributions (e.g., attributable deaths, excess deaths, years of life lost due to exposure) can be estimated. Concepts that depend on the joint distribution of age at death under different exposures, such as etiologic deaths and years of life lost due to exposure and conditional on age or cause of death, cannot be estimated because they are not statistically identified.

The result that the etiologic fraction is not statistically identified should not be equated with the fact that epidemiological studies cannot identify individuals or groups of individuals who die (or develop a disease) because of exposure. It is stronger: one cannot even determine the number of such individuals.

If attributable deaths are reported, authors should take care to prevent readers from confusing this statistic with a count of the number of deaths caused by exposure. Numbers of deaths 'attributable to exposure' are the result of an attribution algorithm; they are not equivalent to and should not be interpreted as 'numbers of deaths caused by exposure'.

When evaluating the environmental burden of disease (EBD), total life years lost in a population due to exposure can be estimated but cannot be disaggregated into categories defined by age at death or specific cause (e.g., disease). EBD studies typically disaggregate life years lost assuming the fractions of deaths by age or cause of death that are due to exposure are equal to the hazard fractions. The fraction of deaths that are advanced because of exposure (the etiologic fraction) can be smaller or much larger than the attributable fraction; it can be as large as one and as small as about 37 percent of the attributable fraction (if the association is causal and zero otherwise). Estimates of the number of cases of disease caused by exposure, used to calculate years lived with disability, are limited by the same problem of identification. EBD studies assume the fraction of cases of a disease that are due to exposure is equal to the attributable fraction, but the etiologic fraction can be smaller or much larger. In contrast, total or average life years lost due to exposure in the population is a useful measure of burden; it is identified and can be calculated from mortality data.

As explained by Greenland (2015, p. 160), the paper that introduced the term 'etiologic fraction' (Miettinen 1974) introduced confusion by equating etiologic, excess, and attributable cases (though without using the latter term). Although Miettinen's conceptual misunderstandings were later clarified

(Greenland and Robins 1988) the three terms are still often equated, at least implicitly, as illustrated by quotations from recent literature in Section 1.1.

The monetary value of a reduction in exposure-related mortality risk, as defined for benefit-cost analysis, can be estimated using the marginal distributions that are identified using mortality data. The monetary value for a population is the sum of individuals' values, which depend only on the difference between the mortality hazard functions if exposed or unexposed; the probability distribution of the difference in age of death between exposure conditions is irrelevant. Inability to know the fraction of deaths that are etiologic does not invalidate conventional estimates of the monetary value of a change in exposure-related population mortality hazard (e.g., due to a change in air pollution). But the common practice of describing the sum of the reductions in individuals' hazards as 'statistical lives saved' can lead to misinterpretation.

Economic values are always conditional on information; if all that is known for an individual is that her change in hazard can be viewed as a random draw from the population distribution of individuals' changes in hazard, then the relevant measure for valuation is the mean change in population hazard. If more information about the changes in individuals' mortality hazards is available, it can affect their valuations and hence the total population value (though the effect on the population value will be small if individuals' changes in risk are small). Such information can also reduce uncertainty about how many deaths are etiologic.

Because any change to an individual's hazard function implies a unique change to her life expectancy, the individual's monetary value can be expressed as the product of a value per statistical life year (VSLY) and change in life expectancy. VSLY is individual-specific and may depend on the change in hazard. An individual may have preferences between alternative changes to her hazard function that yield the same increase in life expectancy (e.g., a small risk reduction over many years and a larger risk reduction over a few years); if so, she would value these differently.

To improve communication about environmental, health, and safety risks, we recommend that authors take care when describing estimates of mortality effects, health burden, and economic values of changes in mortality risk. The confusion induced by equating premature, etiologic, excess, and attributable deaths is not a trivial semantic issue. Two decades ago, Greenland (1999) characterized the issue (in his subtitle) as "a methodologic error that has become a social problem."

Appendix

The limiting values of the ratio of the lower bound to the hazard fraction as R approaches 1 and infinity

can be derived as follows. The hazard fraction $HF = \frac{R-1}{R}$ and the lower bound $LB = \frac{R-1}{R^{\frac{R}{R-1}}}$. The ratio

$$\frac{HF}{LB} = \frac{R-1}{R} \frac{R^{\frac{R}{R-1}}}{R-1} = R^{\frac{R}{R-1}-1} = R^{\frac{1}{R-1}}$$

and the logarithm of the ratio $\ln\left(\frac{HF}{LB}\right) = \frac{\ln(R)}{R-1}$. By l'Hôpital's

$$\text{rule, } \lim_{R \rightarrow 1} \left[\ln\left(\frac{HF}{LB}\right) \right] = \lim_{R \rightarrow 1} \frac{\frac{d}{dR} \ln(R)}{\frac{d}{dR} (R-1)} = \lim_{R \rightarrow 1} \frac{\frac{1}{R}}{1} = 1. \text{ Hence } \lim_{R \rightarrow 1} \left[\frac{HF}{LB} \right] = \exp(1) = e \text{ and the limiting}$$

value of LB/HF as $R \rightarrow 1 = 1/e$.

To determine the limiting value of the ratio of the lower bound to the hazard fraction as R approaches infinity, change the limiting value of R in the expression using l'Hôpital's rule above to obtain

$$\lim_{R \rightarrow \infty} \left[\ln\left(\frac{HF}{LB}\right) \right] = \lim_{R \rightarrow \infty} \frac{\frac{1}{R}}{1} = 0. \text{ Hence } \lim_{R \rightarrow \infty} \left[\frac{HF}{LB} \right] = \exp(0) = 1 \text{ and the limiting value of LB/HF as } R \rightarrow$$

$\infty = 1$.

Acknowledgments

We thank John S. Evans for encouraging us to write this paper. For helpful comments and suggestions we thank the anonymous reviewers, the editor-in-chief (Tony Cox) and area editor (Warner North), seminar participants at the Toulouse School of Economics, Newcastle University, the Society for Benefit-Cost Analysis annual conference (George Washington University), and the Advances in Decision Analysis 2019 conference (Bocconi University).

References

- Alolayan MA, Evans JS, Hammitt JK. Valuing mortality risk in Kuwait: stated-preference with a new consistency test. *Environmental and Resource Economics*, 2017; 66: 629-646.
- Aldy JE, Viscusi WK. Adjusting the value of a statistical life for age and cohort effects. *Review of Economics and Statistics*, 2008; 90: 573–581.
- Anenberg SC, Miller J, Minjares R, Du L, Henze DK, Lacey F, Malley CS, Emberson L, Franco V, Klimont Z, Heyes C. Impacts and mitigation of excess diesel-related NO_x emissions in 11 major vehicle markets. *Nature*, 2017; 545: 467-471.
- Anenberg S, Miller J, Henze D, Minjares R. A global snapshot of the air pollution-related health impacts of transportation sector emissions in 2010 and 2015. ICCT report, 2019, Washington, D.C. <https://www.theicct.org/publications/health-impacts-transport-emissions-2010-2015>.
- Atkinson RW, Kang S, Anderson HR, Mills IC, Walton HA. Epidemiological time series studies of PM_{2.5} and daily mortality and hospital admissions: a systematic review and meta-analysis. *Thorax*, 2014; 69: 660–665.
- Beyea J, Greenland S. The importance of specifying the underlying biological model in estimating the probability of causation. *Health Physics*, 1999; 76: 269-274.
- Burnett R, Chen H, Szyszkowicz M, et al. Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. *Proceedings of the National Academy of Sciences USA*, 2018; 115: 9592-9597.
- Cameron TA. Euthanizing the value of a statistical life. *Review of Environmental Economics and Policy*, 2010; 4: 161-178.
- Cohen AJ, Brauer M, Burnett R, Anderson HR, Frostad J, Estep K, Balakrishnan K, Brunekreef B, Dandona L, Dandona R, Feigin V, Freedman G, Hubbell B, Jobling A, Kan H, Knibbs L, Liu Y, Martin R, Morawska L, Pope CA, Shin H, Straif K, Shaddick G, Thomas M, van Dingenen R, van Donkelaar A, Vos T, Murray CJL, Forouzanfar MH. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. *Lancet*, 2017; 389: 1907-1918.

- COMEAP. COMEAP: mortality effects of long-term exposure to particulate air pollution in the UK. A report by the Committee on the Medical Effects of Air Pollutants (COMEAP). Chilton, Health Protection Agency, 2010. <https://www.gov.uk/government/publications/comeap-mortality-effects-of-long-term-exposure-to-particulate-air-pollution-in-the-uk>.
- Dikensoy O. Mesothelioma due to environmental exposure to erionite in Turkey. *Current Opinion in Pulmonary Medicine*, 2008; 14: 322–325.
- Dominici F, McDermott A, Daniels M, Zeger SL, Samet JM. Revised analyses of the national morbidity, mortality, and air pollution study: mortality among residents of 90 cities. *Journal of Toxicology and Environmental Health, Part A*, 2005; 68: 1071–1092.
- Drèze J. L'Utilité sociale d'une vie humaine. *Revue Française de Recherche Opérationnelle*, 1962; 6: 93–118.
- Eeckhoudt LR, Hammitt JK. Background risks and the value of a statistical life. *Journal of Risk and Uncertainty*, 2001; 23: 261-279.
- Environmental Protection Agency, Review of the national ambient air quality standards for particulate matter: policy assessment of scientific and technical information. OAQPS staff paper, EPA-452 \ R-96-013, July 1996.
- GBD 2017 Risk Factors Collaborators. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990–2017: A systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*, 2018; 392: 1923–1994.
- Goodkind AL, Tessum CW, Coggins JS, Hill JD, Marshall JD. Fine-scale damage estimates of particulate matter air pollution reveal opportunities for location-specific mitigation of emissions. *PNAS*, 2019; 116: 8775-8780.
- Greenland S. Relation of probability of causation to relative risk and doubling dose: a methodologic error that has become a social problem. *American Journal of Public Health*, 1999; 89: 1166-1169.
- Greenland S. Underestimating effects: Why causation probabilities need to be replaced in regulation, policy, and the law. *Bulletin of the Atomic Scientists*, 2012; 68: 76–83.

- Greenland S. Concepts and pitfalls in measuring and interpreting attributable fractions, prevented fractions, and causation probabilities. *Annals of Epidemiology*, 2015; 25: 155-161.
- Greenland S, Robins JM. Conceptual problems in the definition and interpretation of attributable fractions. *American Journal of Epidemiology*, 1988; 128: 1185-1197.
- Greenland S, Robins JM. Epidemiology, justice, and the probability of causation. *Jurimetrics*, 2000; 40: 321-340
- Hammitt JK. Valuing changes in mortality risk: lives saved versus life years saved. *Review of Environmental Economics and Policy*, 2007; 1: 228–240.
- Hammitt JK, Robinson LA. The income elasticity of the value per statistical life: transferring estimates between high and low income populations. *Journal of Benefit-Cost Analysis*, 2011; 2(1): Article 1.
- Hammitt JK, Treich N. Statistical vs. identified lives in benefit-cost analysis. *Journal of Risk and Uncertainty*, 2007; 35: 45-66.
- Hammitt JK, Tunçel T. Preferences for life-expectancy gains: sooner or later? *Journal of Risk and Uncertainty*, 2015; 51: 79-101.
- Hammitt JK, Geng F, Guo X, Nielsen CP. Valuing mortality risk in China: comparing stated-preference estimates from 2005 and 2016,” *Journal of Risk and Uncertainty*, 2019; 58: 167–186.
- Hanemann, WM. Willingness to pay and willingness to accept: how much can they differ? *American Economic Review*, 1991; 81: 635-647.
- Hoek G, Krishnan RM, Beelen R, Peters A, Ostro B, Brunekreef B, Kaufman JD. Long-term air pollution exposure and cardio- respiratory mortality: a review. *Environmental Health*, 2013; 12: 43.
- Howard RA. On fates comparable to death. *Management Science*, 1984; 30: 407-422.
- Héroux E, Brunekreef B, Anderson HR, Atkinson R, Cohen A, Forastiere F, Hurley F, Katsouyanni K, Krewski D, Krzyzanowski M, Künzli N, Mills I, Querol X, Ostro B, Walton H. Quantifying the health impacts of ambient air pollutants: recommendations of a WHO/Europe project. *International Journal of Public Health*, 2015; 60: 619-627.

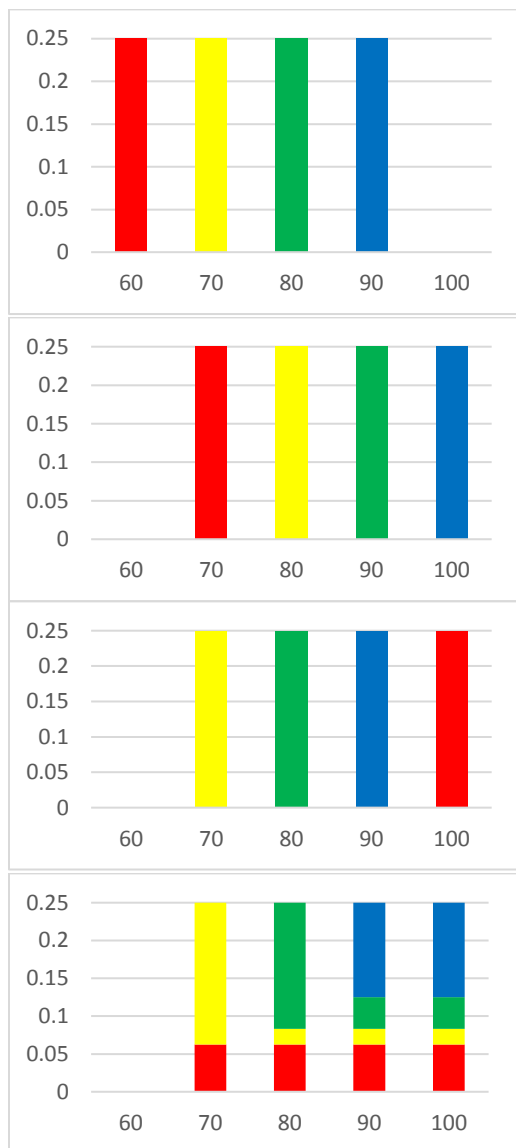
- Héroux ME, Anderson HR, Atkinson R, Brunekreef B, Cohen A, Forastiere F, Hurley F, Katsouyanni K, Krewski D, Krzyzanowski M, Künzli N, Mills I, Querol X, Ostro B, Walton H. Response to: Premature deaths attributed to ambient air pollutants: let us interpret the Robins–Greenland theorem correctly. *International Journal of Public Health*, 2017; 62: 339-341.
- HM Treasury. *The green book: appraisal and evaluation in central government*. London: TSO, 2003.
- HRAPIE. *Health risks of air pollution in Europe: HRAPIE project recommendations for concentration–response functions for cost–benefit analysis of particulate matter, ozone and nitrogen dioxide*. World Health Organization, 2013.
- International Agency for Research on Cancer (IARC). *Overall evaluations of carcinogenicity: an updating of IARC monographs volumes 1 to 42. IARC monographs on the evaluation of carcinogenic risks to humans*. Supplement / World Health Organization, International Agency for Research on Cancer. Lyon, France, 1987; 7: 1–440.
- Jamison DT, Summers LH, Alleyne G, Arrow KJ, Berkley S, Binagwaho A, et al. *Global health 2035: a world converging within a generation*. *Lancet*, 2013; 382: 1898-955.
- Jasani B, Gibbs A. Mesothelioma not associated with asbestos exposure. *Archives of Pathology and Laboratory Medicine* 2012; 136: 262–267. Dockery DW, Pope CA III, Xu X, Spengler JD, Ware JH, Fay ME, Ferris BG Jr, Speizer FE. An association between air pollution and mortality in six U.S. cities. *New England Journal of Medicine*, 1993; 329: 1753–1759.
- Jones-Lee MW. The value of changes in the probability of death or injury. *Journal of Political Economy*, 1974; 82: 835–849.
- Jonson JE, Borcken-Kleefeld J, Simpson D, Nyiri A, Posch M, Heyes C. Impact of excess NO_x emissions from diesel cars on air quality, public health and eutrophication in Europe. *Environmental Research Letters*, 2017; 12: 094017.
- Krupnick, A. Mortality-risk valuation and age: stated preference evidence. *Review of Environmental Economics and Policy*, 2007; 1: 261-282.
- Landrigan PJ, Fuller R, Acosta NJR, et al. *The Lancet commission on pollution and health*. *Lancet*, 2017. [http://dx.doi.org/10.1016/S0140-6736\(17\)32345-0](http://dx.doi.org/10.1016/S0140-6736(17)32345-0).

- Lelieveld J, Evans JS, Fnais M, Giannadaki D, Pozzer, A. The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature*, 2015; 525: 367-371.
- Lim SS, Vos T, Flaxman AD, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*, 2012; 380: 2224-2260.
- Masterman CJ, Viscusi WK. The income elasticity of global values of a statistical life: stated preference evidence. *Journal of Benefit-Cost Analysis*, 2018; 9: 407–434.
- Miettinen OS. Proportion of disease caused or prevented by a given exposure, trait, or intervention. *American Journal of Epidemiology*. 1974; 99: 325-332.
- Murray CJ. Quantifying the burden of disease: the technical basis for disability-adjusted life years. *Bulletin of the World Health Organization*, 1994; 72: 429–445.
- Murray CJL, Ezzati M, Lopez AD, Rodgers A, Vander Hoorn S. Comparative quantification of health risks: Conceptual framework and methodological issues. *Population Health Metrics*, 2003; 1: 1-20.
- Ostro, B. Outdoor air pollution: assessing the environmental burden of disease at national and local levels, *Environmental Burden of Disease series, No. 5*, World Health Organization, Geneva, 2004.
- Poole, C. A history of the population attributable fraction and related measures. *Annals of Epidemiology*, 2015; 25: 147-154.
- Pope CA III, Rodermund DL, Gee MM. Mortality effects of a copper smelter strike and reduced ambient sulfate particulate matter air pollution. *Environmental Health Perspectives*, 2007; 115: 679–683.
- Pope CA III, Schwartz J, Ransom MR. Daily Mortality and PM₁₀ pollution in Utah Valley. *Archives of Environmental Health*, 1992; 47: 211-217.
- Prüss-Üstün A, Mathers C, Corvalán C, Woodward A. Introduction and methods: Assessing the environmental burden of disease at national and local levels. *Environmental burden of disease Series, No. 1*. World Health Organization, 2003.
- Rabl A. Interpretation of air pollution mortality: Number of deaths or years of life lost? *Journal of the Air & Waste Management Association*, 2003; 53: 41-50.

- Robins JM, Greenland S. Estimability and estimation of excess and etiologic fractions. *Statistics in Medicine*, 1989a; 8: 845-859.
- Robins JM, Greenland S. The probability of causation under a stochastic model for individual risk. *Biometrics*, 1989b; 45: 1125-1138.
- Robins JM, Greenland S. Estimability and estimation of expected years of life lost due to a hazardous exposure. *Statistics in Medicine*, 1991; 10: 79-93.
- Robinson LA, Hammitt JK. Valuing reductions in fatal illness risks: implications of recent research. *Health Economics*, 2016; 25: 1039–1052.
- Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology*. 3rd Edition. Philadelphia: Lippincott, Williams and Wilkins, 2008.
- Rowe AK, Powell KE, Flanders WD. Why population attributable fractions can sum to more than one. *American Journal of Preventive Medicine*, 2004; 26: 243-249.
- Silverman DT, Samanic CM, Lubin JH, Blair AE, Stewart PA, Vermeulen R, Coble JB, Rothman N, Schleiff PL, Travis WD, Ziegler RG, Wacholder S, Attfield MD. The diesel exhaust in miners study: a nested case–control study of lung cancer and diesel exhaust. *Journal of the National Cancer Institute*, 2012; 104: 855-868.
- Suzuki E, Yamamoto E, Tsuda T. On the relations between excess fraction, attributable fraction, and etiologic fraction. *American Journal of Epidemiology*, 2012; 175:567-575.
- Testa JR, Cheung M, Pei J, et al. Germline BAP1 mutations predispose to malignant mesothelioma. *Nature Genetics*, 2011; 43: 1022-1025.
- Tsiatis A. A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Sciences*, 1975; 72: 20-22.
- Viscusi WK, Masterman CJ. Income elasticities and global values of statistical life. *Journal of Benefit-Cost Analysis*, 2017; 8: 226–250.
- Wang J, Xing J, Mathur R, Pleim JE, Wang S, Hogrete, Gan C-M, Wong DC, Hao J. Historical trends in PM2.5-related premature mortality during 1990-2010 across the Northern Hemisphere. *Environmental Health Perspectives*, 2017; 125: 400–408.

Weinstein MC, Shepard DS, Pliskin JS. The economic value of changing mortality probabilities: a decision-theoretic approach. *Quarterly Journal of Economics*, 1980; 94: 373-396.

WHO. Ambient air pollution: a global assessment of exposure and burden of disease. World Health Organization, 2016.



A. Probability distribution of age of death in Smokeville.

B. Probability distribution of age of death in Clearville if all deaths are advanced (by 10 years).

C. Probability distribution of age of death in Clearville if only deaths at age 100 are advanced (by 40 years).

D. Probability distribution of age of death in Clearville if only attributable deaths are advanced. Fractions of deaths that are etiologic are 1 , $1/4$, $1/3$, and $1/2$ at ages 60, 70, 80, and 90, respectively.

Figure 1. Probability distribution of age of death showing alternative possible fractions of deaths advanced by age. Colors correspond to age at which individual would die in Smokeville.

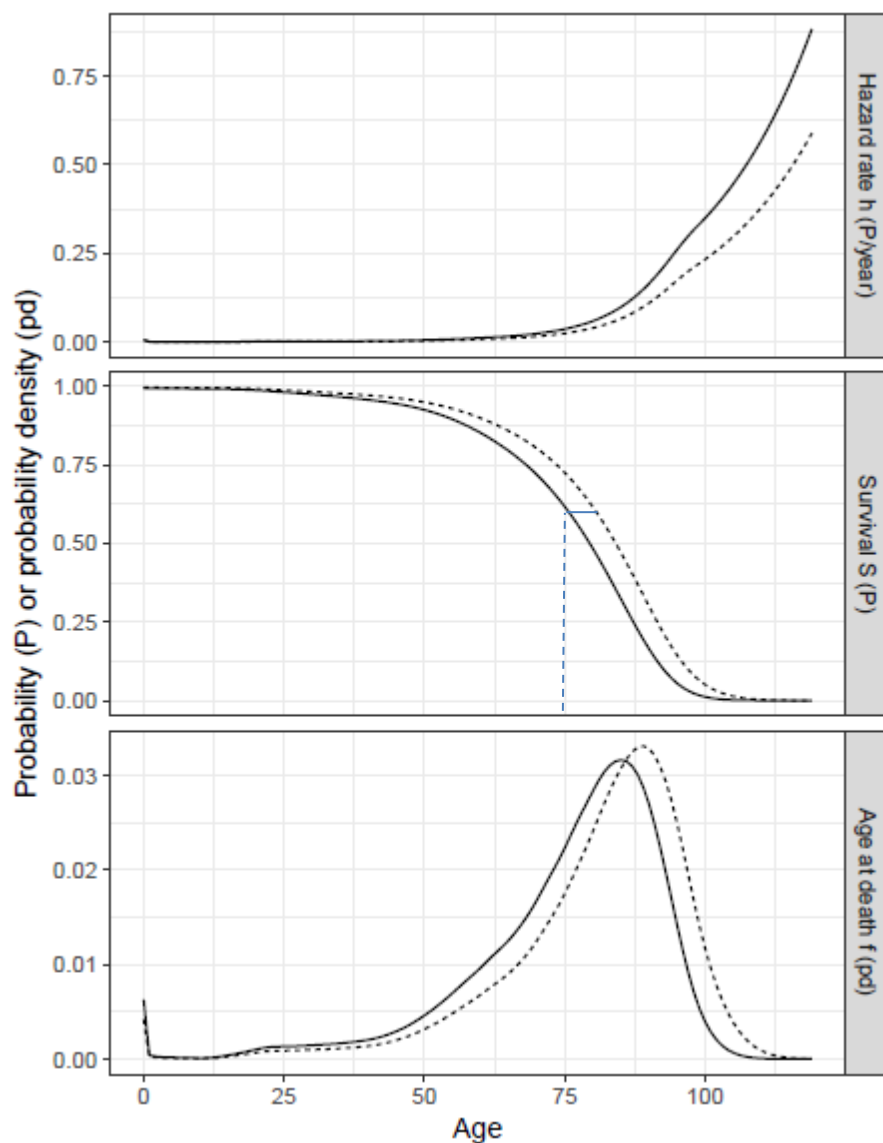


Figure 2. Annual hazard, survival, and marginal distribution of deaths by age. Exposed (h_1 , solid curves) are for U.S. males (2014 period lifetable www.ssa.gov/oact/STATS/table4c6.html). Unexposed (h_0 , dashed curves) are for annual hazard = (2/3) annual hazard if exposed. The short horizontal line segment in the middle panel is the years of life lost by individuals dying at age 75 when all deaths are etiologic, i.e., $S_0^{-1}[S_1(75)] - 75$ in expression (7).

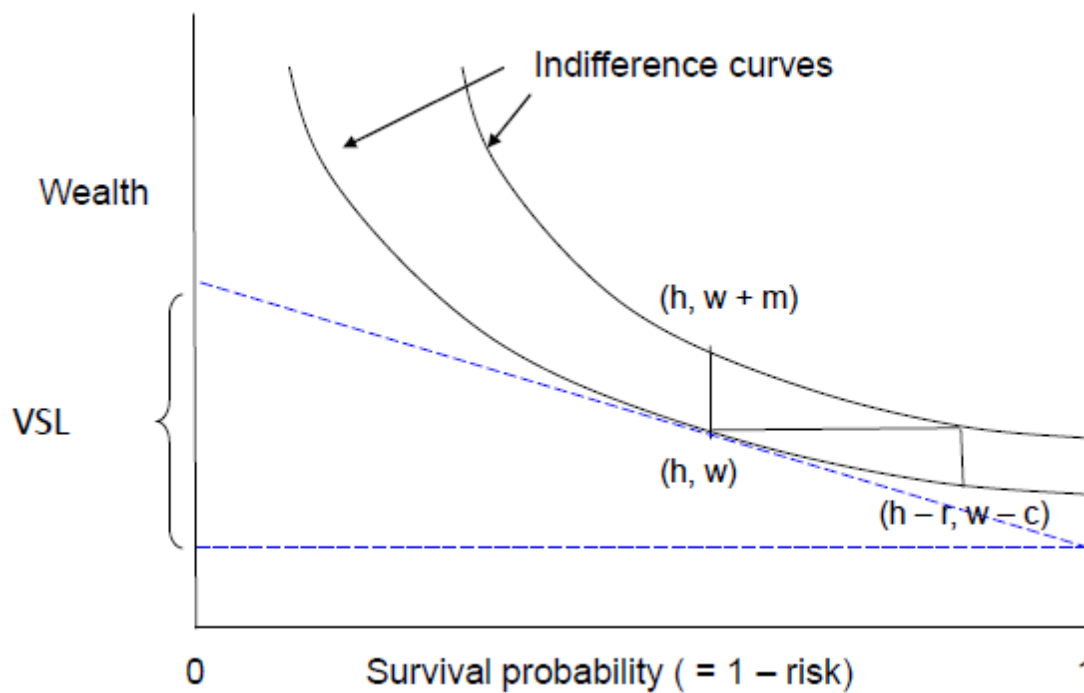


Figure 3. VSL is the slope of the individual's indifference curve at current wealth w and survival probability $1 - h$. The value of a reduction in risk from h to $h - r$ can be measured as compensating variation c or equivalent variation m . For $r \approx 0$, $c \approx m \approx r \cdot \text{VSL}$. (For clarity, the value of r in the figure is much larger than is usually relevant.)

Table 1. Summary of concepts

Concept	Definition	Comments
Relative risk	$R = (\text{hazard in exposed population}) / (\text{hazard in unexposed population})$.	R is usually > 1 and is often assumed to be independent of age ('constant proportional hazard').
Hazard fraction	$(R - 1) / R$, based on relative risk R.	If R is age-dependent, so is the hazard fraction. Often called 'attributable fraction' or 'population attributable fraction'.
Attributable fraction	Fraction of deaths in the exposed population attributed to exposure, equal to the hazard fraction.	
Etiologic fraction	Fraction of deaths in the exposed population caused by exposure (i.e., that occur earlier than if these individuals were never exposed), under the assumption that exposure is never protective.	Cannot be identified from mortality data. Lower bound depends on R: $(R-1)/R^{R/(R-1)}$ (assuming $R > 1$ and independent of age) and upper bound equals 1 (assuming $R > 1$).
Attributable deaths (in a specified time period)	Expected number of deaths in the exposed population that would not have occurred in the period if the hazard during only that period were equal to the hazard in the unexposed population.	Calculated as (deaths in exposed population during period) x (hazard fraction). Always > 0 if $R > 1$. Equals decrease in deaths in exposed population if suddenly unexposed and hazard depends only on current exposure. Cannot be meaningfully summed over periods.
Statistical lives saved (in a specified time period)	Expected reduction in the number of deaths in the period for a specified population and reduction in hazard. Can be < 0 if hazard is increased.	Equal to the sum over individuals of the reductions in individuals' hazards in the period. If the hazard in the exposed population is instantaneously reduced to the hazard in the unexposed population, statistical lives saved = attributable deaths for that instant.
Excess deaths	Expected steady-state difference in number of deaths between the exposed and unexposed populations, i.e., the number of deaths that would not have occurred absent exposure.	The sum of excess deaths across all ages is zero. Excess deaths are typically > 0 for younger ages and < 0 for older ages. Excess deaths at age $t <$ attributable deaths at age t (if exposure reduces chance of surviving to t).
Etiologic deaths	Number of deaths in the exposed population that occur earlier than if these individuals were never exposed. Number of deaths caused by exposure, under the assumption that exposure is never protective.	Cannot be identified from mortality data. Etiologic deaths at age $t \geq$ minimum of excess deaths at age t and zero and \leq all deaths in exposed population at age t . Lower bound depends on R but is no smaller than $1/e$ (≈ 0.37) times attributable deaths (assuming $R > 1$ and independent of age); upper bound is all deaths in exposed population at age t .

Premature deaths	No technical definition.	population (assuming $R > 1$). Seems to imply etiologic deaths, often used to describe attributable deaths.
Expected years of life lost (YLL) among the exposed	Sum over age of (exposed deaths) x (residual life expectancy in exposed population).	Can be calculated in total and identified by covariates like age or cause of death.
Expected years of life lost (YLL) due to exposure	Difference between life expectancy for unexposed and exposed populations. Equal to and often calculated as sum over age of (attributable deaths) x (residual life expectancy if unexposed).	Can be identified from mortality data for entire population but not for subpopulations defined by age of death or cause (disease).
Expected years lived with disability (YLD) due to exposure	YLD among the exposed that are attributed to exposure. YLD among the exposed are a function of prevalence and/or incidence of sequelae and disability weights.	Calculated as (YLD among the exposed) x (hazard fraction). Assumes elimination of all disability among individuals affected by exposure if exposure is reduced, i.e., assumes no disability among these individuals even if they would suffer the disease absent exposure. Potentially biased estimate of YLD due to exposure = YLD among the exposed – YLD among the exposed had exposure been absent.
Expected disability adjusted life years (DALYs) due to exposure	DALYs among the exposed that are attributed to exposure. DALYs among the exposed = YLL among the exposed + YLD among the exposed.	DALYs due to exposure by age of death or cause (disease) cannot be identified from epidemiological data (see YLL due to exposure). Are potentially biased (see YLD due to exposure).
Value per statistical life (VSL)	Individual-specific rate of substitution between money and infinitesimal change in survival probability for a specified (short) period.	Monetary value to an individual of an increase in period survival probability \approx VSL x (probability change). VSL may depend on wealth, age, life expectancy, health, and other factors. The approximation is good for monetary values that are small compared with the individual's wealth or income.
Value per statistical life year (VSLY)	Individual-specific rate of substitution between money and infinitesimal change in life expectancy.	Monetary value to an individual of an increase in life expectancy \approx VSLY x (change in life expectancy). VSLY depends on the same factors as VSL. VSLY need not be the same for all changes to the hazard function yielding the same increase in life expectancy. Similar to VSL, the approximation is best for monetary values that are small compared with wealth or income.

Table 2. Value of risk reduction in Smokeville

t	S(t)	Type A			Type B		
		h(t)	r(t)	Value	h(t)	r(t)	Value
60	1	1/4	1/4	$v^A(60) \cdot 1/4 \cdot 1 \cdot \rho^{(a-60)}$	1/4	1/4	$v^B(60) \cdot 1 \cdot 1 \cdot \rho^{(a-60)}$
70	3/4	1/3	1/3	$v^A(70) \cdot 1/3 \cdot 3/4 \cdot \rho^{(a-70)}$	0	0	0
80	1/2	1/2	1/2	$v^A(80) \cdot 1/2 \cdot 1/2 \cdot \rho^{(a-80)}$	0	0	0
90	1/4	1	1	$v^A(90) \cdot 1 \cdot 1/4 \cdot \rho^{(a-90)}$	0	0	0
100	0	0	0	0	1	0	0
Total				$V^A \leq v^A(60)$			$V^B = 1/4 v^B(60)$

Note: $V^A = v^A(60)$ if $v^A(t)$ is independent of t and $\rho = 1$ (zero discount rate).