Harvard Center for Risk Analysis

May 2020

Risk in Perspective

Premature deaths, statistical lives, and the economic value of mortality risk

James K. Hammitt

"Eliminating fine-particulate air pollution (PM2.5) in the United States would prevent 100,000 deaths per year at an annual value of $1 trillion."[1]

Mortality risk associated with air pollution and other environmental hazards is often described by the number of "premature deaths" and the economic value of reducing the hazard as equal to the number of "lives saved" multiplied by the "value per statistical life." These descriptions are misleading. The number of deaths caused by many environmental hazards is unknowable from epidemiological information and the economic value is the sum of many individuals' willingness to pay for small reductions in age-specific mortality risk.

Consider a simple example: There are two cities that are identical in every respect, except air pollution and mortality. In Smokeville, the air is polluted. One-quarter of the population dies at age 60, one-quarter at 70, one-quarter at 80, and one-quarter at 90 years. In Clearville, the air is clean and one-quarter of the population dies at ages 70, 80, 90, and 100 years. Assume that individuals live their entire lives in their city of birth, that the difference in mortality is caused by air pollution, and that if pollution in Smokeville were eliminated the age-specific mortality rates would instantaneously fall to those in Clearville.

How many people in Smokeville die because of air pollution? How much would it be worth to them to reduce air pollution and mortality risk to the levels in Clearville? Does the economic value depend on how many deaths are caused by air pollution?

Terms like "premature deaths" and "lives saved" are misleading – the number of deaths advanced by exposure cannot be determined from mortality data alone.

James K. Hammitt

Professor of Economics and Decision Sciences

Director, Harvard Center for Risk Analysis

[1] Apocryphal but consistent with current science for the 200 million residents aged 30 years and older.

Etiologic deaths

Etiologic deaths are those in which a specified hazard plays some role, in the sense that the deaths would have occurred later if the hazard were not present. Deaths in Smokeville "caused" by air pollution are deaths of individuals who would have lived longer in Clearville.

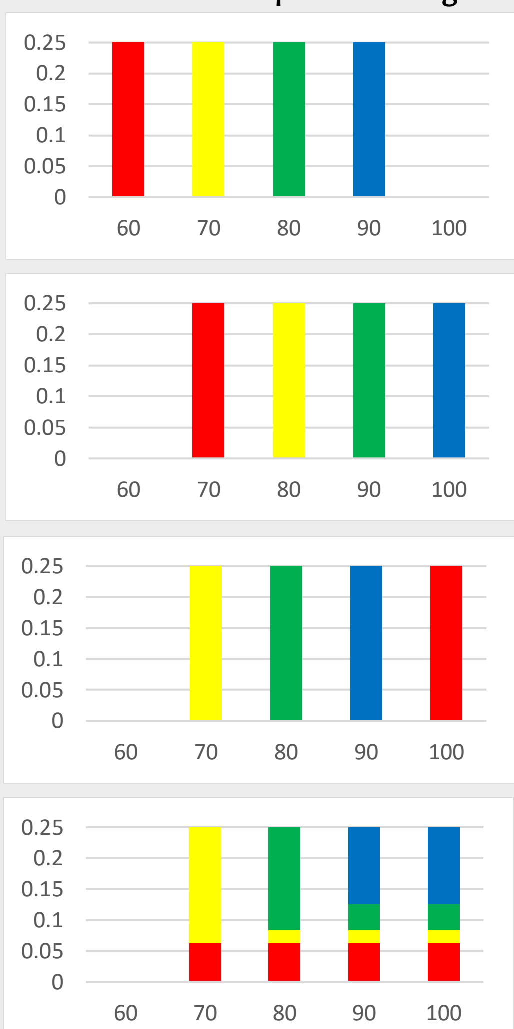
Epidemiological studies and randomized controlled trials are used to estimate the mortality effect of environmental and other hazards.

A perfect epidemiological study that compared Smokeville and Clearville would accurately measure the frequency distributions of age at death in the two cities and would confirm that there are no observable differences between the populations and environments except air pollution. But if the specific individuals who die because of air pollution cannot be identified (if there is no marker of death caused by air pollution), the fraction of deaths in Smokeville that are etiologic can be any number between one and one-quarter.

To clarify, assume that each individual has two potential ages at death, t_s if she lives in Smokeville and t_c if she lives in Clearville. A Smokeville resident's death is etiologic if and only if she would have lived longer in Clearville, $t_c > t_s$. The inferential problem is that only one of these ages at death can be observed; the other is counterfactual.

Figure 1 illustrates the frequency distributions of age at death in Smokeville and Clearville. Panel A is for Smokeville and panels B, C, and D are for Clearville. The colors denote age at death in Smokeville.

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



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, 52 percent of total deaths.

Mortality data can determine:

- *The fraction of deaths attributable to a hazard*
- *Total life years lost to a hazard*

It cannot determine:

- *The fraction of deaths advanced by a hazard*
- *Life years lost to a hazard by age or cause of death*

The economic value of reducing exposure does not depend on how many lives are lost to a hazard

- *Unless those whose deaths are advanced can be (partially) identified*

Panel B shows the case in which everyone in Smokeville would have lived exactly 10 years longer in Clearville; the etiologic fraction is one. Panel C shows the case in which people who die at 60 would have lived to 100 in Clearville and people who die at 70, 80, or 90 would have died at the same age in Clearville; the etiologic fraction is one-quarter.

Panel D shows an intermediate case: deaths "attributable" to air pollution. The attributable fraction of people that would have lived longer in Clearville depends on their age of death in Smokeville. The fraction is one for deaths at age 60, one-quarter at age 70, one-third at age 80, and one-half at age 90, about 52 percent in total. Among those who would have lived longer, their age at death in Clearville is equally likely to be any of the possible ages. That is, people who would have lived beyond 60 would be equally likely to die at 70, 80, 90, and 100; those who would have lived beyond 70 would be equally likely to die at 80, 90, and 100, and so on.

If the people who would have lived longer in Clearville cannot be identified, it is impossible to distinguish between the different possibilities. The etiologic fraction could be one (Panel B; everyone would have lived 10 years longer); it could be one-quarter (Panel C; only the one-quarter of people who die at 60 would have lived longer, by 40 years); it could be 0.52 (Panel D).

Indeed, the etiologic fraction could be any number between one-quarter and one. Imagine a series of buckets of colored sand, labelled 60, 70, 80, 90, and 100. Initially, the sand is distributed as in Panel A. On the assumption that air pollution never causes someone to live longer, then any of the infinite number of ways of leaving sand in its current bucket or shifting it to buckets for older ages that completely fills the buckets for ages 70, 80, 90, and 100 is possible.

Technically, the issue is one of statistical identification. By observing the ages at death for everyone in the two cities, one can estimate the marginal distributions of age at death in Smokeville (t_s) and in Clearville (t_c). Equivalently, one can estimate the survival functions (probability of surviving to each age) and hazard functions (probability of dying at age t conditional on being alive then) for the two cities. But the joint distribution of the two potential ages at death and the distribution of the difference between them ($t_c - t_s$) are not statistically identified and cannot be estimated from mortality data alone.

The total loss of life expectancy due to air pollution is statistically identified: it is 10 years. But the loss of life expectancy conditional on age at death is not: individuals who die at age 60 in Smokeville lose between 10 years and 40 years. Individuals who die at age 70 in Smokeville lose between zero and 30 years, and so on.

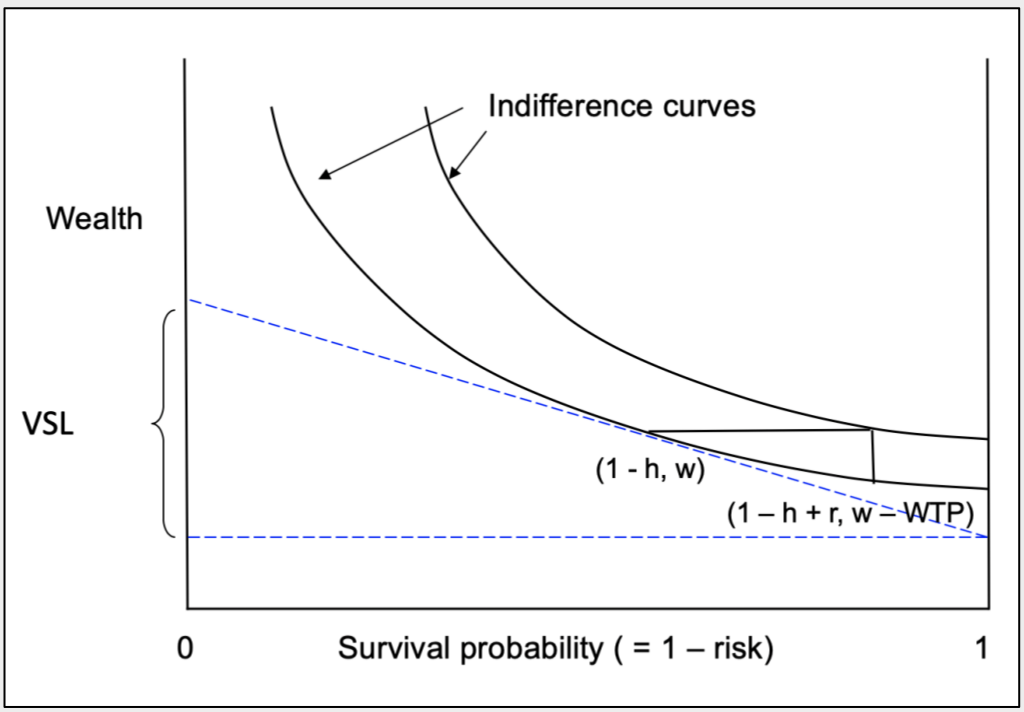
Similarly, the loss of life expectancy conditional on cause of death is not identified. Assume that some of the deaths at age 60 but none of the deaths at older ages are by heart attack. The people who die of heart attack lose between 10 and 40 years.

Value of risk reduction

What is the economic value of reducing air pollution in Smokeville to the levels in Clearville? Does it depend on how many people would live longer?

The economic value of a reduction in mortality risk depends on individuals' willingness to exchange money for small changes in mortality risk in a specified short time period (e.g., the current year). As illustrated in Figure 2, an individual's tradeoffs between her probability of surviving the current period and wealth (that can be spent now or later) can be characterized by a set of indifference curves that are downward sloping and convex to the origin. Her willingness to pay for a small increase in survival probability is approximately equal to the slope of the indifference curve passing through her current position multiplied by the gain in survival probability. The slope is conventionally measured in units of dollars per one-unit change in survival probability and called her "value per statistical life" (VSL). As shown in Figure 2, an individual's VSL depends on her wealth and survival probability; it also depends on factors that affect the benefit of surviving the current period, such as future health and life expectancy.

Figure 2. VSL is the slope of the individual's indifference curve at current wealth w and survival probability 1 – h. Willingness to pay for a reduction in risk from h to h – r can be approximated as r • VSL.



The value to a population of reducing mortality risk in the current period is the sum of their individual values, $\sum r_i v_i$, where r_i is individual i's risk reduction and v_i is her VSL, and the summation is over all individuals. This value is often calculated and described as $\bar{v} \sum r_i$, i.e., the product of average VSL and the sum of the risk reductions (which equals the expected reduction in the number of deaths in the period, so-called "lives saved"). This second formula is an approximation of the first formula that is exact only if the risk reductions and VSLs in the population are uncorrelated.

The value of a continuing change in mortality risk, such as from permanently reducing air pollution in Smokeville, is the sum of the individuals' values for this continuing change. The general formula for an individual's value is $\sum v(t)r(t)\,s(t)d^{t-a}$, where the sum is over future ages t, v(t), r(t), and s(t) are her VSL, risk reduction, and probability of surviving to t, d is a discount factor, and her current age is a.[2] Note that both VSL and risk reduction may depend on age.

For a person in Smokeville younger than 60, the value of eliminating air pollution does not depend on the fraction of deaths that are etiologic. To see this, assume she could know whether she is of type B or C, corresponding to panels B and C of Figure 1. If she is of type B, eliminating air pollution will postpone her death by 10 years (from 60 to 70, or from 70 to 80, and so on). If she if of type C, eliminating air pollution will postpone her death by 40 years (from 60 to 100) with probability one-quarter and with probability three-quarters it will have no effect. Her values of the risk reduction conditional on being of each type are calculated in Table 1.

The value of the risk reduction is the same for the two types. For type C, the value of reducing risk at age 60, $v^C(60)$, is the value of postponing death from 60 to 100. For type B, the value of reducing risk at age t, $v^B(t)$, is the value of postponing death from age t to age t + 10. Hence the summation shown in the last row for type B is the sum of the values of postponing death from 60 to 70, from 70 to 80, from 80 to 90, and from 90 to 100, which is the same as the value of postponing death from 60 to 100.

An alternative way to see that the value of eliminating air pollution does not depend on the etiologic fraction is to recognize that each individual substitutes a lottery with equal chances of dying at ages 70, 80, 90, and 100 for a lottery with equal chances of dying at ages 60, 70, 80, and 90. Since she will confront only one of these lotteries, why would it matter how the other lottery would have turned out?

Table 1. Value of risk reduction in Smokeville

		Type B		Type C	
t	s(t)	r(t)	Value	r(t)	Value
60	1	1/4	$v^B(60) \cdot 1/4 \cdot 1 \cdot d^{60-a}$	1/4	$v^C(60) \cdot 1/4 \cdot 1 \cdot d^{(60-a)}$
70	3/4	1/3	$v^B(70) \cdot 1/3 \cdot 3/4 \cdot d^{(70-a)}$	0	0
80	1/2	1/2	$v^B(80) \cdot 1/2 \cdot 1/2 \cdot d^{(80-a)}$	0	0
90	1/4	1	$v^B(90) \cdot 1 \cdot 1/4 \cdot d^{(90-a)}$	0	0
100	0	0	0	0	0
Total			$V^B = \frac{1}{4}\sum_{t=60}^{90} v^B(t)d^{t-a}$		$V^C = \frac{1}{4}v^C(60)d^{60-a}$

[2] More accurately, the term v(t) is the average rate of substitution between wealth and survival probability for the risk reduction r(t); if r(t) is small, v(t) is the VSL (i.e., the marginal rate of substitution). The discount factor $\boldsymbol{d^t} = \left(\frac{\mathbf{1}}{\mathbf{1+n}}\right)^{\boldsymbol{t}}$ where n is the discount rate.

Conclusion

Terminology commonly used to describe the mortality effects and economic value of reducing environmental hazards is misleading. The number of people who die earlier than they would have because of exposure to air pollution or another hazard is not knowable from epidemiological data (or from randomized controlled trials) alone. Measures of the effect of exposure to a hazard that are commonly reported, like the number of premature deaths, the burden of disease, and the lives saved by reducing exposure, are better described as quantities that are "attributable" to the hazard. Attribution is based on comparing deaths in the exposed population with those that would have occurred if the age-specific mortality rates had suddenly fallen to the rates that would exist without exposure to the hazard.

Describing valuation as the number of (attributable) lives saved multiplied by the value of a life is also misleading; what is valued is not a big change (saving a life) for a few people but a small change (slightly increasing survival probability for a time) for many.

The quote that begins this essay describes the benefits of eliminating fine-particulate air pollution as preventing "100,000 deaths per year at an annual value of $1 trillion." It would be more accurate and less misleading to describe the benefit as "decreasing annual mortality risk by 5/10,000 on average (increasing life expectancy by 4 months) with an average individual value of $5,000."

For further reading

This essay is based on J.K. Hammitt, P. Morfeld, J.T. Tuomisto, and T.C. Erren, "Premature Deaths and Statistical Lives: Identification, Quantification, and Valuation of Mortality Risks," *Risk Analysis* 40(4): 674-695, April 2020. The paper is available (open access) at https://onlinelibrary.wiley.com/doi/full/10.1111/risa.13427

Peer reviewer

John S. Evans, Harvard Center for Risk Analysis

HARVARD UNIVERSITY

Harvard Center for Risk Analysis

Center for Health Decision Science

Harvard T.H. Chan School of Public Health

https://www.hsph.harvard.edu/hcra