EPIDEMIOLOGY, JUSTICE, AND THE PROBABILITY OF CAUSATION

Sander Greenland
James M. Robins*

ABSTRACT: The concept of "probability of causation" forms the basis of important legal standards, legislation, and compensation schemes, which in turn use epidemiologic data to estimate the probability of causation. This usage is a misapplication of epidemiology, because it has been shown that without imposing restrictive biologic assumptions, epidemiologic data cannot supply estimates of the probability of causation. Although the misapplication of the probability of causation concept responds to the need to resolve cases in a rational and consistent manner, this need does not justify continued misuse of epidemiologic data in compensation decisions. Compensation schemes and legal standards must recognize that an upper bound on the probability of causation cannot be determined from epidemiologic data alone; biologic models also are needed. Although equitable compensation schemes can be formulated without reference to the probability of causation, all schemes must deal with fundamental methodologic uncertainties in estimation.

CITATION: Sander Greenland and James M. Robins, Epidemiology, Justice, and the Probability of Causation, 40 Jurimetrics J. 321–340 (2000).

This article concerns the distinction between the excess incidence caused by an exposure (the "attributable fraction") and the probability that the exposure caused an individual's disease (the "probability of causation"). Our points are not

---

*Sander Greenland is Professor, Department of Epidemiology, UCLA School of Public Health, Los Angeles, CA 90095-1772. James M. Robins is Professor, Departments of Epidemiology and Biostatistics, Harvard School of Public Health, Boston, MA 02115-9957. The authors are grateful to Ben Armstrong, Jan Beyea, Geoffrey Howe, Hal Morgenstern, Mark Parascandola, Diana Petitti, Jonathan Samet, Noel Weiss, and an anonymous reviewer for their valuable comments on this manuscript.

new. They were anticipated in early risk-analysis literature,[1] were thoroughly described by the late 1980s,[2] and have been repeated since.[3] Nonetheless, many epidemiologists and health physicists serving as expert consultants and witnesses continue to equate attributable fractions with the probability of causation.[4]

In part, this may be because equating the probability of causation to the attributable fraction leads to systematic underestimation of the potential range for the probability. Those whose personal interests or values make them sympathetic to the defense would have to go against those interests or values to recognize the distinction.[5]

Most of our points are important only for cases where disease costs are sensitive to the time of disease incidence.[6] For example, consider a mother of two children who gave birth at ages 35 and 40. We ordinarily would expect fatal

---

[1] *See* Louis A. Cox, Jr., *Probability of Causation and the Attributable Risk*, 4 RISK ANALYSIS 221 (1984) [hereinafter Cox, Jr., *Probability of Causation*]; Louis A. Cox, Jr., *Statistical Issues in the Estimation of Assigned Shares for Carcinogenesis Liability*, 7 RISK ANALYSIS 71 (1987) [hereinafter Cox, Jr., *Statistical Issues*]; Fritz A. Seiler & Bobby R. Scott, *Mixtures of Toxic Agents and Attributable Risk Calculations*, 7 RISK ANALYSIS 81 (1987); Stephen W. Lagakos & Frederick Mosteller, *Assigned Shares in Compensation for Radiation-Related Cancers*, 6 RISK ANALYSIS 345 (1986).

[2] *See generally* Sander Greenland & James M. Robins, *Conceptual Problems in the Definition and Interpretation of Attributable Fractions*, 128 AM. J. EPIDEMIOLOGY 1185 (1988) [hereinafter Greenland & Robins, *Conceptual Problems*]; James M. Robins & Sander Greenland, *Estimability and Estimation of Excess and Etiologic Fractions*, 8 STAT. MED. 845 (1989) [hereinafter Robins & Greenland, *Estimability and Estimation*]; James M. Robins & Sander Greenland, *The Probability of Causation Under a Stochastic Model for Individual Risks*, 45 BIOMETRICS 1125 (1989) [hereinafter Robins & Greenland, *Probability of Causation*]; James M. Robins & Sander Greenland, *Estimability and Estimation of Expected Years of Life Lost Due to a Hazardous Exposure*, 10 STAT. MED. 79 (1991) [hereinafter Robins & Greenland, *Hazardous Exposure*].

[3] *See* Sander Greenland, *Relation of Probability of Causation to Relative Risk and Doubling Dose: A Methodologic Error that Has Become a Social Problem*, 89 AM. J. PUB. HEALTH 1166 (1999); Jan Beyea & Sander Greenland, *The Importance of Specifying the Underlying Biologic Model in Estimating the Probability of Causation*, 76 HEALTH PHYSICS 1 (1999); Mark Parascandola, *What's Wrong with the Probability of Causation?*, JURIMETRICS J. 29 (1999).

[4] *See, e.g.*, Linda A. Bailey et al., *Reference Guide on Epidemiology*, *in* REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 168–69 (Federal Judicial Center ed., 1994); P. Cole, *Causality in Epidemiology, Health Policy, and Law*, 27 ELR NEWS ANALYSIS 10279 (1997); FREDERICK A. METTLER & ARTHUR C. UPTON, MEDICAL EFFECTS OF IONIZING RADIATION ch. 10 (2d ed. 1995); B. Armstrong & G. Thériault, *Compensating Lung Cancer Patients Occupationally Exposed to Coal Tar Pitch Volatiles*, 53 OCC. ENV. MED. 160 (1996); Noel Weiss, *General Concepts of Epidemiology*, *in* MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY 321 (David L. Faigman et al. eds., 1997); Richard Wakeford et al., *A Review of the Probability of Causation and Its Use in a Compensation Scheme for Nuclear Industry Workers in the United Kingdom*, 74 HEALTH PHYSICS 1 (1998).

[5] *See generally* Ernst Wynder et al., *The Wish Bias*, 43 J. CLIN. EPIDEMIOLOGY 619 (1990); Ernst Wynder, *Tobacco as a Cause of Lung Cancer: Some Reflections*, 146 AM. J. EPIDEMIOLOGY 687 (1997). This article was stimulated by our work for plaintiffs' lawyers. Like defense experts, we have been paid by interested parties, although (as discussed below) the issues we raise reveal weaknesses in plaintiff as well as defense arguments.

[6] Our points regarding effect reversal and estimation uncertainty apply even when the time of disease is not important.

cancer in this woman to result in greater emotional and financial loss the earlier it occurs. Even one year of life lost can translate into a considerable cost. Undoubtedly, a just compensation scheme should be sensitive to incidence time. Yet, the probability of causation, even if correctly estimated, is insensitive to the impact of exposure on disease timing. In our view, this insensitivity creates perverse compensation schemes based on the probability of causation, *even if the latter were known exactly*, unless damages are awarded in proportion to years of life lost.[7]

Some authors have defended the probability of causation as the best available quantity for making administrative and judicial decisions.[8] Such claims are simply wrong when incidence time is important. It is possible to develop compensation schemes that are based on correct interpretations of epidemiologic relations and that are appropriately sensitive to incidence time and years of life lost.[9]

Before discussing these approaches, we begin by describing how epidemiologic data does *not* determine the probability of causation.

I. "MORE PROBABLE THAN NOT" IS NOT THE SAME AS A RATE RATIO ABOVE TWO

Consider a hypothetical case of a woman diagnosed at age 45 with bone cancer (a very rare disease) after 25 years of employment in a nuclear waste reprocessing facility. She files suit against her employer, claiming her disease was due to occupational radiation exposure. According to current standards, judgment should favor the plaintiff if and only if "it is more probable than not" that the exposure "causally contributed to" or "was a substantial contributing factor to" her disease.[10] This standard is often rephrased as a requirement that the "probability of causation" exceed 50%, where "probability of causation" is the probability that exposure causally contributed to the development of the individual's disease.

To give an objective meaning to the word "probability," one must provide some sort of frequency or sampling framework. A common interpretation here is that it represents the frequency with which cases having the same exposure and measured-covariate history as the case at issue (i.e., same dose, age, sex, etc.) have exposure as a cause. With this interpretation, the probability of causation is equal to the "etiologic fraction," the fraction of cases with that exposure and covariate history for which the exposure played a role in disease etiology.[11]

---

[7] *See* Robins & Greenland *Hazardous Exposure*, *supra* note 2

[8] *See* Victor P Bond, *The Need for Probabilities in Cancer Litigation*, 29 NUCLEAR NEWS 62 (1986)

[9] *See* Robins & Greenland, *Hazardous Exposure*, *supra* note 2. A few such schemes are outlined *infra* Section IV

[10] *See* Bailey et al., *supra* note 4, Cole, *supra* note 4, Bond, *supra* note 8, David E. Lilienfeld & Bert Black, *The Epidemiologist in Court Some Comments*, 123 AM J EPIDEMIOLOGY 961 (1986)

[11] *See* Greenland & Robins, *Conceptual Problems*, *supra* note 2, Robins & Greenland, *Estimability and Estimation*, *supra* note 2, Robins & Greenland, *Probability of Causation*, *supra*

A. The Common Error

Although any estimate would be disputed, let us suppose that we know that the occupational exposure suffered by the plaintiff increases the bone cancer rate in 45-year-old women by 1.5-fold, or 50% above what would have occurred absent the occupational radiation.[12] In other words, the *causal* rate ratio[13] or "true relative risk" for the elevation in the rate above the natural background among 45-year-old women is 1.5. According to many authorities,[14] knowledge of this rate ratio allows one to compute the probability of causation, using the following formula.

$$PC = (I_1 - I_0)/I_1 = (IR - 1)/IR, \tag{1}$$

where PC stands for "probability of causation," I_1 and I_0 are the incidence rates that 45-year-old women would experience with and without exposure, and $IR \equiv I_1/I_0$ is the incidence-rate ratio (often called the "relative risk"). If we define $RF = (IR - 1)/IR$, where RF stands for "rate fraction," then formula (1) simply asserts that $PC = RF$, or that the probability of causation equals the rate fraction. Applying this formula to our plaintiff yields $(1.5 - 1)/1.5 = 33\%$ for the probability that occupational radiation causally contributed to the plaintiff's cancer.

Most epidemiologists and statisticians would label RF with more familiar terms such as "attributable fraction"or "attributable risk" among the exposed.[15] However, these terms are used to refer to many different quantities,[16] and we use the more specific term "rate fraction"[17] to indicate that I_1 and I_0 represent numbers of cases per woman year at risk and are therefore incidence *rates*,[18] not probabilities. The quantity RF has also been called the "assigned share" in the legal and risk analysis literature by those who recognize failings of the $PC = RF$ equation.[19]

Given the above formula, the "more probable than not" standard requires $PC > 0.50$ for the plaintiff to prevail. If one asserts that $PC = RF$, this standard

---

note 2, Sander Greenland & Kenneth J. Rothman, *Measures of Effect and Measures of Association*, *in* MODERN EPIDEMIOLOGY 47–64 (Kenneth J. Rothman & Sander Greenland eds., 2d ed. 1998).

[12] The *incidence rate* is the number of new cases that develop per woman per year, or "per woman-year." For example, if 3 cases develop over a year in a population of women whose average size is 30,000 over the year, the rate that year is 3 cases per 30,000 women per year, or 3/30,000 = 0.0001 cases per woman per year.

[13] *See* Robins & Greenland, *Probability of Causation*, *supra* note 2; Greenland & Rothman, *supra* note 11.

[14] Authorities cited, *supra* notes 4, 8 & 10; Bert Black & David E. Lilienfeld, *Epidemiologic Proof in Toxic Tort Litigation*, 52 FORDHAM L. REV. 732 (1984).

[15] *See* Wynder et al., *supra* note 5.

[16] *See* Greenland & Robins, *Conceptual Problems*, *supra* note 2.

[17] *See id.*; Robins & Greenland, *Estimability and Estimation*, *supra* note 2; Robins & Greenland, *Probability of Causation*, *supra* note 2; Greenland & Rothman, *supra* note 11.

[18] *See* Wynder et al., *supra* note 5; Greenland & Rothman, *supra* note 11.

[19] *See supra* note 1.

requires that $(IR - 1)/IR > 0.5$, which reduces to $IR > 2$. Thus, in a trial where the judge imposes $IR > 2$ as the "more probable than not" standard,[20] studies finding a rate ratio less than 2 would be taken as evidence *against* causation of the plaintiff's disease by the exposure at issue. Our plaintiff has $IR = 1.5$ and therefore fails this standard.

The error in this reasoning is that, even if the effect IR of exposure on the rate is known, *the causal rate fraction does not equal or even approximate the probability of causation except under restrictive assumptions*.[21] Perhaps the simplest assumption is the following independence-of-background (IOB) assumption about mechanisms of exposure action:[22] The incidence of cases caused by exposure is independent of the incidence of cases not caused by exposure (i.e., cases due to background causes).

This assumption corresponds to the "no interaction with background risks" assumption that was delineated by Cox[23] in his pioneering criticisms of the $PC = RF$ formula. Many plausible disease mechanisms do not satisfy this assumption. For example, if the disease is a tumor that requires a sequence of distinct mutations and the exposure at issue causes only one of the mutations, following the classic Armitage-Doll multistage biologic model of carcinogenesis,[24] the incidence of cases caused by exposure will be positively associated with incidence from other causes, because any factor that increases the rate of other necessary mutations will multiply the incidence of cases caused and not caused by exposure to the same degree. We know of no cancer or other important chronic disease for which current biomedical knowledge allows one to exclude mechanisms that violate the assumptions needed to claim that $PC = RF$.

The result is that judicial applications of the formula $PC = RF$ have been without foundation in fact.[25] Furthermore, absent indefensible assumptions, PC and RF may be as far apart as their logical limits allow. It is possible for exposure to have causally contributed to every case of disease, even if the exposure elevates the rate only slightly.[26] That is, it is possible to have $PC = 1$ even when the causal rate ratio IR is close to 1 and so RF is close to zero. However, the

---

[20] Such a standard has been applied in both legislation (Orrin G. Hatch, *Medical-Legal Aspects of Radiation Induced Cancer*, HEALTH PHYSICS SOC. NEWSLETTER, Dec. 1983, at 6-8; S. 921, 98th Cong., 129 CONG. REC. 7111 (1983)) and judicial decisions (Hall v. Baxter Healthcare, 947 F. Supp. 1387 (D. Or. 1996); Pick v. American Med. Sys., 958 F. Supp. 1151 (E.D. La. 1997)).

[21] *See* Greenland & Robins, *Conceptual Problems*, *supra* note 2; Robins & Greenland, *Estimability and Estimation*, *supra* note 2; Robins & Greenland, *Probability of Causation*, *supra* note 2.

[22] *See* Robins & Greenland, *Estimability and Estimation*, *supra* note 2; Greenland, *supra* note 3.

[23] *See* Cox, Jr., *Probability of Causation*, *supra* note 1.

[24] *See* Peter Armitage & Richard Doll, *Stochastic Models for Carcinogenesis*, *in* PROCEEDINGS OF THE FOURTH BERKELEY SYMPOSIUM (Jerzy Neyman ed., 1961).

[25] *See* Greenland, *supra* note 3; Parascandola, *supra* note 3; Diana Petitti, *Reference Guide on Epidemiology*, 36 JURIMETRICS J. 159 (1996).

[26] *See* Greenland & Robins, *Conceptual Problems*, *supra* note 2, at 1186; Robins & Greenland, *Probability of Causation*, *supra* note 2, at 1132.

converse is false. Epidemiologic data *does* place a nonzero lower bound on the probability of causation when RF > 0. This lower bound is always less than RF when IR is constant, although under commonly used assumptions, RF approximates this lower bound.[27] The Appendix contains a detailed description of these relationships.

In summary, current judicial standards use the formula PC = RF, which relies on assumptions that are unwarranted in typical cases. The assumptions are biological, not methodological, and have nothing to do with disease rarity. Furthermore, there are no validity issues here because IR as defined above *is* the effect of exposure on the incidence rate in the population of 45-year-old women. Violating the assumptions can lead to a complete breakdown of the PC = RF formula, with RF likely to underestimate PC.[28]

In sworn statements and declarations by epidemiologists and statisticians enrolled as expert witnesses, we found only one case where the expert was aware that biologic assumptions were required to equate PC to RF.[29] Even in that case, the expert attempted to rationalize PC = RF with a number of incorrect assertions.[30]

B. How the Standard Fails

There are many technical details that arise in a rigorous attempt to connect the probability of causation or etiologic fraction to rate ratios.[31] Nonetheless, some simple examples illustrate conditions in which the probability of causation is near or far from the rate fraction.

Suppose that we have assembled a large cohort of women who, with respect to known risk factors for bone cancer, are indistinguishable from the plaintiff in our example. In particular, suppose these women have had the same occupational and background radiation exposure. Suppose further that this cohort experienced three cases of bone cancer at age 45, but would have experienced only two cases at that age absent the occupational exposure. Finally, suppose that the occupational exposure has only a small effect on the total number of person years contributed by women in the study.[32] We then have a causal rate ratio of 3/2 = 1.5 and a causal rate fraction of (1.5 − 1)/1.5 = 33%, as before.

All this information is quite a lot to be given—a cohort observed without error in which the exposure effect on the rate is known. It is far more than we

---

[27] *See* Robins & Greenland, *Probability of Causation*, *supra* note 2, at 1129

[28] *See* Greenland & Robins, *Conceptual Problems*, *supra* note 2, at 1192–93, Robins & Greenland, *Estimability and Estimation*, *supra* note 2, at 850–51, Cox, Jr , *Statistical Issues*, *supra* note 1, at 72–73, Lagakos & Mosteller, *supra* note 1, at 374; Greenland, *supra* note 3, at 1167

[29] *See* Hall v Babcock & Wilcox Co , 69 F Supp.2d 716 (W D Pa 1999)

[30] *See id* at 722–26

[31] *See generally* Robins & Greenland, *Estimability and Estimation*, *supra* note 2, Robins & Greenland, *Probability of Causation*, *supra* note 2.

[32] This assumption is not essential to the point, but it is reasonable and greatly simplifies the mathematics

ever have in reality. Yet, it is not all we need to calculate the probability of causation. The reason is simple: we cannot determine the etiology of any individual case of bone cancer. Consequently, we cannot tell if the three cases that occurred at age 45 in this exposed cohort overlap with the two cases that would have occurred at this age absent exposure.

Maybe the overlap is complete. Perhaps the two cases that would have occurred without exposure are women whose cancer was unaffected by exposure. If so, they must be two of the three cases that did occur at age 45. The remaining case at that age, 33% of the total, must have had exposure involved in her disease etiology because she would not have developed cancer at age 45 without the exposure. Put another way, it could be that two of the three cases at age 45 were "background" cases that occurred independently of exposure. Thus, if we picked one of the three at random, there would only be a 33% chance it would be a case with exposure as a contributory cause. Epidemiologists and others must have this independent background model in mind when they assert that PC = RF. Nonetheless, the IOB model is not the only reasonable possibility.

To comprehend just how wrong the PC = RF assertion can be, consider that the three cases that occurred at age 45 *may not overlap at all* with the two cases that would have occurred at this age absent exposure. For example, it is possible that exposure interacts with background factors to advance the incidence time of *all* bone cancer cases. This would happen if the cancer is the endpoint of a pathologic process whose rate is accelerated by radiation exposure. Thus, it could be that the two background cases (the two women who would have gotten bone cancer at age 45 even without exposure) instead got their cancer years earlier because of exposure; while the three cases that occurred at age 45 would not have occurred until years later absent exposure. In fact, it could be that exposure causally contributed to *all* cancers at *all* ages by accelerating *all* the incidence times. If so, the probability of causation would be 100%. Yet, despite this ubiquity of harm, the causal rate ratio for this cohort would remain 3/2 = 1.5 and the causal rate fraction would thus remain 33%.

C. Facing an Epidemiologic Limit

While there is considerable literature on radiation carcinogenesis, the literature is incapable of demonstrating beyond a reasonable doubt that radiation induction of human bone cancer does or does not follow either the IOB or "affects all cases" models. Nor can the literature rule out models that would yield a probability of causation anywhere between 33% and 100% when the rate ratio is 1.5.

More generally, population data alone cannot establish an upper bound for the probability of causation if the factor in question has a net causal effect (IR > 1).[33] This limitation is an important manifestation of the fact that data on

---

[33] *See* Greenland & Robins, *Conceptual Problems*, *supra* note 2, at 1189–90, Robins &

incidence and prevalence will always be compatible with a wide variety of underlying causal mechanisms, even if the data are free from all error and bias.[34] Only further information on biologic mechanisms enables us to narrow the possibilities beyond those allowed by the population data.

Along with these logical limitations of epidemiologic data, there are always the methodologic limitations of random errors and systematic biases. And, while there are always numerous speculative theories, there is rarely an abundance of data on the mechanisms of induction for noninfectious human diseases, such as typical neoplastic, cardiovascular, connective-tissue, and neurologic diseases. The only logical conclusion is that, in most of these cases, scientific attempts to rule out a probability of causation above 50% are futile. Claims of success, at best, have rested on questionable assumptions and, at worst, have no foundation at all. Thus, when an exposure is known to be harmful in some cases, available data from epidemiology and biology are simply incapable of telling us whether a given case was "more probably than not" harmed by exposure.

Epidemiology can modestly contribute to estimating the probability of causation in that epidemiologic data can be used to estimate *lower bounds* for the probability of causation, subject to assumptions that must be checked against the data.[35] Under certain modeling assumptions, for example, rare exposure effects coupled with independent and infrequent censoring, the rate fraction can approximate this lower bound.[36] Nonetheless, even if the assumptions are accepted, the resulting lower bound does no more than rule out PC values below the bound: it still leaves possible any larger PC value, including 100%. Unfortunately, the probability of causation controversy involves cases where IR < 2, and hence, RF < ½. In such cases, PC can easily be two or more times RF; more generally, establishing that RF < ½ does not establish, or even render it probable, that PC < ½.

Greenland, *Estimability and Estimation*, *supra* note 2, at 850–51; Robins & Greenland, *Probability of Causation*, *supra* note 2, at 1129–30.

[34] *See* references cited, *supra* note 33; Jack Siemiatycki & Duncan C. Thomas, *Biological Models and Statistical Interactions: An Example from Multistage Carcinogenesis*, 10 INT'L J. EPIDEMIOLOGY 383, 386 (1981); Sander Greenland & Charles Poole, *Invariants and Noninvariants in the Concept of Interdependent Effects*, 14 SCAND. J. WORK ENV. HEALTH 125 (1988); W. Douglas Thompson, *Effect Modification and the Limits of Biological Inference from Epidemiologic Data*, 44 J. CLIN. EPIDEMIOLOGY 221, 228 (1991).

[35] *See* Robins & Greenland, *Estimability and Estimation*, *supra* note 2, at 850–51; Robins & Greenland, *Probability of Causation*, *supra* note 2, at 1126.

[36] *See* Robins & Greenland, *Probability of Causation*, *supra* note 2, at 1129–30.

II. THE FALSITY OF VARIOUS UNSUPPORTED ASSERTIONS

One aspect of the probability of causation issue that is most distressing is the frequent neglect of basic canons of scientific reasoning.[37] To our knowledge, every precept of scientific and mathematical investigation teaches that a theory (such as the assertion PC = RF) should not be treated as fact or even treated as likely unless overwhelming supportive evidence in the form of validated mathematical proofs or extensive data has accumulated and any apparently contrary evidence has been explained within the theory. In other words, the burden of proof should be entirely on the proponents of an assertion.

Instead, however, we have observed expert assertions treated as if they were true unless proven false. For example, an expert will assert that PC = RF, either as an axiom or subject to some assumption, without offering any mathematical proof or data to support this assertion. The assertion nonetheless is treated as a fact and the burden is placed on any critics to show that the assertion is false.

This section provides some examples of false variations on the PC = RF assertion that have been put forth and accepted as facts, despite a lack of supporting evidence.

A. Lifetime Risks and Rare Diseases

Some authors have mistakenly assumed that if the lifetime risk of the disease is low, the rate fraction will approximate the probability of causation.[38] The previous bone cancer example provides a counterexample to this assumption. Because the rate when exposed can be arbitrarily close to the rate when unexposed, even when PC = 1, it follows that lifetime risk when exposed can be arbitrarily close to the lifetime risk when unexposed, even when PC = 1. The situation is similar when risk up to a particular age is considered.[39] This phenomenon would occur, for example, when the disease is the endpoint of a cumulative process and exposure accelerates either the process itself or the exhaustion of repair mechanisms.[40] Such mechanisms cannot be ruled out for most chronic diseases and are especially consistent with the physiology of diseases defined by cumulative pathologic changes, such as atherosclerosis, systemic sclerosis, and many neurologic conditions.

---

[37] *See generally* CARL G. HEMPEL, PHILOSOPHY OF NATURAL SCIENCE (1966).

[38] *See* Armstrong & Thériault, *supra* note 4, at 162.

[39] *See* Robins & Greenland, *Estimability and Estimation*, *supra* note 2, at 855–56; Robins & Greenland, *Probability of Causation*, *supra* note 2, at 1125–26.

[40] *See* Greenland, *supra* note 3, at 1167.

B. Multiplicative Models and PC

Some expert witnesses have asserted that PC = RF if the effect of exposure is simply to multiply the disease rate by a constant amount.[41] This assertion is unsupported by mathematics or data. Results in Robins and Greenland[42] show that the incidence rates can follow a multiplicative model and yet the probability of causation can still be 100%. For example, suppose the baseline disease rate is of Weibull form[43] and that exposure shortens time to disease by a constant proportion. It then follows from the equivalence of accelerated-failure and proportional-hazards models under a Weibull hazard[44] that the exposure effect will be multiplicative *and* PC = 1.

C. Age-Specific Incidence Curves

Some expert witnesses have claimed that the age-specific incidence curves for radiogenic cancers are incompatible with biologic models where the probability of causation exceeds RF. These claims are false. For *any* situation in which the age-specific incidence curve under exposure always exceeds that under nonexposure (including all radiogenic cancer cases), there will be biologic models that perfectly predict the observed incidence and that imply a probability of causation of one.[45] In particular, any such pair of curves can be generated by a process in which exposure advances the incidence time of every case.

D. Biologic Arguments

We have shown that epidemiologic data cannot reject the possibility that exposure harmed everyone (and hence PC = 1) if exposure is positively associated with risk. Under the same conditions, epidemiologic data cannot reject values for the probability of causation anywhere in a range from less than the rate fraction to 1.[46] Therefore, any scientific attempt to narrow the possible range for PC must depend on biologic evidence.

We have seen experts assert that biology supports the claim that PC = RF. However, we have seen no data to support these claims. There are few, if any, noninfectious diseases for which the mechanisms of exposure action have been delineated in enough detail to establish the relation of PC to RF. In light of this absence of data, scientific arguments can only bear on the relative plausibility of

---

[41] This is called a "multiplicative model."

[42] *See* Robins & Greenland, *Probability of Causation*, *supra* note 2, at 1131.

[43] *See* JOHN D. KALBFLEISCH & ROSS L. PRENTICE, STATISTICAL ANALYSIS OF FAILURE TIME DATA (1980).

[44] *See* DAVID R. COX & D. OAKES, ANALYSIS OF SURVIVAL DATA (1984).

[45] *See* Robins & Greenland, *Estimability and Estimation*, *supra* note 2, at 850–51; Robins & Greenland, *Probability of Causation*, *supra* note 2, at 1134.

[46] *See* authorities cited, *supra* note 45.

various possible relations. Unfortunately, there are no generally accepted methods for determining relative plausibility in the absence of decisive evidence.

Suppose one could argue that a biologic model where PC = 1 is implausible. Such an argument would not support the hypothesis that PC = RF. On the contrary, models where the probability of causation equals the rate fraction could be just as implausible as those in which PC = 1. As an illustration, consider the "independence-of-background" assumption in evaluating the impact of fenfluramine on the occurrence of symptomatic valvular regurgitation.[47] In this context, IOB corresponds to assuming that this impact is unrelated to pre-existing valvular abnormalities—a highly implausible assumption. Similar considerations undermine the credibility of the IOB assumption in most examples of carcinogenesis and atherogenesis. For example, there are good reasons to suspect that radiation interacts with genetic susceptibility in carcinogenesis.[48] Yet it seems that many experts make the IOB assumption cavalierly or implicitly when interpreting the rate fraction, and they assert that evidence supports the assumption when no such evidence exists.

III. FURTHER REASONS WHY RF WILL NOT EQUAL PC

A. Heterogeneity of Background Rates

We have shown that there can be large bias in equating the causal rate fraction RF to the probability of causation PC and that the probability of causation for each cohort member can exceed the true (causal) rate fraction for the cohort.[49] These problems may be viewed as stemming from the fact that heterogeneity of background rates within a population can lead to completely counterintuitive relations among rates in differently exposed subpopulations. In some instances, there can be apparent effect reversal where none exists.[50] Further unrealistic assumptions, such as independence of exposure effects from background risk across individuals, are needed to ensure that such phenomena do not occur.[51]

---

[47] *See* Richard B. Devereux, *Appetite Suppressants and Valvular Heart Disease*, 339 NEW ENGLAND J. MED. 765, 766 (1998).

[48] *See* E. J. Hall, *Molecular Biology in Radiation Therapy: The Potential Impact of Recombinant Technology on Clinical Practice*, 30 INT. J. RAD. ONCOL. BIOL. PHYS. 1019, 1028 (1994).

[49] *See* Robins & Greenland, *Probability of Causation*, *supra* note 2, at 1132–33.

[50] *See* James W. Vaupel & Anatoli I. Yashin, *Heterogeneity Ruses: Some Surprising Effects of Selection on Population Dynamics*, 39 AM. STAT. 176 (1985).

[51] *See* Robins & Greenland, *Estimability and Estimation*, *supra* note 2, at 855–56; Robins & Greenland, *Probability of Causation*, *supra* note 2, at 1126; Vaupel & Yashin, *supra* note 50.

B. Competing Risks

Up to this point, we have ignored competing risks, as does the literature we criticize. In nearly all situations, however, there will be other endpoints, such as death from lung cancer or heart disease, that remove a person from risk of the disease at issue. The impact of such "competing risks" on the probability of causation can be large,[52] and yet it is not adequately accounted for by the rate fraction formula. This problem becomes especially worrisome at older ages, when competing risks become common.

To get a sense of the possibilities, suppose in a given cohort three cases of bone cancer would have occurred among women ages 65 to 74 without occupational radiation exposure, but the exposure caused two of these cases to occur earlier (but still in the 65-74 age range) and caused one of these cases to die of leukemia before getting bone cancer. Both of the two bone cancer cases would have exposure involved in their etiology, for a probability of causation of 100%. Yet the causal rate fraction for bone cancer among these women would be approximately $(2 - 3)/2 = -0.50$. This negative value only reflects the fact that the exposure killed a potential bone cancer case before the cancer occurred. Thus, even though the rate fraction properly reflects the fact that exposure reduced the *rate* of bone cancer among women ages 65 to 74, it conceals the fact that exposure harmed every one of the cases.

The preceding example illustrates a more general problem with assuming that ordinary rate calculations or hazard models automatically and successfully account for competing risks. Standard methods assume that competing risks occur independently of the study disease, an assumption that is dubious in many situations and questionable in most.[53] For example, it is plausible that individuals especially susceptible to chemical or radiation carcinogenesis in one organ system are also especially susceptible to cancers of other systems because they possess unmeasured genetic risk factors or poorly measured dietary risk factors. This correlation of susceptibilities leads to dependencies like those just illustrated and can render useless any conventional estimate of the probability of causation *even if exposure acts independently of the background.*[54]

C. Effect Reversal

So far, we have assumed that exposure never prevents or forestalls the disease at issue. Occasionally, however, the assumption may be challenged. For example, some authors have suggested that low dose radiation exposures may

---

[52] *See* Robins & Greenland, *Estimability and Estimation*, *supra* note 2, at 855, Robins & Greenland, *Probability of Causation*, *supra* note 2, at 1132.

[53] *See* Robins & Greenland, *Probability of Causation*, *supra* note 2, at 1126, KALBFLEISCH & PRENTICE, *supra* note 43

[54] *See* Robins & Greenland, *Estimability and Estimation*, *supra* note 2, at 856; Robins & Greenland, *Probability of Causation*, *supra* note 2, at 1131

protect some individuals from the same cancer types caused by radiation.[55] If a predominantly causal exposure is sometimes preventive, the rate fraction can underestimate the probability of causation even if the exposure effects are independent of the background.[56] If the frequency of these preventive effects is not negligible compared to the frequency of causal effects, this underestimation is important even if the time of disease incidence is not relevant to damages.

To illustrate this point, suppose that the causal and preventive effects of exposure occur disjointly of one another and independently of background. Any preventive effects will reduce the total number of exposed cases without reducing the number of causal effects (i.e., cases caused by exposure). Consequently, these preventive effects will decrease the incidence rate among the exposed and decrease the rate ratio and rate fraction. Yet, by decreasing the total number of cases, they will increase the proportion of cases caused by exposure (the etiologic fraction). In this fashion, the coexistence of causal and preventive effects will drive the rate fraction and probability of causation further apart than they would have been otherwise.

IV. COMPENSATION SCHEMES

If the $PC = RF$ formula is so deeply flawed, why has it become entrenched in the American legal system? One reason, of course, is that its flaws remain largely unrecognized. Another is that, like many other incorrect beliefs, it satisfies a pressing societal need. Workers' compensation programs demand a standardized method of awarding compensation to persons potentially injured by an exposure. Although stricter standards are needed in tort cases, there is nonetheless a need for some standard.

In this section, we argue that the worst consequence of the $PC = RF$ formula stems from its translation into a dichotomous decision rule that misinterprets a relative risk greater than 2 to mean "more probable than not." We further argue that, once the distinction between the probability of causation and rate fraction is recognized, the probability of causation may provide a poor basis for equitable compensation schemes.

A. The More-Probable-Than-Not Rule

A common scheme based on $PC = RF$ is the "all or nothing" rule mentioned earlier: If $IR > 2$ so that $RF > \frac{1}{2}$, the jury will fully compensate the plaintiff because $PC > \frac{1}{2}$, and thus it is "more probable than not" that exposure caused the plaintiff's disease. Conversely, if $IR < 2$ so that $RF < \frac{1}{2}$, the jury will deny compensation or decide for the defense on the grounds that $PC < \frac{1}{2}$, and thus it

---

[55] *See* Marvin Goldman, *Cancer Risk of Low-level Exposure*, 271 SCIENCE 1821, 1821 (1996).

[56] *See* Robins & Greenland, *Probability of Causation*, *supra* note 2, at 1130–31.

is "more probable than not" that exposure had no role in the etiology of the plaintiff's disease.[57]

The irrationality of this compensation scheme follows from the fact that there is no basis for inferring PC < ½ from epidemiologic data in any of the cases at issue. Even if the probability of causation were *known*, however, the scheme would remain unjust because it would never compensate anyone harmed by an exposure that caused fewer than half of the cases. This scheme shields from liability the party responsible for a harmful exposure or product, as long as the number of cases caused by the exposure does not exceed the number of cases that would have occurred among the exposed had they not been exposed.[58]

For example, if it were established that PC = RF for an environmental contaminant, an industrial facility could feel free to release the contaminant into its environs, as long as it carefully monitored its releases to insure that the surrounding community did not receive a dose capable of doubling the rate of any disease. Under the "all or nothing rule," a release known to be capable of increasing the childhood leukemia rate by 90% would not subject the facility to any liability, even if the release caused a local epidemic. In its defense, the facility could simply point out that its action did not double the childhood leukemia rate, so no case was "more probably than not" caused by its action. A similar rationale could be used by employers to allow workplace exposures to known hazards. For example, an employer could dismiss workers from jobs with hazardous exposures before workers accumulated an exposure dose known to elevate rates by twofold or more. The only current safeguards against such abuses are that courts may consider intent and ignore PC estimates when deciding cases.

Conversely, the "all or nothing" rule can excessively reward plaintiffs. For example, suppose it were established that PC = RF = 51% for some class of plaintiffs. All of these plaintiffs could receive full compensation even though nearly half suffered no damages from the exposure. Consequently, a defendant could be held liable for up to twice the damages it actually caused.

B. Partial Compensation Schemes

To avoid the problems of the "all or nothing" rule, other schemes have been developed to provide partial compensation for PC below one half.[59] The simplest scheme compensates plaintiffs in direct proportion to their PC. For example, suppose a given leukemia case would receive $1 million from a defendant *if* it were established with certainty that the defendant's action caused the plaintiff's disease. Instead, if the probability that the defendant's action caused the disease was known to be 0.30, the case would receive 30% of full compensation, or $300,000.

---

[57] Hatch, *supra* note 20

[58] *See, e.g.*, STEVEN SHAVELL, ECONOMIC ANALYSIS OF ACCIDENT LAW 116–17 (1987) (describing advantages and disadvantages of recovery in proportion to the probability of causation)

[59] *See* Armstrong & Theriault, *supra* note 4, Wakeford et al., *supra* note 4

While arguably more just than "all or nothing" schemes,[60] partial payment schemes based on PC still falter because the probability of causation is not estimable without strong and controversial biologic assumptions. Even if the disease biology is known, there is an aspect of these schemes that is arguably unfair to defendants. As described earlier, the probability of causation can be 1 for every case even if the rate fraction IR is close to 0. If research established that PC = 1, a defendant could be liable to fully compensate every single case of a disease, regardless of how small the impact of the defendant's actions is on the disease rate. Although such a judgment might in practice be precluded by consideration of *de minimis* harm, the possibility reveals a fundamental weakness in compensation proportional to PC.

C. Schemes Based on Rate Fractions

Recognizing the profound problems associated with PC schemes, some commentators have proposed a proportional compensation scheme with the rate fraction replacing the probability of causation in the compensation formula.[61] In the above example, the plaintiff would receive $300,000 if it were established that the rate fraction was 0.30. Without the PC misinterpretation of RF, there is no basis for claiming that a rate ratio above 2 corresponds to the "more probable than not" judicial standard, and the "all or nothing" rule is more easily seen as arbitrary. The rate fraction simply becomes the share of full compensation that the plaintiff receives or his "assigned share."[62]

Compensation schemes based on the rate fraction are severely flawed. Even if the disease is rare, the exposure is never preventive, and there is no confounding or modification of the rate ratio, the value of the causal rate ratio, and hence, the causal rate fraction, can increase as the degree of stratification by predictors of disease increases.[63] In particular, although it will not increase under an IOB model, it must increase under an accelerated-incidence model.[64] For example, it may be that the causal rate ratio for women age 45 is 1.5, but upon stratification by a given genotype, this ratio becomes 3 at each genotype level. Upon

---

[60] For discussions of the desirability of proportional compensation in various situations, see David Kaye, *The Limits of the Preponderance of the Evidence Standard: Justifiably Naked Statistical Evidence and Multiple Causation*, 1982 AM. BAR FOUNDATION RES. J. [LAW & SOCIAL INQUIRY] 487, *reprinted in* EVIDENCE AND PROOF (William Twining & Alex Stein eds., 1992); Shavell, *supra* note 58, at 116–17.

[61] *See* Lagakos & Mosteller, *supra* note 1, at 346–47.

[62] *See id.*; Cox, Jr., *Statistical Issues*, *supra* note 1, at 71–73.

[63] *See generally* Robins & Greenland, *Estimability and Estimation*, *supra* note 2, at 855–58; Robins & Greenland, *Probability of Causation*, *supra* note 2, at 1134–38; COX & OAKES, *supra* note 44; Sander Greenland, *Absence of Confounding Does Not Correspond to Collapsibility of the Rate Ratio or Rate Difference*, 7 EPIDEMIOLOGY 498 (1996) [hereinafter Greenland, *Absence of Confounding*]; Sander Greenland et al., *Confounding and Collapsibility in Causal Inference*, 14 STAT. SCIENCE 29-46 (1999) [hereinafter Greenland et al., *Confounding and Collapsibility*].

[64] *See generally* Robins & Greenland, *Estimability and Estimation*, *supra* note 2; Robins & Greenland, *Probability of Causation*, *supra* note 2.

stratification on a further genotype, the ratio becomes 5, and so on. This phenomenon of noncollapsibility of the causal rate ratio[65] precludes establishing stable compensation schemes based on the rate fraction.[66] Schemes with payment proportional to years of life lost do not suffer from this problem.[67]

D. Schemes Based on Years of Life Lost

Unlike the probability of causation, the rate fraction varies directly with the disease-free years of life lost due to exposure. Unfortunately, this relation will not be one of direct *proportionality*, except under implausible mathematical restrictions. Such proportionality is arguably desirable when damages are proportional to years of life lost.

One might use several compensation schemes to approximate such proportionality. One approach is to make compensation directly proportional to an estimate of the age-specific expected years of life lost under a biologic model. Unfortunately, as with probability of causation, the *age-specific* years of life lost due to exposure cannot be estimated from epidemiologic data without restrictive biologic assumptions.[68]

Despite their limitations, compensation schemes based on years of life lost help ensure that all cases suffering loss from exposure receive some compensation. They also limit defendants' liability to the total years of life lost due to exposure in the whole population, because errors caused by using an incorrect biologic model will tend to average out across age-specific estimates of years of life lost.[69] In contrast, such model errors will *not* average out across age-specific estimates of the probability of causation except under very special circumstances.[70] For example, if IR = 1.5 and PC = 1 at all ages and an IOB model is used to estimate PC, the expected PC estimate from unbiased data will be only 33% at all ages.[71]

V. UNCERTAINTY IN ESTIMATES

Assume now that all parties understand the basic conceptual problem and abandon the probability of causation in favor of an epidemiologically estimable quantity, such as the expected years of life lost. We then must confront the methodologic problems that contribute to our uncertainty about the chosen quantity. These problems are widely recognized, but often are not quantified realistically. Usually, the sole quantification of uncertainty in a study estimate is based on assuming that the data arose from random sampling of a target

---

[65] *See* Greenland, *Absence of Confounding*, *supra* note 63

[66] *See* Robins & Greenland, *Probability of Causation*, *supra* note 2

[67] *See generally* Robins & Greenland, *Hazardous Exposure*, *supra* note 2

[68] *See id.* at 80–82

[69] *See id.* at 79–93

[70] *See id.*

[71] For a more detailed discussion of years of life lost and compensation schemes, see *id.*

population in which exposure was randomized within levels of adjustment variables.[72] Even the most elementary sensitivity analyses reveal that the resulting statistics seriously understate the uncertainty one should have about observational estimates of effects.[73] Meta-analytic summary estimates are no better in this regard—they simply add an unrealistic assumption that the effects estimated by studies are homogeneous for fixed-effects summaries or constitute a random sample of effects from some abstract "population of effects" for random-effects summaries, and further understate uncertainty.[74]

The preceding cautions apply to *any* use of epidemiologic data to estimate causal effects, regardless of whether incidence time is important. Estimation of effects and quantification of uncertainty based on observational data are topics surrounded by heated philosophical disputes and formidable technical problems.[75] Hence, there is no complete and generally accepted methodology for causal inference from observational data. There are even disputes about inference from randomized trials, especially those that suffer from severe nonrepresentativeness or noncompliance.[76] Even if there were such an accepted methodology, however, another difficult problem would remain – that of how a court or agency should account for uncertainty in the estimates.[77]

---

[72] *See* Sander Greenland, *Randomization, Statistics, and Causal Inference*, 1 EPIDEMIOLOGY 421 (1990).

[73] *See generally id.*; David A. Freedman, *As Others See Us: A Case Study in Path Analysis*, 12 J. EDUC. STAT. 101, 122–25 (1987); David Draper, *Assessment and Propagation of Model Uncertainty*, 57 J. ROY. STAT. SOC. B 45, 47–49 (1995); PAUL R. ROSENBAUM, OBSERVATIONAL STUDIES 87–133 (1995); Sander Greenland, *Basic Methods of Sensitivity Analysis*, 25 INT. J. EPIDEMIOLOGY 1107–16 (1996); John B. Copas & H. G. Li, *Inference for Non-Random Samples*, 59 J. ROY. STAT. SOC. B 55, 55–57 (1997); James M. Robins et al., *Sensitivity Analysis for Selection Bias in Missing Data and Causal Inference Models*, *in* STATISTICAL MODELS IN EPIDEMIOLOGY, THE ENVIRONMENT, AND CLINICAL TRIALS 1–92 (M. Elizabeth Halloran & Donald Berry eds., 2000); John B. Copas, *What Works? Selectivity Models and Meta-analysis*, 162 J. ROY. STAT. SOC. A 95, 97, 107–08 (1999).

[74] *See* Copas, *supra* note 73, at 95–97; Sander Greenland, *A Critical Look at Some Popular Meta-Analytic Methods*, 140 AM. J. EPIDEMIOLOGY 290, 290–94 (1994).

[75] *See* Greenland, *supra* note 72; Freedman, *supra* note 73, at 101–04; Draper, *supra* note 73, at 65–67; Robins et al., *supra* note 73; Copas, *supra* note 73, at 95–98; Joseph Gastwirth et al., *How a Court Accepted an Impossible Explanation*, 48 AM. STAT. 313, 314–15 (1994); Charles Poole & Sander Greenland, *How a Court Accepted a Possible Explanation: A Comment*, 51 AM. STAT. 112 (1997); Joseph L. Gastwirth et al., *Unquestionably Impossible - Reply*, 51 AM. STAT. 115 (1997); Jim Mintz & Wilfrid Dixon, *Objection Overruled*, 51 AM. STAT. 117 (1997); Paul Humphreys & David A. Freedman, *The Grand Leap*, 47 BRIT. J. PHIL. SCI. 113 (1996); Kevin B. Korb & Chris S. Wallace, *In Search of the Philosopher's Stone*, 48 BRIT. J. PHIL. SCI. 543 (1997); Peter Spirtes et al., *Reply to Humphreys and Freedman's Review of "Causation, Prediction and Search,"* 48 BRIT. J. PHIL. SCI. 555 (1997); Mark Parascandola, *Chances, Individuals and Toxic Torts*, 14 J. APPL. PHIL. 147, 148–51 (1997); COLIN HOWSON & PETER URBACH, SCIENTIFIC REASONING: THE BAYESIAN APPROACH (2d ed. 1993); Sander Greenland, *Probability Logic and Probabilistic Induction*, 9 EPIDEMIOLOGY 322–32 (1998); JUDEA PEARL, CAUSALITY: MODELS, REASONING, AND INFERENCE (2000); *see also* Copas & Li, *supra* note 73.

[76] *See* HOWSON & URBACH, *supra* note 75.

[77] *See* Donald E. Jose, *The Limitations of Probability of Causation*, 29 NUCLEAR NEWS 39, 41 (1986); CARL F. CRANOR, REGULATING TOXIC SUBSTANCES: A PHILOSOPHY OF SCIENCE AND THE

* * *

Over the past two decades, the concept of "probability of causation" and equating this concept with attributable-fraction measures have become entrenched in law, policy, and the health sciences. This entrenchment occurred despite early and extensive explanations of why the equation is flawed. These explanations have mostly been ignored, improperly cited, or misunderstood. The perpetuation of this conceptual error has promoted logically unsound uses of epidemiologic data to make compensation decisions.

LAW 44–48 (1993) (discussing public policy issues of courts and agencies using epidemiological data)

APPENDIX: BOUNDS FOR THE PROBABILITY OF CAUSATION AND THEIR RELATION TO RATE FRACTIONS

For simplicity, we assume that disease is nonrecurrent, exposure is never preventive, and there are no competing risks.[78] With these assumptions, the average disease risk up to time t if exposure had not occurred, $R_0(t)$, cannot exceed the average risk given that exposure did occur, $R_1(t)$.

A. Lower Bounds

Suppose the biologic mechanism of exposure action minimizes the fraction of cases affected by exposure under the risk distributions $R_1(t)$ and $R_0(t)$. If $R_1(t)$ and $R_0(t)$ have probability densities $f_1(t) = dR_1(t)/dt$ and $f_0(t) = dR_0(t)/dt$, the minimum is attained if the differences $f_1(t) - f_0(t)$ are solely due to exposure shifting case occurrences from times t with $f_1(t) < f_0(t)$ to earlier times u with $f_1(u) > f_0(u)$. Under such a mechanism, the chance of getting disease at time t *because* of exposure, given that exposure occurred, is $f_1(t) - f_0(t)$ if $f_1(t) > f_0(t)$, and 0 otherwise. We then obtain the minimum probability of causation among exposed cases occurring at time t, m(t), by conditioning on (dividing by) the probability of disease at t among the exposed, so that

$$m(t) = [f_1(t)) - f_0(t)]/f_1(t) = 1 - f_0(t)/f_1(t) \quad \text{if} \quad f_1(t) > f_0(t),$$

and m(t) is zero otherwise. Furthermore, the minimum fraction of exposed cases affected by exposure is the sum (integral) of the minimum time-specific probabilities,

$$\int_G [f_1(u) - f_0(u)]du = \int_G m(u)f_0(u)du,$$

where G is the set of all times t with $f_1(t) > f_0(t)$.

Now, let $S_1(t) \equiv 1 - R_1(t)$ and $S_0(t) \equiv 1 - R_0(t)$ be the survival distributions with and without exposure, and let $I_1(t)$ and $I_0(t)$ be the corresponding time-specific incidence rates, given by[79]

$$I_i(t) = f_i(t)/S_i(t), \; i = 0, 1.$$

Then $$f_i(t) = I_i(t)S_i(t) \quad \text{and so}$$

$$m(t) = 1 - \frac{I_0(t)S_0(t)}{I_1(t)S_1(t)}.$$

---

[78] Without these assumptions the discrepancies between RF and PC can be even larger than described here, as explained in the text and Robins & Greenland, *Estimability and Estimation*, *supra* note 2, at 847–48; Robins & Greenland, *Probability of Causation*, *supra* note 2, at 1131.

[79] *See* KALBFLEISCH & PRENTICE, *supra* note 43, COX & OATES, *supra* note 44.

If, as is often true, the exposure effect on survivorship to t is small, so that $S_0(t)/S_1(t) \approx 1$ (often less accurately described as a "rare-disease" assumption), then

$$m(t) = 1 - I_0(t)/I_1(t) = [I_1(t) - I_0(t)]/I_1(t) ,$$

which is the time-specific rate fraction.[80] This relation is why we claim the rate fraction, while not a reasonable estimate of PC, can sometimes be a reasonable *lower bound* for PC. Without the given assumptions, however, RF may be a poor lower bound, as well as a poor estimate of PC.[81]

B. Upper Bounds

Any pair of risk functions $R_1(t)$ and $R_0(t)$ with $R_1(t) > R_0(t)$ at all t are compatible with an accelerated-life[82] mechanism of exposure effect, in which a person who would have gotten disease at t_0 when exposed instead gets disease at $t_1 = S_1^{-1}[S_0(t_0)]$ when unexposed.[83] With such a mechanism, $t_1 < t_0$ for all $t_0 > t_e$, where

$$t_e = \inf\{u: R_1(u) > R_0(u)\}$$

is the earliest time at which exposure effects appear, and so the probability of causation will be 1 for exposed cases occurring after t_e, and 0 for cases occurring before t_e. In general, t_e will be unknown. But if t_u is an upper bound for t_e, then PC will also be 1 for exposed cases occurring after t_u. For cases occurring before t_u, t_e is identifiable if $R_1(t)$ and $R_0(t)$ are identifiable. In theory, one could develop a posterior distribution $F_e(t_e)$ for t_e from epidemiologic data. Under the accelerated-life model, $F_e(t)$ would be the corresponding posterior probability of causation among cases occurring at time t. $F_e(t)$ is the posterior probability that these cases occurred after the unknown time t_e.

---

[80] *See* Robins & Greenland, *Estimability and Estimation*, *supra* note 2, at 848–54; Robins & Greenland, *Probability of Causation*, *supra* note 2, at 1129–31.

[81] *See id.*

[82] *See* COX & OATES, *supra* note 44, at 75–76.

[83] *See* Robins & Greenland, *Estimability and Estimation*, *supra* note 2; Robins & Greenland, *Probability of Causation*, *supra* note 2.