UNC SCHOOL OF LAW

NORTH CAROLINA LAW REVIEW

Volume 71 | Number 1 Article 14

11-1-1992

Causal Inference in Epidemiology: Implications for Toxic Tort Litigation

Melissa Moore Thompson

COMMENT

Causal Inference in Epidemiology: Implications for Toxic Tort Litigation

> *Alice had been looking over his shoulder with some curiosity. "What a funny watch!" she remarked. "It tells the day of the month, and doesn't tell what o'clock it is!"*
>
> *"Why should it?" muttered the Hatter. "Does your watch tell you what year it is?"*
>
> *"Of course not," Alice replied very readily; "but that's because it stays the same year for such a long time together."*
>
> *"Which is just the case with mine," said the Hatter.*
>
> *Alice felt dreadfully puzzled. The Hatter's remark seemed to her to have no sort of meaning in it, and yet it was certainly English. "I don't quite understand you," she said, as politely as she could.*[1]

It seems that the "Mad Hatter" was more a reflection of reality than a caricature of fiction. Bernardino Ramazzini, a sixteenth century Italian physician, observed that certain diseases tend to cluster within particular occupational groups.[2] Ramazzini implicated mercury, historically used by hatmakers, as a cause of nervous system disorders—hence the Hatter's madness.[3] Observational studies like Ramazzini's were the progenitors of the science of epidemiology and portents of the harm engendered by certain occupational and environmental exposures.

Epidemiology is a discipline concerned with the causes and occurrence of disease in populations.[4] One of the main goals of epidemiology is disease prevention. The aim is "[t]o provide the basis for developing

---

1. LEWIS CARROLL, ALICE'S ADVENTURES IN WONDERLAND 60 (Holt, Rinehart and Winston 1985) (1865).

2. HARVEY CHECKOWAY ET AL., RESEARCH METHODS IN OCCUPATIONAL EPIDEMIOLOGY 4 (1989) (citing BERNARDINO RAMAZZINI, DE MORBIS ARTIFICUM (DISEASES OF WORKERS) (W.C. Wright trans., Hafner 1964) (1700)).

3. *See id.* In hatters, mercury caused "toxic mental changes, called madness, hence the phrase 'mad as a hatter.' " STANLEY L. ROBBINS ET AL., PATHOLOGIC BASIS OF DISEASE 452 (3d ed. 1984).

4. *See* ABRAHAM M. LILIENFELD & DAVID E. LILIENFELD, FOUNDATIONS OF EPIDEMIOLOGY 3 (2d ed. 1980) ("Epidemiology is concerned with the patterns of disease occurrence in human populations and of the factors that influence these patterns."); KENNETH J. ROTHMAN, MODERN EPIDEMIOLOGY 23 (1986) ("The fundamental task in epidemiologic research is thus to quantify the occurrence of illness. The goal is to evaluate hypotheses about the causation of illness and its sequelae and to relate disease occurrence to characteristics of people and their environment."); Reuel A. Stallones, *To Advance Epidemiology*, 1 ANN. REV. PUB.

and evaluating preventive procedures and public health practices."[5] As a science, epidemiology is in the early phases of development; many of its basic tenets and methods have evolved only in the last thirty years.[6] Epidemiological studies have increased public awareness of risk from exposure to harmful substances, and in a rather circular manner this heightened concern has provided an impetus for the continued development of epidemiology and its use in the courtroom.

Epidemiological evidence can be crucial to the outcome of toxic tort litigation. Attempts to apply epidemiology in the courtroom unfortunately have resulted in confusion and inconsistent holdings.[7] One commentator, Michael Dore, notes that "[c]ourts have treated epidemiological evidence inconsistently, largely as a result of their failure to recognize its limitations in proving causation."[8] To eliminate this confusion, judges and lawyers must develop a basic understanding of epidemiology. For, as Dr. Harold Ginzburg writes, "in litigation, the court is the ultimate interpreter of epidemiologic data."[9]

This Comment provides an introduction to the logic of causal inference in epidemiology. It emphasizes that epidemiologists consider more than mere statistical association when determining causation, including factors such as the degree of consistency among studies and the strength of the temporal relationship between cause and effect. This Comment urges that the courts employ a similar analysis. It focuses primarily on single-defendant situations[10] under a claim of negligence, in which exposure to one substance[11] for which the defendant is responsible allegedly

---

HEALTH 69, 71 (1980) (The "territory of epidemiology . . . is presently generally agreed to be the concern for the occurrence of disease (and, by inference, health) in groups of people.").

5. LILIENFELD & LILIENFELD, *supra* note 4, at 4.

6. ROTHMAN, *supra* note 4, at 2-3. *See generally* Stallones, *supra* note 4, at 69-70 (discussing the development of epidemiology in this century). A current and widely cited textbook of epidemiology, *Modern Epidemiology*, by Dr. Kenneth J. Rothman, will form the basis for much of the discussion in this Comment.

7. Michael Dore, *A Commentary on the Use of Epidemiological Evidence in Demonstrating Cause-in-Fact*, 7 HARV. ENVTL. L. REV. 429, 435 (1983); *see also* Troyen A. Brennan, *Untangling Causation Issues in Law and Medicine: Hazardous Substance Litigation*, 107 ANNALS INTERNAL MED. 741, 746 (1987) (stating that "[j]udges and juries are confused by epidemiologic and probabilistic evidence of causation").

8. Dore, *supra* note 7, at 440.

9. Harold M. Ginzburg, *Use and Misuse of Epidemiologic Data in the Courtroom: Defining the Limits of Inferential and Particularistic Evidence in Mass Tort Litigation*, 12 AM. J.L. & MED. 423, 428 (1986).

10. The presence of multiple defendants complicates the analysis of causation. Cases involving multiple defendants are beyond the scope of this Comment. For a discussion of multiple defendants, see Andrew G. Celli, Jr., *Toward a Risk Contribution Approach to Tortfeasor Identification and Multiple Causation Cases*, 65 N.Y.U. L. REV. 635, 635-39 (1990).

11. Although not addressed in this Comment, methods that analyze multiple causes of

has caused a particular disease in the plaintiff. It notes that courts have applied epidemiology in this context to address two related issues.[12] The first issue is whether exposure to a particular substance is capable of causing the disease. This is a population-level analysis that can be used to determine negligence—whether the defendant has subjected the plaintiff to an unreasonable risk of harm. The second issue is an individual-level analysis—whether the exposure has caused the plaintiff's harm. This Comment maintains that epidemiological evidence is limited in its legal applications.[13] It concludes that the results generated from epidemiological studies are particularly relevant to the determination of general population-level causation, but that when used to determine causation at the level of the individual, their relevance decreases and the potential for prejudice increases.[14] This Comment recommends two guidelines for courts to consider when evaluating the admissibility of evidence based on epidemiology to prove legal cause: (1) whether the plaintiff can prove facts from which an inference of population-level causation could be made, examining factors in addition to mere statistical association; and (2) whether the plaintiff can prove facts to support an adequate degree of external validity, *i.e.*, that the epidemiological study population was similar enough to the plaintiff to make extrapolation reliable.[15] This Comment proposes the following guideline to be used when evaluating the sufficiency of epidemiological evidence: the plaintiff has proved the facts in (1) and (2) above, as well as additional facts that, when taken together, make it more likely than not that exposure to a particular substance caused the plaintiff's disease.[16]

I. CAUSAL INFERENCE IN EPIDEMIOLOGY

Determining causation in epidemiology is a complex process. A common misconception seems to be that epidemiologists rely solely on statistical measures of association, such as the risk ratio, to determine

---

disease are currently being debated in epidemiology. *See, e.g.*, ROTHMAN, *supra* note 4, at 311-26 (acknowledging the current debate and describing a conceptual model based on the notion of sufficient cause). Epidemiologists often use multiple regression analysis, a class of statistical methods designed to study multiple causes of disease. DAVID G. KLEINBAUM ET AL., APPLIED REGRESSION ANALYSIS AND OTHER MULTIVARIABLE METHODS 1-6 (2d ed. 1988) (discussing biostatistical concepts and examples of research involving regression analysis). *But cf.* ROTHMAN, *supra* note 4, at 4-5 (objecting to the notion that "[s]tatistical methodology in multivariate modeling has often been transferred wholesale to epidemiology without giving sufficient thought to the underlying epidemiologic concepts").

12. *See infra* notes 150-51 and accompanying text.
13. *See infra* notes 183-98 and accompanying text.
14. *See infra* notes 183-84 and accompanying text.
15. *See infra* notes 215-30 and accompanying text.
16. *See infra* notes 231-35 and accompanying text.

causation.[17] Although epidemiologists use statistics, the logic of causation extends well beyond the numbers alone. Statistical association is but one of several factors considered in the analysis.[18]

A. Measures of the Association between Cause and Effect

Epidemiologists derive mathematical measures of association between a putative cause and its effect from population-based data. Measures of association serve two basic functions. First, they are used to determine the strength of the association, one of several guidelines used in a determination of causation.[19] In addition, they are used to gauge risk—the likelihood that exposure to a particular substance will cause disease. Because much of the logic of causation in epidemiology depends upon these measures, legal professionals using epidemiology should develop a conceptual understanding of them. Courts often cite two basic measures, the risk ratio and the attributable risk proportion. The following sections present a brief description of these measures and several illustrations of their use in the courtroom.

1. The Risk Ratio

In epidemiology, risk is quantitative—a number derived using a mathematical formula. Because of factors such as study bias and error due to chance,[20] the resulting number represents, at best, a close approximation of the actual risk. The risk ratio, also known as relative risk,[21] may be defined as the risk of disease in a population segment exposed to a particular substance, divided by the risk of disease in the rest of the population.[22] The risk ratio represents how much more likely an exposed

17. *See generally* Dore, *supra* note 7, at 431 (defining epidemiology as "the statistical study of disease in human populations"); Jeffrey Trauberman, *Statutory Reform of "Toxic Torts": Relieving Legal, Scientific, and Economic Burdens on the Chemical Victim*, 7 HARV. ENVTL. L. REV. 177, 198 (1983) ("[Epidemiological] studies address issues of causality in terms of statistical probabilities.").

18. Epidemiological causation determinations involve the balancing of many factors. *See infra* notes 103-48 and accompanying text.

19. *See infra* notes 108-10 and accompanying text.

20. For a discussion of these factors, see *infra* notes 46-84 and accompanying text.

21. ROTHMAN, *supra* note 4, at 37. Other terms for the risk ratio are the relative rate, rate ratio, and incidence rate ratio. *Id.*

22. The risk ratio is the measure of association commonly derived from a cohort study. ROBERT H. FLETCHER ET AL., CLINICAL EPIDEMIOLOGY 195-97 (2d ed. 1988). In a cohort study, the researcher defines a population (such as all workers at a particular type of manufacturing plant) and then divides the population into an exposed group (those exposed to the substance in question) and an unexposed group (those not exposed to the substance). The number of people in each group who contract the disease in question is recorded. *Id.* at 195. The risk ratio is derived from this data. Two other epidemiological study types deserve men-

person is to contract the disease than an unexposed person.[23] A comparison is made between the proportion of persons with the disease in a group of people exposed versus the proportion of persons with the disease in the group not exposed. If the substance is a cause of the disease, the exposed group should have a larger proportion of people with the disease. The risk ratio may be calculated as follows:[24]

$$\frac{\text{risk of disease in the exposed group}}{\text{risk of disease in the unexposed group}}$$

or, equivalently,

$$\frac{\text{number of exposed persons developing disease (over time)}}{\text{total number of exposed people}}$$

divided by

$$\frac{\text{number of unexposed persons developing disease (over time)}}{\text{total number of unexposed people}}$$

This establishes a ratio of rates, representing the average likelihood of contracting the disease after being exposed to a particular substance relative to the average likelihood of contracting the disease if not exposed.

When the risk ratio equals one, there are equal proportions of persons with the disease in both the exposed group and the unexposed group. This indicates that exposed persons do not contract the disease more often than unexposed persons. In other words, there is no observed association between the exposure and disease. When the risk ratio is greater than one, exposed persons contract the disease more often than those not exposed. In such a situation, an association exists between exposure to the substance and the disease.[25] When the risk ratio equals three, for example, exposed persons are three times more likely to develop the disease than unexposed persons. When the risk ratio equals

---

tion here: the case-control study and the prevalence study. These two study types produce an odds ratio, which closely approximates the risk ratio under certain conditions. *Id.* at 196-97. For a discussion of study design in epidemiology, see *id.* at 79-84, 96-101, 192-94, and 2 MICHAEL DORE, LAW OF TOXIC TORTS § 25.05 (1992). The odds ratio closely approximates the risk ratio only when the disease is rare in a population—when less than 10% of the population has contracted the disease by the end of the study period. Sander Greenland & Duncan C. Thomas, *On the Need for the Rare Disease Assumption in Case-Control Studies*, 116 AM. J. EPIDEMIOLOGY 547, 547-53 (1982).

23. To extrapolate accurately from population data to the level of the individual, one would need to assume that *all* persons were nearly identical. In reality, this assumption cannot be made. This raises the broader question of whether population data may ever be extrapolated to the individual. *See infra* notes 41-45.

24. *See generally* ROTHMAN, *supra* note 4, at 36-37 (calculating the risk ratio); CHECKOWAY et al., *supra* note 2, at 99 (defining risk ratio).

25. FLETCHER et al., *supra* note 22, at 197.

eight, exposed persons are eight times more likely to develop the disease, and so forth. When the risk ratio is less than one, exposure to the substance exhibits a negative association with the disease and may actually protect against the disease.[26]

A large risk ratio signifies a strong association, which is highly indicative, although not determinative, of a causal relationship.[27] Conversely, a nonexistent or extremely weak association, evidenced by a risk ratio close to one, suggests, but does not prove, the absence of causation. Some courts have viewed risk ratios greater than one as capable of proving causation. For example, the court in *Oxendine v. Merrell Dow Pharmaceuticals, Inc.*[28] viewed testimony based on both an epidemiological study with a risk ratio between 1.3 and 1.8 and other data sufficient to present to the jury a question on the issue of causation.[29] Although any risk ratio greater than one logically would support the notion of a causal relationship, it is very important to understand that epidemiologists consider any risk ratio less than *three* to indicate a weak association.[30] Risk ratios less than three can be generated entirely by factors such as study bias and lack of precision.[31]

2. The Attributable Risk Proportion

The attributable risk proportion (ARP)[32] represents "the proportion of exposed [persons with the disease in the study] for whom the disease is

---

26. *Id.*

27. In addition to its utility in the determination of epidemiological causation, the risk ratio also may be used in risk assessment. In this context, the risk ratio provides a rough measure of the risk that exposure to a particular substance could result in a given disease. Risk is an important concept in health care; it indicates whether physicians should recommend certain treatments (such as surgery versus chemotherapy for certain cancers) or preventive measures (such as low salt diets for persons with high blood pressure). Risk also is important in the legal context of evaluating what reasonable care requires. *See infra* note 151 and accompanying text.

28. 506 A.2d 1100 (D.C. 1986), *cert. denied*, 493 U.S. 1074 (1990).

29. *Id.* at 1108, 1110.

30. Ernst L. Wynder, *Guidelines to the Epidemiology of Weak Associations*, 16 PREVENTIVE MED. 139, 139 (1987).

31. For a discussion of these factors, see *infra* notes 46-84 and accompanying text.

32. The ARP is also termed the attributable proportion, the etiologic fraction in the exposed, and the attributable risk percent. ROTHMAN, *supra* note 4, at 38 (citing Philip Cole & Brian MacMahon, *Attributable Risk Percent in Case-Control Studies*, 25 BRIT. J. PREVENTIVE & SOC. MED. 242, 242-44 (1971); Olli S. Miettinen, *Proportion of Disease Caused or Prevented by a Given Exposure, Trait or Intervention*, 99 AM. J. EPIDEMIOLOGY 325, 325-32 (1974)). The attributable risk proportion also has been called the "probability of causation." Junius C. McElveen, Jr. & Pamela S. Eddy, *Cancer and Toxic Substances: The Problem of Causation and the Use of Epidemiology*, 33 CLEV. ST. L. REV. 29, 43 (1984-85). Probability of causation is unfortunate terminology because the attributable risk proportion, without more, does not prove causation and generally may not be extrapolated to the individual.

attributable to the exposure."[33] Assuming a cause-and-effect relationship exists at the population level, the ARP also represents the probability of randomly selecting, from the exposed, diseased group in the study, a person whose disease was caused by the exposure. The ARP is simply a mathematical transformation using the risk ratio. It is employed after causation has been established at the population level and is meaningful only after a determination of population-level causation first has been made. The ARP is calculated as follows:[34]

$$\text{Attributable Risk Proportion} = \frac{\text{risk ratio} - 1}{\text{risk ratio}}$$

As the following calculation shows, a risk ratio of two will yield an ARP of .50:

$$\text{ARP} = \frac{2 - 1}{2} = \frac{1}{2} = .50$$

An ARP of .50 means that half of the disease in the exposed study population is likely to have been caused by the exposure and half is not likely to have been caused by the exposure. This Comment will represent the ARP as a percentage rather than a proportion. For example, an ARP of .50 will be represented as 50%. As such, ARP will represent the attributable risk *percentage*, rather than the attributable risk *proportion*. In this context, an ARP of 50% means there is a 50% chance that the disease of a person selected at random from the exposed group *in a particular study* was due to the exposure. Notably, it does not mean that there is a 50% chance that the disease of a given person *outside the study* was caused by the exposure.

Before applying the ARP, causation should be determined at the population level.[35] Otherwise, the ARP merely represents the probability that a person's disease is associated with, but not necessarily caused by, the exposure.[36] Accordingly, until population-level causation has been determined, the concept of an ARP has no meaning. A court that uses the risk ratio or ARP to assess individual causation prior to a determination of population-level causation violates this premise. Table I presents risk ratio and ARP estimates and indicates the approximate strength of association at each level.

---

33. ROTHMAN, *supra* note 4, at 38.

34. *Id.*

35. *See infra* notes 112-48 and accompanying text.

36. An association is *not* tantamount to causation, but rather is one of several factors epidemiologists consider in an overall causation assessment. *See infra* notes 103-48 and accompanying text.

TABLE I. AN APPROXIMATION OF THE STRENGTH OF ASSOCIATION EXHIBITED BY VARIOUS RISK RATIOS.

Relative Risk	(ARP)	Approximate Strength of the Association	Examples
1.1 to 3	(9%-67%)	Weak or nonexistent	Bendectin and birth defects[37]
3 to 8	(67%-87%)	Moderate	Swine flu vaccine and Guillain-Barre syndrome[38]
8 to 16	(87%-94%)	Strong	Cigarette smoking and death due to lung cancer[39]
16 to 40+	(94%-97.5%)	Extremely Strong	High levels of radon and death due to lung cancer[40]

3. Use of the Risk Ratio and ARP by Courts

After causation is found at the population level, the ARP can be probative, but not determinative, of causation in an individual. As the following example illustrates, problems arise when risk estimates are applied to a particular individual. Epidemiological studies can predict the risk of death from riding bicycles on roadways. But can the value of the risk be applied with confidence to an individual cyclist in a particular situation? The answer is certainly no. Such a value merely averages risk and does not take into account potentially pertinent factors such as the number of hours ridden per week, safety equipment worn, level of training, or type of roadway. Of course, studies can be designed to address such factors. The more factors studied and the greater their similarity to an individual's cycling style, the more confident one should be in applying risk estimates to the individual. As a population-based measure, however, risk should be viewed as an estimation rather than a precise quantification for a particular individual.

Some courts have focused on ARPs greater than 50% or, equivalently, risk ratios greater than two, as the cut-off point for a "more likely than not" burden of proof for legal cause in an individual.[41] These

37. Patricia H. Shiono & Mark A. Klebanoff, *Bendectin and Human Congenital Malformations*, 40 TERATOLOGY 151, 151-55 (1989).

38. Alvarez v. United States, 495 F. Supp. 1188, 1203-05 (D. Colo. 1980).

39. FLETCHER et al., *supra* note 22, at 102 (estimated data from Richard Doll & Austin B. Hill, *Mortality in Relation to Smoking: Ten Years' Observations of British Doctors*, 1 BRIT. MED. J. 1399, 1402-03 (1964)).

40. ENVIRONMENTAL PROTECTION AGENCY, RADON REDUCTION TECHNIQUES FOR DETACHED HOUSES 3 (1986) (radon exposure level of 0.2 WL). There is no single risk ratio published for radon and lung cancer—the risk ratio varies with the level of radon exposure, from a risk ratio of three to above 75. *Id.*

41. *See, e.g.*, Manko v. United States, 636 F. Supp. 1419, 1437 (W.D. Mo. 1986) (stating that when the risk ratio is greater than two, exposure to a substance is "more likely than not" a

courts incorrectly assume this translates into a greater than 50% chance that the *plaintiff's* disease was caused by exposure to the substance. The use of the ARP and risk ratio in this manner is misguided. James Robins and Sander Greenland examined this practice and concluded that it could not be substantiated mathematically.[42] They stated, "the probability of [individual] causation is non-identifiable."[43] In other words, statistic-based epidemiological study results should not be applied directly to establish the likelihood of causation in an individual plaintiff. According to these authors, a population-based statistical measure represents a *range* of values, not a single value.[44] Using this reasoning, a risk ratio greater than two, corresponding to an ARP greater than 50%, should not be used without other evidence to satisfy a "more likely than not" standard for legal cause in an individual.

A risk ratio of two simply means that persons in a study who have been exposed to a particular substance are twice as likely to contract disease than persons not exposed.[45] Consider the following example:

> Assume a hypothetical population of 100,100 people, 100 of whom were exposed to a substance thought to cause a particular disease. This establishes two groups: an exposed group (100 people) and an unexposed group (100,000 people). The epidemiologist would count the number of people in each group who developed the disease over a particular period of time. A ratio would be set up using the risk ratio formula. Assume the following numbers: of the 100 exposed persons, two develop the disease, and of the 100,000 unexposed persons, 1,000 develop the disease. Applying the formula:

---

cause of the disease), *aff'd in part*, 830 F.2d 831 (8th Cir. 1987); Marder v. G.D. Searle & Co., 630 F. Supp. 1087, 1092 (D. Md. 1986) (indicating that a risk ratio of two is "an important showing for plaintiffs to make because it is the equivalent of the required legal burden of proof—a showing of causation by the preponderance of the evidence"), *aff'd sub nom.* Wheelahan v. G.D. Searle & Co., 814 F.2d 655 (4th Cir. 1987); Cook v. United States, 545 F. Supp. 306, 308 (N.D. Cal. 1982) (stating that risk ratio of two would "sustain[ ] plaintiff's burden of proof on causation"). For a discussion of statistical evidence used in *Manko* and *Marder*, see DAVID W. BARNES & JOHN M. CONLEY, STATISTICAL EVIDENCE IN LITIGATION § 10.1.1, at 87-89 (Supp. 1989).

42. James Robins & Sander Greenland, *The Probability of Causation Under a Stochastic Model for Individual Risk*, 45 BIOMETRICS 1125, 1126 (1989).

43. *Id.*

44. *Id.*

45. *See supra* note 24 and accompanying text.

$$\frac{\text{2 exposed persons with the disease}}{\text{100 exposed persons}}$$

divided by

$$\frac{\text{1000 unexposed persons with the disease}}{\text{100,000 unexposed persons}}$$

This results in:

$$\frac{.02}{.01} = \text{2, a risk ratio of "2"}$$

The mathematics support the notion that an *association* exists between exposure and disease, but *causation* decisions are not made based on mathematics alone. Although a risk ratio of two indicates that people are twice as likely to contract the disease when exposed, epidemiologists actually consider such numbers to be only weak support for causal inference. In the above example, two persons from a group of 100 exposed contracted the disease. Had there been no exposure at all, one of these persons would have been expected to contract the disease anyway. Intuitively, it is hard to look at such numbers and state that exposure to the substance can cause the disease. By extension, courts certainly should not use automatically a risk ratio of two as sufficient to prove legal cause.

4. Error as a Source of Uncertainty in the Risk Ratio

Error can seriously compromise an epidemiological study and impair its usefulness in the courtroom. If population-level analyses contain significant error, individual-level application to a plaintiff arguably would compound the error. In an attempt to sift out potential sources of error, epidemiologists address four general areas of concern: statistical significance, lack of precision, scientific bias, and external validity.

a. Statistical Significance

Statistical significance represents the likelihood that the results of an epidemiological study are due entirely to chance or random error.[46] The level of significance is represented usually by a "p-value" or, alternatively, by a "confidence interval."[47] Statistical significance is a function of study size and the amount of variance in the study population. A near consensus of biological scientists, including epidemiologists, recognizes statistical significance at a level that provides a 95% certainty that the results generated are not due to chance alone. In most instances, this is expressed as a p-value that is less than or equal to .05 or a 95% confi-

46. ROTHMAN, *supra* note 4, at 116.

47. *Id.* at 116, 119.

dence interval that does not include the number one.[48] Epidemiologists often report study results as a risk ratio, a chi-square test statistic, and a p-value, with or without a confidence interval.

A number of courts insist that epidemiological evidence be statistically significant if used to prove causation.[49] This probably reflects a suitable measure for acceptance of a particular study in the scientific community. However, statistical significance is not determinative of causation. It is important to understand that statistical significance merely reflects the likelihood that study results are due to chance or random error. Furthermore, most statistical methods rely on the assumption that the population being tested is a random sample.[50] Therefore, statistical significance does not address other types of error that tend to produce nonrandom samples, such as scientific bias.[51] For both of these reasons, a study could be statistically significant yet still yield an incorrect result due to scientific bias. Conversely, because statistical significance depends on study size, a small study could yield an unbiased, correct result yet suffer from a lack of statistical significance. Epidemiologists do not view single studies, even if statistically significant, as establishing a causal relationship. Similarly, they do not view small studies with results that are not statistically significant as definitive proof of the lack of causation.

b. Precision: Reduction of Error Due to Chance

Precision is a term that refers to "the extent to which repeated measurements of a relatively stable phenomenon fall closely to each other."[52] It reflects the reliability or reproducibility of a study.[53] In other words, if the study or portions of it were repeated, precision represents how close the results of latter studies would be to those of the earlier study. Classically, lack of precision has been thought to be due to the presence of chance or random error.[54] However, other factors, such as inconsistency in study measurements and the inherent variability of human subjects,

48. *Id.* at 117. For a discussion of these concepts, see *id.* at 115-25; RICHARD D. REMINGTON & M. ANTHONY SCHORK, STATISTICS WITH APPLICATIONS TO THE BIOLOGICAL AND HEALTH SCIENCES 166-75 (2d ed. 1985).

49. *E.g.*, Brock v. Merrell Dow Pharmaceuticals, Inc., 874 F.2d 307, 312-13 (5th Cir. 1989) (rejecting plaintiff's claims because lack of statistical significance in available epidemiological studies evidenced by a confidence interval that included the number one), *cert. denied*, 110 S. Ct. 1511 (1990).

50. For a discussion of random sampling in statistical analysis, see REMINGTON & SCHORK, *supra* note 48, at 71-77.

51. For a discussion of these types of errors, see *infra* notes 65-78 and accompanying text.

52. FLETCHER et al., *supra* note 22, at 23.

53. *Id.*

54. ROTHMAN, *supra* note 4, at 78.

may contribute to lack of precision in epidemiological studies.[55]

Lack of precision is of greatest consequence when the association between exposure to a substance and the disease is weak (risk ratio less than three).[56] In such a situation, lack of precision can defeat any attempt to establish population-level causation. As a result, the study may not be statistically significant, and it may appear that there is no association when an association in fact exists.

Enlarging the study size is the dominant vehicle for correcting a lack of precision.[57] The strength of the association that can be detected at a given level of statistical significance depends on the size of the study; a small study generally is able to detect only strong associations, but a large study can detect both strong and weak associations.[58] The litigation initiated by Vietnam War veterans exposed to the defoliant Agent Orange illustrates this concept.[59] In *In re Agent Orange Product Liability Litigation*, the epidemiologic evidence revealed weak associations that

---

55. Humans are a diverse species; each person has different disease determinants and susceptibilities because of random, environmental, or genetic factors. *See generally* Leon Gordis & Moyses Szklo, *Causation From Groups to Individuals*, 1 CTS. HEALTH SCI. & L. 362, 367-68 (1991) (discussing interaction between disease-causing agent and population subgroups in order to identify those who are more susceptible to the agent). Both environmental and genetic variables differentiate human subjects. As an example, consider oral contraceptives, an environmental disease determinant. Women taking oral contraceptives have different risk profiles from those not taking contraceptives. In particular, women taking contraceptives are at increased risk of disorders that can lead to a heart attack. *Id.* at 367. Susceptibility to high blood pressure is an example of a genetic determinant of disease. "[S]ome individuals, when exposed to a high salt intake . . . develop high blood pressure, whereas the same would not be true for other individuals lacking the relevant genetic trait." *Id.*

One epidemiologist writes:

> I believe that we have a central axiom, not subject to proof, but upon which epidemiology is based, and without which no epidemiology is possible.
>
> *Axiom*: Disease does not distribute randomly in human populations.
>
> *Corollary 1*: Nonrandom aggregations of human disease are manifested along axes of measurement of time, of space, of individual personal characteristics, and of certain community characteristics.
>
> *Corollary 2*: Variations in the frequency of human disease occur in response to variations in the intensity of exposure to etiologic agents or other more remote causes, or to variations in the susceptibility of individuals to the operation of those causes.

Stallones, *supra* note 4, at 80.

56. *See supra* notes 30-31 and accompanying text.

57. ROTHMAN, *supra* note 4, at 79.

58. The study size needed to detect differences between populations can be calculated. REMINGTON & SCHORK, *supra* note 48, at 181. This calculation requires that researchers first specify the desired significance level, usually a p-value less than or equal to .05, and estimate in advance the size of the risk ratio. *Id.*

59. *In re* Agent Orange Prod. Liab. Litig., 597 F. Supp. 740, 787-94 (E.D.N.Y. 1984), *aff'd*, 818 F.2d 145 (2d Cir. 1987), *cert. denied*, 484 U.S. 1004 (1988).

overall were not statistically significant.[60] The court concluded that the epidemiological studies did not "furnish sufficient support for plaintiffs' casuality [sic] claims . . . because of their small size, self-selective nature and other defects."[61]

Study size also was crucial in *Hoffman v. Merrell Dow Pharmaceuticals, Inc. (In re Bendectin Litigation)*, which consolidated claims of more than 800 plaintiffs who alleged that the defendant's morning sickness drug had caused birth defects in children whose mothers took the drug.[62] The Sixth Circuit Court of Appeals upheld a jury verdict for the defendant.[63] The court reviewed the following testimony by the plaintiffs' experts that addressed the issue of lack of precision: "the numbers . . . were too small" and "the study was incapable of detecting a relative [risk] for limb reductions that was smaller than sixfold."[64] As the court noted, such evidence can be an important factor in the overall assessment of admissibility.

c. Internal Validity—Lack of Scientific Bias

A valid study lacks significant scientific bias. Scientific bias is defined as

> any systematic error in the design, conduct, analysis, or interpretation of a study that tends to produce an incorrect assessment of the nature of the association between an exposure . . . and the occurrence of disease. More generally, bias has been defined as "any process at any stage of inference which tends to produce results or conclusions that differ systematically from the truth."[65]

Bias leads to a deviation of the study results, causing the perceived association to differ from the actual association. Bias is particularly troublesome for studies with weak statistical associations (risk ratio less than

60. *Id.* at 790. For a discussion of statistical significance, see *supra* notes 46-51 and accompanying text.

61. *Agent Orange Prod. Liab. Litig.*, 597 F. Supp. at 787; *see also* Troyen A. Brennan, *Can Epidemiologists Give Us Some Specific Advice?*, 1 CTS. HEALTH SCI. & L. 397, 397 (1991) (discussing the *Agent Orange* litigation); Daniel A. Farber, *Toxic Causation*, 71 MINN. L. REV. 1219, 1235 (1987) ("The key flaw in the plaintiffs' case was that government epidemiological studies . . . [and] [s]tudies by the Air Force, the CDC, and the Australian government all had concluded that no [significant] health effects had been demonstrated.").

62. 857 F.2d 290 (6th Cir. 1988), *cert. denied*, 488 U.S. 1006 (1989).

63. *Id.* at 326.

64. *Id.* at 318-19.

65. Manning Feinleib, *Biases and Weak Associations*, 16 PREVENTIVE MED. 150, 150 (1987) (citations omitted). For a discussion of bias, see CHECKOWAY et al., *supra* note 2, at 77-96; FLETCHER et al., *supra* note 22, at 7-14; and ROTHMAN, *supra* note 4, at 82-94.

three).[66] In such a study, if bias inflates the measure of association, the study's risk ratio is larger than it should be, and an innocuous substance may appear to cause disease. Alternatively, when bias decreases the measure of association, the study's risk ratio is smaller than it should be, leading to the conclusion that a substance does not cause disease when the converse is true.

Bias falls into three general categories: confounding, selection bias, and information bias.[67] As illustrated by *Bendectin Litigation*,[68] courts generally treat testimony regarding study bias as relevant and admissible.[69] However, evidence of bias may not be sufficient to refute values that are either very strong or very weak. Even though the plaintiffs' witnesses identified multiple sources of potential study bias, the judge in *Bendectin Litigation* viewed the absence of a statistical association as determinative, ruling that the plaintiffs failed to meet their burden of proving causation.[70]

The first general category of bias, confounding, "occurs when two factors or processes are [statistically] associated or 'travel together,' and the effect of one is confused with or distorted by the effect of the other."[71] Confounders can alter study results if they are distributed unequally in the population groups being compared. The court in *Bendectin Litigation* reviewed testimony that illustrates confounding.[72] In *Bendectin Litigation*, the plaintiffs offered expert testimony that there was no mechanism to control "for age or for other drugs."[73] The two potential confounders in this instance were age and the use of drugs other than Bendectin. Either of these factors could have increased the rate of birth defects, thereby skewing study results.

Bendectin Litigation also illustrates a second category of bias, selection bias. Selection bias results when improper procedures are used to select study subjects.[74] A Bendectin expert witness testified that "some patients were included as both Bendectin [patients] and controls."[75]

66. Feinleib, *supra* note 65, at 150.

67. *See, e.g., id.* ("[T]he varieties of biases that have been described in the literature are legion.").

68. *Bendectin Litig.*, 857 F.2d at 318-19.

69. *Id.*

70. *Id.* at 314.

71. FLETCHER et al., *supra* note 22, at 8. *See generally* 2 DORE, *supra* note 22, § 25.02(4)(a) (discussing bias).

72. *Bendectin Litig.*, 857 F.2d at 318.

73. *Id.*

74. ROTHMAN, *supra* note 4, at 83; *see also* 2 DORE, *supra* note 22, § 25.02(4)(b) (discussing bias); Feinleib, *supra* note 65, at 155-60 (same).

75. *Bendectin Litig.*, 857 F.2d at 319.

Such selection procedures could have skewed the results.

The third category of bias, information bias, distorts the study results because of "errors in obtaining . . . needed information."[76] The plaintiffs in *Bendectin Litigation* offered the following testimonial evidence: "[P]atients that were supposed to have used Bendectin . . . may not have in fact ingested it because there was nothing to record that they had ever been administered the drug."[77] This illustrates information bias because it relates to information-gathering techniques. Specifically, the study lacked documentation of whether patients actually took the drug.

Bias is a pervasive factor in epidemiological studies. Although researchers utilize complex means for preventing or mathematically correcting for bias, unrecognized bias can affect even well-designed studies. Successfully controlling bias generally increases a study's internal validity, which is defined as "the validity of the inferences drawn as they pertain to the actual subjects in the study."[78] Internal validity focuses on the actual subjects *within* the study, not on persons *outside* the study. If a study appropriately controls for significant bias, the risk ratio and ARP are likely to be valid for subjects within the study. This distinction is legally important since the plaintiffs usually are outside the actual study. Applicability to persons outside the study falls under the rubric of external validity, discussed in the next section. Before assessing external validity, however, the internal validity of the study must be established.

d. External validity

External validity concerns the appropriateness of applying epidemiological findings to persons outside the study.[79] It is an important concept because study results do not apply automatically to outside persons. After all, epidemiological study populations are merely samples of larger populations and may not be representative of those outside the study. Sir Austin Bradford Hill, a renowned epidemiologist, noted that the "[study] sample may, indeed be akin to that of the man who, according to Swift, 'had a mind to sell his house and carried a piece of brick in his pocket, which he showed as a pattern to encourage purchasers.' "[80] At times, it

76. ROTHMAN, *supra* note 4, at 84; *see also* 2 DORE, *supra* note 22, § 25.02(4)(c) (discussing information bias); Feinleib, *supra* note 65, at 160-63 (discussing bias).

77. *Bendectin Litig.*, 857 F.2d at 318.

78. ROTHMAN, *supra* note 4, at 82.

79. S. Hernberg, *Validity Aspects of Epidemiological Studies, in* EPIDEMIOLOGY OF OCCUPATIONAL HEALTH 269, 273-74 (M. Karvonen & M.I. Mikheev eds., 1986); ROTHMAN, *supra* note 4, at 82.

80. Austin B. Hill, *The Environment and Disease: Association or Causation?*, 58 PROC. ROYAL SOC'Y MED. 295, 299 (1965).

can be as ridiculous to use a sample to characterize the entire human race as to rely on a single brick to convey the appearance of a home.

The concept of external validity is significant legally because it is rare for a plaintiff to be an actual subject of an epidemiological study. A reasonable degree of external validity as applied to the particular plaintiff should be present before an epidemiological study is considered probative of causation in the plaintiff. Some commentators disagree with this conclusion and express more confidence in the predictive power of a study sample. *McCormick on Evidence* adopts this latter view: "Samuel Johnson once remarked that 'You don't have to eat the whole ox to know the hide is tough.' "[81]

No scientific methodology for analyzing external validity exists; indeed the subject often is not even addressed. External validity is a qualitative determination and is, therefore, "ultimately a matter of informed judgement."[82] Courts are particularly well-suited to make this determination. Essentially, one compares the characteristics of the study population to persons outside the study; similarity in all potentially pertinent factors is the dominant concern.[83] A pertinent factor could be, for example, the gender of the study subjects compared with the plaintiff's gender. If the subjects of the study were male and the plaintiff female, the substance-disease association should be examined to see if gender is a pertinent factor. The following excerpt illustrates such an analysis:

> [F]rom a study of smoking and lung cancer in men, one might generalize the results to a target population of women. To do so presumes that being male is irrelevant to the carcinogenic action that smoking has on lung tissue, a judgment based on knowledge about the likely mechanism of carcinogenesis and the biologic similarity between male and female lungs. On the other hand, a study of diet and [heart attack] in men might not be considered generalizable to women because physiologic difference between the sexes may play a role in the causal process.[84]

An epidemiological study is more likely to be legally probative if the plaintiff is similar to the study subjects with regard to all pertinent characteristics. Expert testimony or other evidence could identify the pertinent characteristics. Absent such evidence, courts could look for similarity in characteristics such as gender, age, and exposure level.

---

81. JOHN W. STRONG, MCCORMICK ON EVIDENCE § 208, at 935-36 (4th ed. 1992) (quoting DAVID S. MOORE, STATISTICS: CONCEPTS & CONTROVERSIES 3 (2d ed. 1985)).

82. ROTHMAN, *supra* note 4, at 95.

83. *Id.*

84. *Id.*

B. Statistical Association and Causation

"Cause" in epidemiology has been defined as "an event, condition, or characteristic that plays an essential role in producing an occurrence of the disease."[85] A principal question asked in epidemiology is whether exposure to a particular substance causes disease at the population level.[86] In considering this question, it is extremely important to understand that *statistical association does not establish causation*.[87] Additional factors must be considered. Hill decried the use of statistics, without more, to determine population-level causation, stating:

> [I]s there any other way of explaining the set of facts before us, is there any other answer equally, or more, likely than cause and effect? . . . No formal tests of [statistical] significance can answer those questions. Such tests can, and should, remind us of the effects that the play of chance can create, and they will instruct us in the likely magnitude of those effects. Beyond that they contribute nothing to the "proof" of our [causal] hypothesis.[88]

The following examples illustrate this point. In 1852, a scientific paper documented a strong association between low altitude and cholera, maintaining that low altitude was the cause of cholera.[89] Subsequent research dispelled this notion, establishing that the *Vibrio cholerae* bacteria, not low altitude, causes cholera.[90] The erroneous conclusion of the 1852 study illustrates the danger of relying solely on statistics to "prove" causation.

A more recent example is seen with the documented association between acquired immune deficiency syndrome (AIDS) and the illicit use of nitrate inhalants. Nitrate inhalants are drugs used by a number of gay men who later contracted AIDS.[91] Based on this association, early stud-

---

85. *Id.* at 11.

86. CHECKOWAY et al., *supra* note 2, at 13. A population-level analysis does not focus on whether exposure to a particular substance causes disease in a given *individual*, but rather on whether the exposure causes illness in a *population*—specifically, whether there are more cases of disease with the exposure than would have occurred without the exposure. *Id.*

87. ROTHMAN, *supra* note 4, at 7-21.

88. Hill, *supra* note 80, at 299.

89. Michael Dore, *A Proposed Standard for Evaluating the Use of Epidemiological Evidence in Toxic Tort and Other Personal Injury Cases*, 28 HOW. L.J. 677, 681 n.8 (1985) (citing William Farr, *Influence of Elevation on the Fatality of Cholera*, 15 J. STAT. SOC'Y LONDON 155 (1852)).

90. *See* Gerald T. Keusch, *Cholera*, *in* 1 HARRISON'S PRINCIPLES OF INTERNAL MEDICINE 632, 632 (Jean D. Wilson et al. eds., 12th ed. 1991).

91. Bert Black, *Matching Evidence About Clustered Health Events With Tort Law Requirements*, 132 AM. J. EPIDEMIOLOGY S79, S82 (Supp. 1990); James J. Goedert, *Recreational Drugs: Relationship to AIDS*, 437 ANNALS N.Y. ACAD. SCI. 192, 197-98 (1984).

ies suggested that the nitrate inhalants may have caused the depressed immune system in AIDS.[92] Scientists then looked at other factors such as the person-to-person transmission of AIDS through sexual contact or needle sharing, hypothesizing that an unknown infectious agent caused the disease.[93] This latter interpretation proved correct: AIDS is now known to be caused by an infectious agent, the human immunodeficiency virus (HIV).[94]

The association between nitrate inhalants and AIDS illustrates that two factors can be related, perhaps through a third factor, without having any cause and effect relationship. Statistical correlations also may exist entirely due to chance. As Professor Paul Sherman writes, "A statistical correlation may exist between two groups of numbers with no causal relationship being even possible."[95] Sherman cites the documented "direct statistical relationship between pig iron production in the United States and the British birth rate."[96] It would be ridiculous to conclude solely from this association that a direct cause and effect relationship exists between these two events.

The previous examples illustrate that statistical association alone cannot establish that a substance is capable of causing disease. Nevertheless, some legal commentators have suggested that at some level statistical association could be strong enough to satisfy a "more likely than not" standard in determining legal cause.[97] This concept is misguided, violating epidemiological tenets[98] as well as the legal doctrine of many courts.[99] Causation is an elusive concept that cannot be determined

92. Goedert, *supra* note 91, at 192-97.

93. *Id.*

94. Anthony S. Favci & Clifford Lane, *The Acquired Immunodeficiency Syndrome (AIDS)*, *in* 2 HARRISON'S PRINCIPLES OF INTERNAL MEDICINE, *supra* note 90, at 1402-03.

95. Paul Sherman, *Agent Orange and the Problem of the Indeterminate Plaintiff*, 52 BROOK. L. REV. 369, 384 n.81 (1986).

96. *Id.* (citing GEORGE W. SNEDOCOR & WILLIAM G. COCHRAN, STATISTICAL METHODS 189 (6th ed. 1967)).

97. *E.g.*, Charles Nesson, *Agent Orange Meets the Blue Bus: Factfinding at the Frontier of Knowledge*, 66 B.U. L. REV. 521, 521-23 (1986) [hereinafter Nesson, *Agent Orange*]. Professor Nesson suggests that "[a]t some point, high probability alone is sufficient to produce an acceptable verdict." *Id.* at 522 n.3.

98. ROTHMAN, *supra* note 4, at 4 ("[S]tatistical hypothesis testing is a mode of analysis that offers less insight into epidemiologic data than alternative methods . . . . [It] has often been transferred wholesale to epidemiology without giving sufficient thought to the underlying epidemiologic concepts.").

99. *See, e.g.*, Crim v. International Harvester Co., 646 F.2d 161, 165 (5th Cir. 1981) (holding that testimony that a significant association exists between a particular occupation and "valley fever does not constitute evidence of a causal connection"); Heckman v. Federal Press Co., 587 F.2d 612, 617 (3d Cir. 1977) ("[S]tatistical data about a group do not establish concrete facts about an individual."); Robinson v. United States, 533 F. Supp. 320, 330 (E.D.

solely by numerical calculations in either epidemiology or the law.

One widely cited logical exercise illustrating the dilemma courts face is the Blue Bus hypothetical, described by Professor Charles Nesson as follows:

> While driving late at night on a dark, two-laned road, a person confronts an oncoming bus speeding down the centerline of the road in the opposite direction. In the glare of the headlights, the person sees that the vehicle is a bus, but cannot otherwise identify it. He swerves to avoid a collision, and his car hits a tree. The bus speeds past without stopping. The injured person later sues the Blue Bus Company. He proves, in addition to the facts stated above, that the Blue Bus Company owns and operates 80% of the buses that run on the road where the accident occurred. Can he win?[100]

Nesson reasons that under a "more likely than not" standard that permits the plaintiffs to rely solely on statistics, a plaintiff would always recover if more than 50% of the buses were owned by the defendant's company.[101] Such a result could be considered unfair to a potentially non-negligent defendant. In reference to the Blue Bus hypothetical, he comments that "the acceptability of a conclusion is not a simple function of mathematical probability, but rather is a complex matter . . . that depends on the nature of the issue, the process of decision, and the purposes and audiences the conclusion serves."[102]

It is important to consider Nesson's comments in context, realizing that his rendering of the hypothetical represents a different situation from the one addressed by this Comment. In Nesson's hypothetical, statistics were particularly suspect, since they were being used to prove the identity of the defendant and, by extension, negligence and legal cause. It is not difficult to reject the use of statistic-based analyses in such a situation. By contrast, in the situation addressed by this Comment, sta-

---

Mich. 1982) ("Statistical evidence cannot establish cause and effect."); *see also* Leonard R. Jaffee, *Of Probativity and Probability: Statistics, Scientific Evidence, and the Calculus of Chance at Trial*, 46 U. PITT. L. REV. 925, 934 (1985) ("A statement of probability is merely an uncertain estimate of actuality."); Trauberman, *supra* note 17, at 198 & n.23 (1983) (stating that "courts have been reluctant to accept probabilistic evidence as showing causation"); Laurence H. Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329, 1341 n.37 (1971) (citing Smith v. Rapid Transit, 317 Mass. 469, 470, 58 N.E.2d 754, 755 (1945) (holding that it is not enough that " 'the mathematical chances somewhat favor the proposition' " to be proved)).

100. Charles Nesson, *The Evidence or the Event? On Judicial Proof and the Acceptability of Verdicts*, 98 HARV. L. REV. 1357, 1378-79 (1985) [hereinafter Nesson, *The Evidence or the Event?*].

101. *Id.*

102. Nesson, *Agent Orange*, *supra* note 97, at 522.

tistic-based evidence appears much later in the analysis. A defendant has been identified already and a plaintiff has been exposed to a particular substance because of the defendant's actions. A Blue Bus hypothetical analogous to this situation might be stated as follows:

> While standing in a large crowd at the bus station, a man becomes aware that a bus from the Blue Bus Company is nearing the crowd. The bus pulls into the station, and the crowd moves to stay clear of the bus. The man, although jostled somewhat by the crowd, does not appear to have been injured. Two years later, he develops pain in his lower back, and x-rays reveal a slipped disc. He sues the Blue Bus Company. He proves, in addition to the facts stated above, that statistics indicate it is dangerous for busses to drive near crowds because this can cause injuries to those in the crowd. In addition, he proves that there are more back injuries in crowd-related accidents than would be normally expected. Can he win?

In this hypothetical, the potentially liable party has been identified; the only remaining questions are negligence and causation. In such a situation, statistics seem less objectionable than in Nesson's stated hypothetical. Nevertheless, a given level of mathematical proof should not be entirely determinative.

C. Guidelines Used to Determine Causation in Epidemiology

There is some disagreement among epidemiologists concerning the procedures to be followed when determining causation. Most epidemiologists agree that there are two absolute requirements for causation: (1) temporality—the putative cause must precede its effect; and (2) association—there must be a statistical association between exposure to a substance and its effect.[103] Satisfaction of temporality and association is not sufficient to establish causation, however, and epidemiologists will consider additional factors, such as the biological plausibility of causation and the consistency of results among studies.

Criteria published in 1965 by Bradford Hill are often used as guidelines for the determination of causation (Table II).[104] Not all epidemiol-

---

103. Mervyn Susser, *What is a Cause and How Do We Know One? A Grammar for Pragmatic Epidemiology*, 133 AM. J. EPIDEMIOLOGY 635, 638 (1991).

104. Hill, *supra* note 80, at 295; *see also* CHECKOWAY et al., *supra* note 2, at 13 (discussing various criteria) (citing Alfred S. Evans, *Causation and Disease: a Chronological Journey*, 108 AM. J. EPIDEMIOLOGY 249 (1978)); James J. Schlesselman, *"Proof" of Cause and Effect in Epidemiologic Studies: Criteria for Judgement*, 16 PREVENTIVE MED. 195, 199-203 (1987) (discussing Bradford Hill criteria and other proposed criteria). Table II can be found at *infra* text accompanying note 111.

ogists use the Bradford Hill criteria, however;[105] some refuse to venture into the realm of causal inference at all.[106] Dr. Kenneth J. Rothman mentions, although does not necessarily advocate, the view that causal inference lies in the domain of public policy.[107]

Despite differing views, the Bradford Hill criteria are widely employed by epidemiologists[108] and will be used in this Comment to illustrate causal inference in epidemiology. The criteria, listed in modified form in Table II, include: (1) statistical association; (2) temporality; (3) biological plausibility and coherence; (4) dose-response gradient; (5) consistency; (6) analogy; (7) experimental evidence; and (8) specificity.[109]

---

105. Several commentators have recommended using a set of causal criteria known as the Henle-Koch postulates. *See* 2 DORE, *supra* note 22, § 25.04(4); Bert Black & David E. Lilienfeld, *Epidemiologic Proof in Toxic Tort Litigation*, 52 FORDHAM L. REV. 732, 763 (1984). Epidemiologists consider the Henle-Koch postulates to be limited, however, to situations involving exposure to infectious microorganisms. Susser, *supra* note 103, at 640; *see also* Marcia A. Mobiliz and Annette M. Rossignol, *The Role of Epidemiology in Determining Causation in Toxic Shock Syndrome*, Fall JURIMETRICS J. 78, 82-84 (1983) (applying the Henle-Koch postulates to toxic shock syndrome). The International Agency for Research on Cancer has suggested five criteria comparable to the Bradford Hill criteria to establish cancer causation. McElveen & Eddy, *supra* note 32, at 44-45 (citing 17 IARC 18 (1978)). The Bradford Hill criteria arose out of an earlier set of causal criteria advanced by the *Surgeon General's Report on Smoking and Health* in 1964. These criteria were consistency, strength, specificity, temporal sequence, and coherence. Feinleib, *supra* note 65, at 151 (citing 1964 SURGEON GEN. REP. ON SMOKING AND HEALTH). Other approaches to causal inference include that of Susser, who has urged consideration of three essential criteria: (1) association: a cause must be associated with its effect; (2) time order: a cause must precede its effect; and (3) direction: a change in an effect must be the result of a change in the putative cause. Susser, *supra* note 103, at 638-39.

106. ROTHMAN, *supra* note 4, at 20. Some epidemiologists reject the process of causal inference and the notion of criteria, instead favoring approaches based on deductive reasoning that arise in part from the philosophical works of Karl Popper. For a discussion of deductive reasoning and the influence of Popper, see Susser, *supra* note 103, at 642-43. *See generally* ROTHMAN, *supra* note 4, at 7-10 (discussing the philosophy of scientific inference).

107. Rothman states:

> Recently, Lanes . . . has proposed that causal inference is not part of science at all, but lies strictly in the domain of public policy. According to this view, since all scientific theories could be wrong, policy makers should weigh the consequences of actions under various theories. Scientists should inform policy makers about scientific theories, and leave the choice of a theory and an action to policy makers. Not many public health scientists are inclined toward such a strict separation between science and policy, but as a working philosophy it has the advantage of not putting scientists in the awkward position of being advocates for a particular theory . . . . Indeed, history shows that skepticism is preferable in science.

ROTHMAN, *supra* note 4, at 20 (citing S. Lanes, *Causal Inference is Not a Matter of Science*, 122 AM. J. EPIDEMIOLOGY 550, 550 (1985); Kenneth J. Rothman & Charles Poole, *Science and Policy Making*, 75 AM. J. PUB. HEALTH, 340, 340-41 (1985)).

108. Joshua E. Muscat & Michael S. Huncharek, *Causation and Disease: Biomedical Science in Toxic Tort Litigation*, 31 J. OCCUPATIONAL MED. 997, 997 (1989).

109. *See* Hill, *supra* note 80, at 295-300.

Not every criterion must be met for a causal relationship to be present; commonly at least one remains unsatisfied.[110] Rather than treating the criteria as a checklist, epidemiologists employ a balancing approach, viewing the criteria as a framework for weighing the evidence of causation. Some criteria are weighted more heavily than others. For example, a strong statistical association, while not sufficient to establish causation, lends more support to causal inference than does a high degree of plausibility and coherence. In addition, strength in one criterion may compensate for weaknesses in others: a strong temporal relationship may balance in favor of causation even in the face of a weak statistical association.

TABLE II: BRADFORD HILL CRITERIA—GUIDELINES OFTEN USED BY EPIDEMIOLOGISTS IN THE ASSESSMENT OF POPULATION-LEVEL CAUSATION.[111]

Statistical Association. There must be some degree of statistical association between a cause and its effect. A strong association (large in magnitude) is more likely to represent causation than a weak association (small in magnitude).
Temporality. A cause must precede its effect. Strength in temporality, such as when a cause immediately precedes its effect, supports an inference of causation.
Biological Plausibility and Coherence. A cause and effect relationship between exposure and disease should be biologically plausible and consistent with other information about the disease or harm.
Dose-Response Effect. Causation is more likely if greater amounts of the putative cause are associated with corresponding increases in the occurrence of disease or harm.
Consistency. When similar findings are generated by several epidemiological studies involving various investigators, causation tends to be supported.
Analogy. Substantiation of relationships similar to the putative causal relationship increases the likelihood of causation.
Experimental Evidence. Causation is more likely if removing the exposure in a population results in a decrease in the occurrence of disease or harm.
Specificity. When there is but a single putative cause for the disease or harm, causation is supported.

1. Statistical Association

Epidemiologists assess association by applying mathematical formulas to the results of well-designed population studies,[112] generating measures of association, such as the risk ratio. The strength of the association

110. As Hill stated, "None of my nine viewpoints can bring indisputable evidence for or against the cause-and-effect hypothesis and none can be required as a *sine qua non*." *Id.* at 299. *But cf. supra* text accompanying note 103 (temporality and association considered absolute requirements for causation).

111. *See* ROTHMAN, *supra* note 4, at 16-20; Hill, *supra* note 80, at 295-99.

112. Poor study design can seriously impair the epidemiologist's ability to assess meaningfully the strength of the association. FLETCHER et al., *supra* note 22, at 215-16; Moyses Szklo, *Design and Conduct of Epidemiologic Studies*, 16 PREVENTIVE MED. 142, 148-49 (1987).

is structured conceptually on a continuum, ranging from measures epidemiologists view as weak to those considered extremely strong.[113]

Epidemiological causal inference requires at least some degree of association. Although epidemiologists may accept a weak association when demonstrated strength exists in the other Bradford Hill criteria, a stronger association provides more evidence for inferring a causal relationship.[114] Stronger associations tend to compensate for the inherent weaknesses in population-based study designs.[115] Strong associations are preferred because they are less likely to be due to errors such as hidden bias.[116] A weak association, by contrast, easily could have been generated by error alone.

2. Temporality

The temporality criterion represents the notion that a cause must precede its effect. Satisfaction of temporality is considered a requirement for causal inference.[117] Strength in temporality, such as when a disease occurs shortly after its supposed cause, can add much credence to an association that is weak in other criteria. For example, in a series of cases the plaintiffs claimed they contracted a nervous system disorder, Guillain-Barre syndrome (GBS), from the swine flu vaccine given in the mid-1970s.[118] A study under the auspices of the Centers for Disease Control in Atlanta determined that the swine flu vaccine was a cause of GBS, but only during a ten-week period immediately following the vaccination.[119] The strength of the temporal relationship was striking, with GBS incidence increasing shortly after vaccination and decreasing to

113. *See supra* Table I, text accompanying notes 37-40.

114. ROTHMAN, *supra* note 4, at 17-18.

115. A physician from the National Cancer Institute explains:

> [Epidemiological studies] are quite weak at identifying the causes of very low levels of risk. Very small differences in risk between a group exposed to some substance versus that in a group not exposed to it could be due to a variety of reasons: for example, chance, or other differences between the exposed and unexposed which we either do not know about or cannot adequately control for.

McElveen & Eddy, *supra* note 32, at 39-40 (quoting Occupational Safety & Health Admin., Identification, Classification, and Regulation of Potential Occupational Carcinogens, 45 Fed. Reg. 5,040 (1980) (testimony of Dr. Robert Hoover)).

116. FLETCHER et al., *supra* note 22, at 217. For a discussion of bias and other factors that can influence epidemiological study results, see *supra* notes 46-84 and accompanying text.

117. *See* ROTHMAN, *supra* note 4, at 19. Rothman describes temporality as "a sine qua non: If the 'cause' does not precede the effect, that indeed is indisputable evidence that the association is not causal." *Id.*

118. *See, e.g.*, cases cited *infra* note 122.

119. Lawrence B. Schonberger et al., *Guillain-Barre Syndrome Following Vaccination in the National Influenza Immunization Program, United States, 1976-1977*, 110 AM. J. EPIDEMIOLOGY 105, 105 n.1 (1979).

normal within 10 weeks. A government compensation program was established for the victims.[120] Before awarding compensation, the government required victims to show proof of swine flu vaccination and onset of GBS within ten weeks of vaccination.[121] After the act was passed, a number of claimants who contracted GBS outside the ten-week window unsuccessfully brought suit against the government.[122]

Although temporality is necessary, it is not sufficient to establish epidemiological or legal causation. For example, several of the lawsuits arising from the swine flu vaccination program were unsuccessful because they were based almost entirely on temporality. The plaintiffs had contracted diseases other than GBS within a short time after being vaccinated for swine flu. Although temporality was strong, statistical association was weak or non-existent.[123] These plaintiffs failed to establish the government's liability. In *Kubs v. United States*,[124] the court sustained a verdict for the government because the plaintiff contracted polymyalgia rheumatica, not GBS, after being vaccinated for swine flu.[125] Similarly, in *Tabaczynski v. United States*,[126] the plaintiff did not prove that his polymyocitis was caused by the disease even though temporality was strong—the onset of his disease had occurred approximately one week after vaccination.[127]

3. Biological Plausibility and Coherence[128]

This criterion embraces the proposition that an inference of causation is supported if a cause and effect relationship between exposure and

120. National Swine Flu Immunization Program of 1976, 42 U.S.C. § 247b(j)-(l) (1978) (amended 1979).

121. Ginzburg, *supra* note 9, at 429.

122. *See, e.g.*, Gates v. United States, 707 F.2d 1141, 1147 (10th Cir. 1983) (eleven months after vaccination); Padgett v. United States, 553 F. Supp. 794, 804 (W.D. Tex. 1982) (sixteen weeks after vaccination); Thompson v. United States, 533 F. Supp. 581, 585-86 (N.D. Okla. 1981) (more than eleven weeks after vaccination); Hixenbaugh v. United States, 506 F. Supp. 461, 471-72 (N.D. Ohio 1980) (fifteen months after vaccination); Alvarez v. United States, 495 F. Supp. 1188, 1207 (D. Colo. 1980) (seven months after vaccination). *But see* Sulesky v. United States, 545 F. Supp. 426, 429-31 (S.D.W. Va. 1982) (relying on medical rather than epidemiological testimony when plaintiff, having contracted Guillain-Barre syndrome 14 weeks after vaccination, introduced conflicting epidemiological evidence that cast doubt on the basic premises of the government study).

123. The Swine Flu Act did not provide compensation for illnesses other than Guillain-Barre syndrome. Farber, *supra* note 61, at 1233.

124. 537 F. Supp. 560 (E.D. Wis. 1982).

125. *Id.* at 563.

126. 529 F. Supp. 156 (E.D. Mich. 1981).

127. *Id.* at 162.

128. In Hill's version, biological plausibility and coherence are separate criteria. Hill, *supra* note 80, at 295-300. They have been combined here because of their inherent similarity.

disease is biologically plausible. Further, if such a relationship is consistent with what is already known about the disease, causation is more likely. Biological plausibility is "often given considerable weight when assessing [epidemiological] causation."[129] Rothman, however, emphasizes that these criteria may be "difficult to judge," given that they are inherently limited by the extent of available knowledge.[130] He illustrates this point with an 1861 quotation about typhus, a disease now known to be caused by a bacteria transmitted by body lice:[131]

> It [would] be . . . ridiculous for the stranger who passed the night in the steerage of an emigrant ship to ascribe the typhus, which he there contracted, to the vermin with which bodies of the sick might be infested. . . . An adequate cause, one reasonable in itself, must correct the coincidences of simple experience.[132]

This early commentator mistakenly relied on common experience to denounce the cause of typhus, demonstrating that plausibility and coherence, though persuasive, can be misleading. Accordingly, they are not considered absolute requirements for the establishment of a causal relationship, and their absence does not disprove causation.

4. Dose-Response Effect

A dose-response effect is present if the occurrence of disease in a population increases as the exposure quantity increases.[133] A dose-response effect logically supports the notion of a cause and effect relationship.[134] As with plausibility and coherence, satisfaction of the dose-response criterion is not required for a determination of causation.[135]

5. Consistency

The consistency criterion asks whether other studies have generated similar results. If other epidemiological studies, using different popula-

---

See, e.g., ROTHMAN, *supra* note 4, at 19 (suggesting that biological plausibility and coherence are similar).

129. FLETCHER et al., *supra* note 22, at 219.

130. ROTHMAN, *supra* note 4, at 18-19.

131. 1 HARRISON'S PRINCIPLES OF INTERNAL MEDICINE, *supra* note 90, at 758.

132. ROTHMAN, *supra* note 4, at 19 (quoting David W. Cheever, 58 BOSTON MED. SURGICAL J. 449, 450 (1861)).

133. FLETCHER et al., *supra* note 22, at 217-18; ROTHMAN, *supra* note 4, at 18 (terming dose-response "biologic gradient").

134. FLETCHER et al., *supra* note 22, at 217.

135. For example, some researchers believe pregnant women's use of DES may cause cervical carcinoma later in the lives of their female fetuses, yet the causal association "show[s] no apparent trend of effect with dose." ROTHMAN, *supra* note 4, at 14, 18.

tions, support an inference of causation, then the association is more likely to be causal.[136] As with some of the other criteria, the presence of consistency makes the causal argument more persuasive, but the lack of consistency does not preclude causal inference.[137]

6. Analogy

The analogy criterion is based on the notion that if substances biologically or chemically similar to the one in question can cause a particular disease, the substance itself is more likely to cause that disease. For example, Hill suggests that if one substance ingested by a pregnant woman can cause birth defects, then, by analogy, a similar substance may cause them as well.[138] Epidemiologists consider satisfaction of this criterion to be very "weak evidence for causation."[139] The law similarly may view analogy to be a weak indicator of legal causation. One court held inadmissible as evidence an analogy nearly identical to that described by Hill.[140]

7. Experimental Evidence

Experimental evidence confirming the epidemiological findings supports an inference of causation. For example, when a hypothesized cause is removed from a population, the occurrence of disease should decrease if a cause and effect relationship exists.[141] Such an effect may be seen when preventive measures are applied to a population. A second type of experiment adds the putative cause in a controlled manner to a previously unexposed population; if a cause and effect relationship exists, the incidence of disease should increase. Although experiments of the latter type doubtlessly yield compelling evidence, usually ethical mores justifiably preclude such studies.

8. Specificity

Epidemiologists often do not consider specificity a prerequisite for

136. FLETCHER et al., *supra* note 22, at 219; ROTHMAN, *supra* note 4, at 18.

137. ROTHMAN, *supra* note 4, at 18 (stating that "some effects are produced by their causes only under unusual circumstances").

138. Hill, *supra* note 80, at 299.

139. FLETCHER et al., *supra* note 22, at 220.

140. Hoffman v. Merrell Dow Pharmaceuticals, Inc. (*In re* Bendectin Litig.), 857 F.2d 290, 321-22 (6th Cir. 1988), *cert. denied*, 488 U.S. 1006 (1989). The plaintiffs offered evidence that Thalidomide, another widely publicized morning sickness drug, caused birth defects. *Id.* The court ruled that references to Thalidomide were inadmissible because of the potential for prejudice. *Id.*

141. FLETCHER et al., *supra* note 22, at 218-19.

causation.[142] Specificity embodies the notion of "one cause, one effect."[143] For example, only exposure to asbestos causes asbestosis, and only exposure to silicon dust causes silicosis. One problem with specificity is that the manner in which a disease is defined determines whether specificity exists.[144] The terms asbestosis and silicosis encompass, by definition, two separate diseases, each with a single cause. Had the more general term "fibrosis" been applied to both diseases, specificity would have been lost. Another problem with the specificity criterion is that a number of diseases have multiple causes, such as ischemic heart disease, high blood pressure, and lung cancer.[145] Conversely, exposure to some substances can cause more than one disease.[146] Therefore, although the *presence* of specificity tends to prove causation, the *absence* of specificity is not meaningful.[147]

In law as in epidemiology, specificity is not required to prove causation. It would be ludicrous to insist that a particular result could have only one possible cause. A house fire, for example, could have a range of possible causes, from children experimenting with matches to a faulty electrical connection. Similarly, it would be ridiculous to insist that a given cause could have only one effect. The faulty electrical connection, for example, could cause anything from a mild odor to a full-fledged explosion. As in epidemiology, however, the *presence* of specificity is highly probative of causation: a victim's wound containing a bullet fired from a particular type of gun makes it highly likely that the victim's wound was caused by that type of gun.

---

142. ROTHMAN, *supra* note 4, at 18 (describing the specificity criterion as "useless and misleading").

143. FLETCHER et al., *supra* note 22, at 220.

144. *Cf.* Feinleib, *supra* note 65, at 151-52 (noting that a disease's definition may change as new nuances of the disease arise).

145. *See generally* Andrew P. Selwyn & Eugene Braunwald, *Ischemic Heart Disease*, *in* 2 HARRISON'S PRINCIPLES OF INTERNAL MEDICINE, *supra* note 90, at 964-65 (discussing multiple causes of ischemic heart disease); Gordon H. Williams, *Hypertensive Vascular Disease*, *in* 2 HARRISON'S PRINCIPLES OF INTERNAL MEDICINE, *supra* note 90, at 1002-03 (discussing multiple causes of high blood pressure); John D. Minna, *Neoplasms of the Lung*, *in* 2 HARRISON'S PRINCIPLES OF INTERNAL MEDICINE, *supra* note 90, at 1103-04 (discussing multiple causes of lung cancer).

146. Fletcher gives the following example. "[L]ung cancer is caused by cigarette smoking, asbestos, and radiation. Cigarettes cause not only lung cancer but also bronchitis, peptic ulcer disease, periodontal disease, and wrinkled skin. So the absence of specificity is not much of a strike against a cause-and-effect relationship." FLETCHER et al., *supra* note 22, at 220.

147. The requirement for specificity "has often been advanced, especially by those seeking to exonerate smoking as a cause of lung cancer. Causes of a given effect, however, cannot be expected to be without other effects on any logical grounds. In fact, everyday experience teaches us repeatedly that single events may have many effects." ROTHMAN, *supra* note 4, at 18.

9. Summary

The term "Bradford Hill criteria" is really a misnomer because it implies that a set of standards must be met before one can infer epidemiological causation.[148] As previously discussed, these criteria are really guidelines, not absolute requirements; epidemiological causation commonly is determined without the presence of every criterion. Only two of the eight criteria are absolute requirements for causal inference—temporality and association. These two criteria demand that the cause precede the effect and that some degree of statistical association between cause and effect be present. Although temporality and association are mandatory, they are not determinative of causation. The other criteria also should be considered, and to the extent they weigh in favor of or against causal inference, a cause and effect interpretation becomes more or less likely.

II. Using Epidemiology to Establish Legal Cause

Causation in law and causation in epidemiology can be analogized to the different methods of measuring time described in the quotation from *Alice in Wonderland* at the beginning of this Comment.[149] The Mad Hatter's watch measured time on a broad-scaled day-by-day basis. Like the Mad Hatter's watch, epidemiology measures causation broadly—at the population level. It examines whether exposure to a substance causes a particular disease in the population. Alice's watch, by contrast, measured time in smaller increments. The law, like Alice's watch, attempts to narrow the scope of causation down to the level of the individual. Just as it would have been difficult for Alice to tell time using the Hatter's day-based watch, courts understandably have struggled to determine legal cause on an individual level using population-based epidemiological analyses.

The discussion of legal cause in this Comment focuses on negligence claims, one of the most common theories under which toxic tort lawsuits are brought. Under a theory of negligence, liability is found when the plaintiff establishes by a preponderance of the evidence that the defendant's conduct was negligent and that this negligence was the legal cause of the plaintiff's injury.

The Second Restatement of Torts defines negligence as "an unreasonable risk of (1) causing harm to a class of persons of which the other is a member and (2) subjecting the other to the hazard from which the

148. "Criterion" is defined as "a standard, rule, or test on which a judgment or decision can be based." THE AMERICAN HERITAGE DICTIONARY 341 (2d ed. 1985).

149. *See supra* text accompanying note 1.

harm results."[150] Population-based epidemiological analyses are useful both in assessing the risk of harm and in deciding the unreasonableness of the risk.[151] Population-level causation proves that the substance was capable of causing disease. This evidence, combined with proof that the plaintiff was exposed at disease-causing levels to the substance for which defendant is responsible, satisfies most of the elements of this tort.

Liability requires a finding not only of negligence, but legal cause as well.[152] The notions of proximate and actual cause, distinct concepts in the First Restatement, were incorporated by the Second Restatement into the single concept of legal cause.[153] The Second Restatement provides guidance as to what constitutes legal cause. Section 431 states that "negligent conduct is a legal cause of harm to another if . . . [the] conduct is a substantial factor in bringing about the harm."[154] This section has two components. The first component embodies the concept of actual cause, also termed "but-for" cause: "[T]he harm would not have occurred had the actor not been negligent."[155] Actual "but-for" causation in cases involving exposure to harmful substances requires inferential reasoning. The Restatement's comment to section 431 suggests that actual cause is necessary but not sufficient. The plaintiff also must show that the defendant's conduct was a substantial factor in bringing about the harm. This second component incorporates the extent of the defendant's participation.[156] In the words of the Restatement:

---

150. RESTATEMENT (SECOND) OF TORTS § 430 cmt. a (1965).

151. *See* Dore, *supra* note 7, at 435 ("Epidemiology can prove that a defendant's conduct put plaintiff at risk."); *supra* note 22 and accompanying text.

152. RESTATEMENT (SECOND) OF TORTS § 430 (1965). The Second Restatement of Torts states that "[for] a negligent actor [to] be liable for another's harm, it is necessary not only that the actor's conduct be negligent toward the other, but also that the negligence of the actor be a legal cause of the other's harm." *Id.*

153. North Carolina continues to use the term "proximate cause," defined in a recent court of appeals' decision as follows:

> [A] cause which in natural and continuous sequence, unbroken by any new and independent cause, produced the plaintiff's injuries, and without which the injuries would not have occurred, and one from which a person of ordinary prudence could have reasonably foreseen that such a result, or consequences of a generally injurious nature, was probable under all the facts as they existed.

Warren v. Colombo, 93 N.C. App. 92, 100, 377 S.E.2d 249, 254 (1989) (quoting Hairston v. Alexander Tank & Equip. Co., 310 N.C. 227, 233, 311 S.E.2d 559, 565 (1984) (citing Kanoy v. Hinshaw, 273 N.C. 418, 160 S.E.2d 296 (1968))); *see* Robert G. Byrd, *Proximate Cause in North Carolina Tort Law*, 51 N.C. L. REV. 951, 954-55 (1973).

154. RESTATEMENT (SECOND) OF TORTS § 431 (1965).

155. *Id.* § 431 cmt. a.

156. *See id.* Cases generally require that the plaintiff establish by a preponderance of the evidence that the defendant's negligence was a substantial factor in causing the plaintiff's injury. *See, e.g.*, Hart v. Ivey, 102 N.C. App. 583, 592, 403 S.E.2d 914, 920 (1991) (stating that liability attaches if the defendant's negligence is a substantial factor in causing injury to the

> The word "substantial" is used to denote the fact that the defendant's conduct has such an effect in producing the harm as to lead reasonable men to regard it as a cause, using that word in the popular sense, in which there always lurks the idea of responsibility, rather than in the so-called "philosophic sense," which includes every one of the great number of events without which any happening would not have occurred.[157]

Restatement section 432(2) notes an exception to the requirement of establishing actual cause. This exception applies when two forces combine to cause indivisible harm, one force being the negligent defendant and the other an outside force.[158] For example, in *Anderson v. Minneapolis, Saint Paul & Sault Ste. Marie Railway Co.*,[159] two fires—one of innocent origin and one due to the defendant's negligence—merged, and the combined fire caused harm to the plaintiff.[160] Under the Restatement exception, if the defendant's negligence was a substantial factor in bringing about the plaintiff's harm, further proof is not needed and the defendant is liable for the whole injury.[161] Prosser explained this exception as follows:

> When the conduct of two or more actors is *so related to an*

---

plaintiff) (quoting Hutchens v. Hankins, 63 N.C. App. 1, 9, 303 S.E.2d 584, 591 (1988)); Wyatt v. Gilmore, 57 N.C. App. 57, 59, 290 S.E.2d 790, 791 (1982) (stating that an element of proximate cause is "whether the cause was a substantial factor in bringing about the result"). *See generally* Trauberman, *supra* note 17, at 177, 197 (stating that the "plaintiff must show by a preponderance of the evidence that the defendant's behavior was a substantial factor in causing his or her injury").

157. RESTATEMENT (SECOND) OF TORTS § 431 cmt. a (1965).

158. *Id.* Comment a to § 430 states that "[e]xcept as stated in § 432(2), [actual cause] is necessary." *See generally* Joseph H. King, Jr., *Causation, Valuation, and Chance in Personal Injury Torts Involving Preexisting Conditions and Future Consequences*, 90 YALE L.J. 1353, 1356 (1981) (discussing use of "substantial factor" test in determining actual causation).

159. 146 Minn. 430, 179 N.W. 45 (1920). In *Anderson*, the defendant was found liable when two fires, one negligently set by the defendant and the other caused by a bolt of lightening, merged, and the combined fire destroyed the plaintiff's property. *Id.* at 440-41, 179 N.W. at 49.

160. *Id.*

161. RESTATEMENT (SECOND) OF TORTS § 432(2) (1965); *see, e.g.*, Rozark Farms v. Ozark Border Elec. Coop., 849 F.2d 306, 311 (8th Cir. 1988) (Liability for a plaintiff's entire damages attaches to any negligent tortfeasor when "plaintiff's damages arise out of an indivisible loss which the defendant's negligence was a substantial factor in causing."). Some courts view this as a weaker standard for establishing liability and restrict its use to specific factual situations. In *Bendectin Litigation*, a weaker standard applied "only to initial negligent actors in determining their liability in the face of action by a subsequent actor, or in determining causation between simultaneous actors, both of whose acts could have been "but for" causes of plaintiffs' injuries." Hoffman v. Merrell Dow Pharmaceuticals, Inc. (*In re* Bendectin Litig.), 857 F.2d 290, 311 (6th Cir. 1988), *cert. denied*, 48 U.S. 1006 (1989). The court also rejected a "related, but somewhat distinct 'increased the risk' standard." *Id.* at 311 n.15 (citing Cooper v. Sisters of Charity, Inc., 27 Ohio St. 2d 242, 272 N.E.2d 97 (1971)).

> *event* that their combined conduct, viewed as a whole, is a but-for cause of the event, and application of the but-for rule to them individually would absolve all of them, the conduct of each is a cause in fact of the event.[162]

The focus here is again on the extent of the defendant's participation and responsibility.

When the plaintiff has been exposed to a disease-causing substance, one could view two or more forces in operation: a negligent force and one or more outside forces. The negligent force would be exposure to the substance, and the outside forces would be other components that combine with the negligent substance exposure to cause the disease.[163] This model first requires proof that exposure to the substance can cause the disease, and, second, that exposure was a substantial factor in the plaintiff's case. The second criterion means that exposure must have been so related to the event that it could be considered a but-for cause. The Missouri Court of Appeals recently used this logic in its analysis:

> [A]lthough our law requires proof of cause to recover in tort, it does not require proof of a single cause. The substantial factor standard—which ascribes liability to a cause which has played an important part in the production of the harm, even though the harm may have occurred absent that cause—is particularly suited to injury from chronic exposure to toxic chemicals where the subsequent manifestation of biological disease may be the result of a confluence of causes.[164]

Consistent with this approach, Rothman analogizes the causes of disease to the pieces of a pie.[165] Each piece represents a cause contributing to the disease. When all of the contributing causes are assembled into a whole pie, the disease occurs. With a disease such as high blood pressure, for example, the contributing causes could be genetic propensity, exposure to salt in the diet, exposure to stress, and other unknown factors. Each person's pie is unique; the size and relative contribution of each contributing cause varies from individual to individual. In some people, the genetic component is particularly strong, yielding a large

---

162. W. PAGE KEETON ET AL., PROSSER AND KEETON ON THE LAW OF TORTS § 41, at 268 (5th ed. 1984) (emphasis added).

163. *See generally* McElveen & Eddy, *supra* note 32, at 32 (articulating a multiple cause scientific theory for cancer). For a discussion of multiple cause involving a negligent defendant and one or more outside innocent causes, see Robert G. Byrd, *Actual Causation in North Carolina Tort Law*, 50 N.C. L. REV. 261, 275 (1972).

164. Elam v. Alcolac, Inc., 765 S.W.2d 42, 174 (Mo. Ct. App. 1988), *cert. denied*, 493 U.S. 817 (1989).

165. For an in-depth description of this model, including the strength of various causes, interaction among causes, and the proportion of disease due to specific causes, see ROTHMAN, *supra* note 4, at 10-16.

piece attributable to genetics that has filled most of the pie. For such an individual, a small piece of any other cause would make a whole pie and result in high blood pressure. Other individuals have such a small genetic component that no amount of dietary salt or stress at work will result in high blood pressure.

This model helps explain why exposure to a substance causes disease in some people but not others and underscores the difficulty of ascribing causation to a single factor. It also emphasizes that population-based epidemiological studies measure averages. Conceptually, everyone's pies are different for any given disease. Some of us require very little exposure to a particular "cause" to develop a disease; others require a larger exposure; still others will never contract the disease regardless of the magnitude of the exposure.

In summary, establishing liability under a claim of negligence generally requires that a plaintiff prove both negligence and legal cause; in other words, the plaintiff must show that the defendant's conduct posed an unreasonable risk of harm, and that the defendant's negligence "was more likely than not a substantial factor in bringing about the plaintiff's injury."[166] There are really two steps in proving legal causation. The first is a population-level analysis which asks whether the substance is "capable of causing the type of harm from which the particular plaintiff suffers."[167] The second is an individual-level determination which examines whether the substance caused harm in the particular plaintiff.[168] Epidemiological studies can be relevant in both steps of the causation analysis. It should be emphasized, however, that any causation determination should involve more than simply applying statistical measures of association. Statistical association by itself does not establish disease causation at either the population or the individual level.[169]

A. Admissibility of Epidemiological Evidence to Prove Legal Cause

In jurisdictions with rules of evidence similar to the Federal Rules of Evidence, epidemiological studies may form the basis for expert testimony or be admitted as evidence through the public records and reports exception to the hearsay rule.[170] Federal Rule of Evidence 702 requires that the expert testimony "assist the trier of fact to understand the evi-

166. Dore, *supra* note 7, at 430.

167. Farber, *supra* note 61, at 1227.

168. *Id.* Dore recognizes these concepts and suggests that causation may be divided into "risk"—the population-level analysis and "occurrence"—the causation issue at the individual level. Dore, *supra* note 7, at 435.

169. *See supra* notes 85-148 and accompanying text.

170. *E.g.*, *In re* Agent Orange Prod. Liab. Litig., 611 F. Supp. 1223, 1240 (E.D.N.Y. 1985)

dence or to determine a fact in issue."[171] An expert may testify with or without offering an opinion, as long as the testimony will "assist the trier of fact."[172] When the expert offers an opinion, no prior disclosure of the basis of the opinion is needed; however, the rules anticipate that the scientific data will be revealed on cross-examination.[173] The expert is permitted to base his or her opinion on inadmissible studies, but they must be "of a type reasonably relied upon by experts in the particular field in forming opinions or inferences upon the subject."[174] The admissibility of expert opinion testimony, therefore, depends on whether the studies are of a type reasonably relied on by experts in the field. As suggested in *McCormick on Evidence*, "[t]he judge and the attorneys may treat the matter in a hearing under Rule 104."[175] Alternatively, doubts as to the bases of an opinion could lead to the testimony being stricken or merely affect the weight of the evidence.[176]

Whether an epidemiological study is used as the basis for expert testimony or proffered as evidence itself, the basic principles of relevancy and prejudice apply.[177] As stated by Professor George James, relevancy

---

(citing FED. R. EVID. 803(8)), *aff'd*, 818 F.2d 187 (2d Cir. 1987), *cert. denied*, 487 U.S. 1234 (1988).

171. FED. R. EVID. 702.

172. The Federal Rules of Evidence provide: "If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise." *Id.* This standard applies in North Carolina. *See, e.g.*, Tomkins v. Log Sys., Inc., 96 N.C. App. 333, 337, 385 S.E.2d 545, 548 (1989) (citing N.C. GEN. STAT. § 8C-1 (1988) (N.C. R. EVID. 702)), *rev. denied*, 326 N.C. 366, 389 S.E.2d 819 (1990).

173. The Federal Rules of Evidence state: "The expert may testify in terms of opinion or inference and give reasons therefor without prior disclosure of the underlying facts or data, unless the court requires otherwise. The expert may in any event be required to disclose the underlying facts or data on cross-examination." FED. R. EVID. 705.

174. The Federal Rules of Evidence provide:

> The facts or data in the particular case upon which an expert bases an opinion or inference may be those perceived by or made known to the expert at or before the hearing. If of a type reasonably relied upon by experts in the particular field in forming opinions or inferences upon the subject, the facts or data need not be admissible in evidence.

FED. R. EVID. 703.

In North Carolina, the phrase "reasonably relied upon by experts" has been construed to mean "inherently reliable." Cherry v. Harrell, 84 N.C. App. 598, 605-06, 353 S.E.2d 433, 438 (1991) (applying "inherently reliable" standard articulated prior to adoption of the rules by State v. Wade, 296 N.C. 454, 462, 251 S.E.2d 407, 412 (1979)), *rev. denied*, 320 N.C. 167, 358 S.E.2d 49 (1987). For an expanded discussion, see Walter J. Blakey, *Examination of Expert Witnesses in North Carolina*, 61 N.C. L. REV. 2, 20-32 (1982).

175. STRONG, *supra* note 81, § 15, at 67 & n.11.

176. *Id.* § 13, at 56 n.15.

177. *E.g.*, *id.* § 203, at 875-76.

"is not an inherent characteristic of any item of evidence but exists as a relation between an item of evidence and a proposition sought to be proved."[178] Therefore, the relevance of epidemiological studies is determined in the context of a relation between the study and the proposition to be proved—usually, legal causation in the individual plaintiff. Federal Rule of Evidence 401 defines relevant evidence as that which has "any tendency to make the existence of any fact that is of consequence to the determination of the action more probable or less probable than it would be without the evidence."[179] The threshold for this determination is low: "[I]t is enough if the item could reasonably show that a fact is slightly more probable than it would appear without that evidence."[180] A finding of relevance, however, does not assure admissibility. Under Federal Rule of Evidence 403, the potential for prejudice and jury misuse may render relevant evidence inadmissible.[181] The relevancy and prejudice analysis has been characterized as a three-step process: first, a relevancy determination; second, an assessment of the potential for prejudice; and third, a balancing of relevancy and prejudice.[182]

1. Assessing the Probative Value of an Epidemiological Study

Epidemiological studies are statistic-based. Because of this, juries often view epidemiological studies as connoting scientific certainty when there may be none and infer individual causation without adequate basis.[183] The probative value of epidemiological calculations breaks down

178. George F. James, *Relevancy, Probability and the Law*, 29 CAL. L. REV. 689, 690 (1941).

179. FED. R. EVID. 401.

180. STRONG, *supra* note 81, § 185, at 776.

181. FED. R. EVID. 403. The rule states in pertinent part that "[a]lthough relevant, evidence may be excluded if its probative value is substantially outweighed by the danger of unfair prejudice, confusion of the issues, or misleading the jury." *Id.*; *e.g.*, STRONG, *supra* note 81, § 185, at 779-81; Paul C. Giannelli, *The Admissibility of Novel Scientific Evidence:* Frye v. United States, *a Half-Century Later*, 80 COLUM. L. REV. 1197, 1239 (1980) (evidence excluded if potential prejudice substantially outweighs probative value).

The North Carolina Supreme Court applied this analysis in a 1987 case addressing the relevance of statistics to the issue of paternity. State v. Jackson, 320 N.C. 452, 461, 358 S.E.2d 679, 683 (1987). The *Jackson* court upheld the admission of testimony about statistical data, but excluded as prejudicial the expert's statement that the defendant " 'probably [was] the father of the child.' " *Id.* at 461, 358 S.E.2d at 683. Notably, this case was decided in the wake of an earlier "probability of paternity" case, Cole v. Cole, 74 N.C. App. 247, 328 S.E.2d 446, *aff'd*, 314 N.C. 660, 335 S.E.2d 897 (1985). In *Cole*, the court overruled a finding of paternity based on a 95.98% probability of paternity when additional evidence showed the man to be sterile. *Id.* at 255, 328 S.E.2d at 451.

182. Giannelli, *supra* note 181, at 1235.

183. *See generally* Dore, *supra* note 7, at 437-38 (stating that juries may be unduly influenced by apparent certainty of statistic-based studies); Tribe, *supra* note 99, at 1331 (discussing the potential for jury misuse of numerical evidence). For a general discussion of the admissi-

in the transition from population to individual-level causation.[184] Further, epidemiological studies carry a substantial potential for jury confusion and misuse.[185] Dore remarks that

> [u]sing epidemiological evidence . . . presents the danger of confusing and prejudicing the jury, which . . . may well confuse a showing that the defendant's conduct increased the plaintiff's risk of disease with proof that the defendant's conduct more likely than not caused the plaintiff's disease. . . . Furthermore, juries may tend to give epidemiological evidence excessive weight.[186]

Dore views epidemiological evidence as statistical evidence.[187] In this context, his statements are in accord with this Comment, which asserts that statistics alone do not establish causation in either epidemiology or the law.[188] Professor Laurence H. Tribe also cautions against the risk of jury misuse of mathematical-based studies:

> [I]n at least some contexts, permitting *any* use of certain mathematical methods entails a sufficiently high risk of misuse, or a risk of misuse sufficiently costly to avoid, that it would be irrational not to take such misuse into account when deciding whether to permit the methods to be employed at all.[189]

Some courts have adopted a particular value for the risk ratio as a cut-off point for sufficiency under a "more likely than not" standard for legal causation in the individual.[190] These courts recognize that when the risk ratio equals two, the ARP is 50%—indicating a two-fold increase in risk. However, these courts incorrectly assume this translates into a 50% chance that the plaintiff's disease was caused by exposure to the substance. The court in *Marder v. G.D. Searle & Co.*[191] stated that "a two-fold increased risk is an important showing for plaintiffs to make because it is the equivalent of the required legal burden of proof—a showing of causation by a preponderance of the evidence or, in other words, a probability of greater than 50%."[192] According to Professor

---

bility of epidemiological evidence, see Richard E. Hoffman, *The Use of Epidemiologic Data in the Courts*, 120 AM. J. EPIDEMIOLOGY 190, 192-93 (1984).

184. Robins & Greenland, *supra* note 42, at 1125-26.

185. Dore, *supra* note 7, at 435.

186. *Id.* at 437.

187. *Id.* at 431.

188. *See supra* notes 85-148 and accompanying text.

189. Tribe, *supra* note 99, at 1331.

190. *See, e.g.*, cases cited *supra* note 41.

191. 630 F. Supp. 1087, 1092 (D. Md. 1986), *aff'd sub nom.* Wheelehan v. G.D. Searle & Co., 814 F.2d 655 (4th Cir. 1987).

192. *Id.* According to one commentator, "Courts have typically equated this required jury determination [i.e., preponderance of the evidence] with a degree of certainty exceeding fifty

Sherman, these courts follow a "so-called 'weak' preponderance rule with respect to statistical proof[,] . . . [an] approach [that] allows a plaintiff to recover upon a statistical showing of likelihood of causation (in excess of fifty percent), even if there is no direct evidence linking the plaintiff or his injuries to the defendant or its product."[193] As documented in epidemiological and statistical literature, however, population-based calculations should not be extrapolated directly to the individual.[194]

Dore notes that "[c]ourts that fail to distinguish the issue of risk from that of actual causation may . . . erroneously permit the evidence of risk to establish causation."[195] Consider the following hypothetical:

> The plaintiff files a lawsuit alleging negligence. Facts in evidence establish that the plaintiff had worked for the defendant for twenty years before contracting lung cancer. He claims that because of the defendant's negligence, he was exposed to disease-causing levels of a harmful substance which caused his cancer. The plaintiff first convinces the court that the substance can cause lung cancer at the population level by offering evidence satisfying most of the Bradford Hill criteria. He utilizes several workplace studies of men who had been exposed to low levels of the substance for approximately ten years. The plaintiff had been exposed to high levels of the substance for twenty years. The plaintiff now seeks to use the studies' relatively low ARPs of approximately 51% to establish legal causation at the level of the individual.

In this hypothetical, the plaintiff seeks to introduce ARPs from several studies to prove legal cause. However, each ARP represents an average value that should not be applied as a definitive cut-off for a given level of proof.[196] Whether a study ARP should be applied to show causation in an individual plaintiff depends on several factors. One such factor is the magnitude of the ARP. A large ARP, in the range of 90-99%, would certainly tend to support an inference of individual cause.

---

percent." David Rosenberg, *The Causal Connection in Mass Exposure Cases: A "Public Law" Vision of the Tort System*, 97 HARV. L. REV. 849, 857 (1984) (citing MICHAEL O. FINKELSTEIN, QUANTITATIVE METHODS IN LAW 66-67 (1978) and CHARLES T. MCCORMICK, HANDBOOK ON THE LAW OF DAMAGES § 31, at 118 (1935)).

193. Sherman, *supra* note 95, at 384.

194. Robins & Greenland, *supra* note 42, at 1134-35. This is largely due to "the unknown mechanisms by which exposure affects disease risk and competing risks[,] . . . the unknown degree of heterogeneity in the background risks of disease, and . . . the unknown degree of dependence between risk of disease and competing risks." *Id.*; *see also supra* notes 65-84 and accompanying text (discussing internal and external validity).

195. Dore, *supra* note 7, at 436.

196. *See supra* notes 41-45 and accompanying text.

But perhaps the most important factor is external validity—the ability to apply study results to those outside the study. External validity as it applies to an individual would depend on the plaintiff's similarity to the subject population; accordingly, a comparison between the study population and the plaintiff is needed.

Returning to the hypothetical, the magnitudes of the studies' ARPs, at approximately 51%, are quite low. It is possible, because of factors such as scientific bias, that the ARPs actually lie below 50%.[197] If this is the case, then one could not even say that the disease of a person *within* one of the study populations was "more likely than not" associated with exposure to the substance. If an ARP cannot be applied to an individual within the study population, then it certainly should not be applied to an individual, such as the plaintiff, outside of the study. Accordingly, an ARP should not be used as evidence of legal cause unless more evidence is presented.

In the hypothetical case, however, the plaintiff can prove additional facts that support an inference of individual causation. For example, he was exposed to the harmful substance for a greater number of years than were the subjects in the studies, during which time his daily exposure levels were higher. These facts support the notion that any ARP applicable to the plaintiff probably would be greater than 51% and evidence based on these studies should be admitted. Ultimately, the decision would rest with the court.[198]

In conclusion, the probative value of an epidemiological study clearly depends on a number of factors. These factors include magnitude of the risk ratio or ARP, adequacy of the study design, scientific bias, precision, external validity, and the presence or absence of other studies. It would be erroneous, therefore, to find the numerical values from epide-

197. Conversely, the ARP could lie *above* 51%.

198. The notion of adding to and subtracting from a probability is one that should be familiar to any court that has used Bayes theorem. For a description of Bayes theorem, see generally David Kaye, *The Laws of Probability and the Law of the Land*, 47 U. CHI. L. REV. 34, 49-53 (1979) (illustrating how the formula can be used to convey to jurors the probative force of quantitative evidence); David Kaye, *The Paradox of the Gatecrasher and Other Stories*, 1979 ARIZ. ST. L.J. 101, 106-08 (1979) (demonstrating how the formula can be used to describe the way knowledge of a new item of evidence would influence a completely rational decision maker's evaluation of the probability of some fact in dispute). With Bayes theorem, evidence supportive of causation would increase the ARP by a specific amount and evidence tending to negate causation would decrease it. However, Bayes theorem creates a statistic that is only as accurate as the numbers on which it was premised. Although in some cases Bayes theorem could provide a closer approximation of the applicable ARP value, this advantage may be negated by the danger of prejudice resulting from the perception of certainty by courts of mathematical formulas.

miological studies relevant to show legal causation in an individual plaintiff without considering these other factors.

2. Conditions for Admissibility of Epidemiological Evidence

a. The Frye Rule

Many courts require that certain conditions be met for evidence based on scientific studies to be admissible.[199] The original test,[200] established in *Frye v. United States*,[201] requires that the scientific evidence have gained "general acceptance in the particular field in which it belongs."[202] Lately, the "Frye rule" has been criticized, and some courts apply the rule only to questionable scientific methodologies, rather than to the results of particular studies.[203] Some commentators object to placing explicit conditions on admissibility, urging that "traditional standards of relevancy and the need for expertise—and nothing more—should govern."[204]

b. Epidemiological Evidence

Several commentators advocate restrictions on the admissibility of epidemiological evidence, including:

> [1.] Restrictions upon the admission of epidemiological studies as to which inadequate discovery has been provided. . . .
> [2.] Admission only of epidemiological evidence which satisfies the Henle Koch Postulates and also demonstrates an attributable risk for the factor in question in excess of 50 percent.
> [3.] Admission only of epidemiological studies which are highly statistically significant.[205]

The restrictions advocated by Professor Bert Black and Dr. David Lilienfeld are the most comprehensive, requiring: (1) a finding of causation at the population level using appropriate criteria, and (2) an ARP of greater than 50%.[206] Black and Lilienfeld argue that if their test is applied, then the legal standard will be satisfied. In other words, they claim

199. *See, e.g.*, STRONG, *supra* note 81, § 203, at 868.

200. *Id.* § 203, at 869.

201. 293 F. 1013 (D.C. Cir. 1923).

202. *Id.* at 1014. For a discussion of the history of the "Frye rule," see Giannelli, *supra* note 181, at 1204-31.

203. *See, e.g.*, STRONG, *supra* note 81, § 203, at 871. *But cf.* Giannelli, *supra* note 181, at 1204-31 (concluding that "Frye may be tottering, but has not yet fallen").

204. STRONG, *supra* note 81, § 203, at 874.

205. 2 DORE, *supra* note 22, § 25.05(2) (citing Black & Lilienfeld, *supra* note 105, at 767); Dore, *supra* note 89, at 691-95.

206. Black & Lilienfeld, *supra* note 105, at 767.

that satisfaction of the test assures that it would be more likely than not that a particular plaintiff's disease was caused by the exposure.[207] This reasoning is unacceptable because it represents an attempt to substitute directly a population-based standard for an individual-based standard. In addition, Black and Lilienfeld recommend using the Henle-Koch postulates in every situation. These postulates, however, usually apply only when an infectious agent is involved.[208] Finally, they downplay the problem of study bias and the crucial concept of external validity.[209]

Rather than espousing specific criteria, it would be more useful for courts to use qualitative guidelines that indicate whether epidemiological evidence is relevant in a given situation.[210] Admittedly, such guidelines should not eliminate completely the use of numerical epidemiological measures, but rather should condition their admission on the presence of additional qualitative evidence.[211]

B. Guidelines for the Use of Epidemiological Evidence in Toxic Tort Lawsuits

The toxic tort lawsuit presents a perplexing problem; its outcome is highly uncertain and its resolution usually involves enormous financial stakes.[212] Epidemiological studies are crucial elements of proof in these lawsuits. Guidelines based on epidemiological tenets could prove extremely useful to a court attempting to extrapolate population-based data to the level of the individual plaintiff. Guidelines could assist a judge in determining whether an epidemiological study is "of a type reasonably relied upon by experts in the particular field," the threshold for admissi-

---

207. *Id.*

208. *See supra* note 105.

209. In a footnote, the authors allude to the notion of external validity as they caution the reader that exposure levels in the epidemiological study should be the same as or greater than those in the plaintiff. Black & Lilienfeld, *supra* note 105, at 767 n.143. This, however, would result in an over-estimation of the risk to the plaintiff. To make an inference of individual causation, the exposure levels in the epidemiological study actually should be as similar as possible to those experienced by the plaintiff.

210. A more qualitative test was proposed by two commentators in 1983 in the context of sufficiency, not admissibility. It would require that the plaintiff establish: "(1) exposure significant enough to trigger disease; (2) a demonstrated, biologically plausible relationship between the chemical and disease; (3) the diagnosis of such disease in the plaintiff; and (4) expert opinion that the plaintiff's disease was . . . consistent with exposure to the chemical." Khristine L. Hall & Ellen K. Silbergeld, *Reappraising Epidemiology: A Response to Mr. Dore*, 7 HARV. ENVTL. L. REV. 441, 445 (1983).

211. *See* Dore, *supra* note 7, at 431 (stating that epidemiological evidence "may help demonstrate that a particular event occurred, but only when accompanied by more specific evidence").

212. DEBORAH H. HENSLER ET AL., TRENDS IN TORT LITIGATION, THE STORY BEHIND THE STATISTICS 31 (1987).

bility of expert opinion testimony.[213] Additionally, guidelines could aid the judge in the determination of basic relevancy and the potential for prejudice.[214] After admissibility issues are decided, similar guidelines could be helpful in determining sufficiency of the evidence.

1. Guidelines for Admissibility

The following guidelines are proposed for evaluating the admissibility of epidemiological evidence when seeking to prove legal cause in an individual plaintiff:[215] (1) whether the plaintiff can prove facts from which an inference of population-level causation could be made, through the examination of factors such as strength of the statistical association, biological plausibility, and consistency among studies; and (2) whether the plaintiff can prove facts to support an adequate degree of external validity—that the pertinent characteristics of the epidemiological study populations, including exposure level, were similar enough to those of the plaintiff to make extrapolation reliable.

The first guideline requires that the plaintiff be able to prove facts from which an inference of population-level causation could be made. Causation at the population level may be determined using traditional epidemiological methods, including application of the Bradford Hill criteria or similar guidelines.[216] As already discussed, two absolute requirements for a finding of causation at the population level are temporality and statistical association.[217] Evidence of strength in these criteria should be shown and, if possible, in other suitable criteria such as biological plausibility and coherence, dose-response gradient, consistency, analogy, experimental evidence, and specificity.[218] Factors such as study bias[219] and lack of precision[220] also should be accorded full consideration. These factors go to the very heart of the study, asking whether the study results really represent what a party claims they do.

---

213. FED. R. EVID. 703.

214. *See generally* STRONG, *supra* note 81, § 203, at 875 (discussing the admissibility of scientific evidence).

215. This Comment offers open-ended guidelines, allowing courts to evaluate fully the case at hand, using the standards and evidentiary devices they prefer.

216. *See supra* notes 103-48 and accompanying text. A recent scientific paper describes the Bradford Hill criteria as "uniform scientific criteria" and recommends their use in toxic tort cases. Muscat & Huncharek, *supra* note 108, at 997-98. However, as previously discussed, the Bradford Hill criteria should be considered guidelines, not explicit criteria. Other criteria may be better suited to certain exposure-disease combinations. *See supra* note 105.

217. *See supra* text accompanying note 103.

218. *See supra* notes 103-48 and accompanying text.

219. *See supra* notes 65-78 and accompanying text.

220. *See supra* notes 52-64 and accompanying text.

This first guideline probably could best be applied as a conditional relevancy test.[221] The relevancy of an epidemiological study to prove individual-level causation depends on a condition—the existence of population-level causation. As such, it should be admitted "upon, or subject to, the introduction of evidence sufficient to support a finding of" population-level causation.[222] Courts and commentators agree that admissibility of epidemiological studies to show individual causation should be conditioned on such proof.[223] Some judges have ordered bifurcated or trifurcated trials,[224] addressing the issue of population-level causation prior to that of legal cause. For example, in the recent *Bendectin Litigation*, the trial judge ordered a trifurcated trial, first addressing the issue of causation.[225] The court of appeals upheld a finding of no causation at the population level under an abuse-of-discretion standard of review and accordingly affirmed that the question of legal cause at the individual plaintiff level was moot.[226] In addition, the court held that trifurcation was not a source of unfair prejudice and was dispositive of the individual-level causation issue.[227]

Alternatively, the first guideline also could be considered under a more traditional relevancy-based approach. Professor George James wrote that a relevancy analysis involves deductive logic, which in turn depends on the existence of a "major premise" for its effect.[228] Accordingly, the following sequence could be established for epidemiological evidence:

> The plaintiff was exposed to substance X through the defendant's negligence. The plaintiff later developed disease Y.

To take the next logical step, that "the apparent probability of [the defendant's negligence] is now greater than before the evidence . . . was received,"[229] the following major premise would need to be established

221. Rule 104(b) states that "[w]hen the relevancy of evidence depends upon the fulfillment of a condition of fact, the court shall admit it upon, or subject to, the introduction of evidence sufficient to support a finding of the fulfillment of the condition." FED. R. EVID. 104(b).

222. *Id.*

223. *See, e.g.*, Hoffman v. Merrell Dow Pharmaceuticals, Inc. (*In re* Bendectin Litig.), 857 F.2d 290, 317 (6th Cir. 1988) (holding that determination that Bendectin did not cause birth defects at the population level was "dispositive of the litigation"), *cert. denied*, 488 U.S. 1006 (1989). *But see* Wendy E. Wagner, Note, *Trans-Science in Torts*, 96 YALE L.J. 428, 430-31 (1986) (urging that population-level causation should not be necessary, just that plaintiff's injury be "consistent with that substance").

224. *See* FED. R. CIV. P. 42(b) (allowing the judge to separate issues at trial).

225. *Bendectin Litig.*, 857 F.2d at 296.

226. *Id.* at 294.

227. *Id.* at 314-17.

228. James, *supra* note 178, at 694-99.

229. *Id.* at 699.

first:

> MAJOR PREMISE: Persons exposed to substance X are more likely to develop disease Y than those not exposed.

This premise could be satisfied through a population-level causation determination.

Under the second guideline, the plaintiff should be able to prove that the subject population is similar enough to the plaintiff to make possible the extrapolation of the study results to the plaintiff. Essentially, this guideline requires the court to perform an external validity analysis. As Dore states, "If the characteristics of the study group differ markedly from those of the plaintiff[,] . . . the results of the study cannot be applied reliably to the plaintiff. For example, studies of occupational risks may reveal little about dangers to the general public."[230] Consider the following two hypothetical plaintiffs:

> The first plaintiff is a twenty-year-old woman who has developed a rare form of cancer and claims the defendant is liable because she was exposed to "substance X" ten years ago when workers used it in her home for a period of two weeks. The second plaintiff is a fifty-year-old man who has been using "substance X" in his job more than thirty hours a week for the past fifteen years and has developed the same rare cancer.

Both plaintiffs seek to use epidemiological studies implicating "substance X" as a cause of the rare cancer. The subjects in the epidemiological studies were men who had been exposed to "substance X" for an average of twenty years in an occupational setting. Although epidemiological study results should not be applied automatically to either plaintiff, clearly there are factors which make the study more generalizable to the second plaintiff. The key is determining which characteristics are pertinent to the issue of causation and which are not. Expert testimony can be helpful here, and, at the very least, a court can look for the existence of characteristics such as gender, age, and level of exposure, which are similar in the plaintiff and the subject population.

2. Guidelines for Sufficiency

This Comment proposes the following guideline to be used when evaluating the sufficiency of epidemiological evidence: whether the plaintiff has proved the facts in (1) and (2), above,[231] as well as additional facts that make it more likely than not that exposure to a particular substance caused the plaintiff's disease. One such additional fact could be a strong

230. Dore, *supra* note 7, at 435-36 (citations omitted).

231. *See supra* text accompanying note 215.

statistical association between exposure and the occurrence of disease at the population level.[232] Epidemiological study results indicating a strong association between exposure and disease would tend to support, but not be determinative of, legal causation. To be considered a strong association, the risk ratio should be greater than or equal to eight (ARP greater than or equal to 87%).[233] Strong associations should not always be required, however. A moderate association[234] coupled with strength in other epidemiological causation guidelines, such as a strong temporal relationship, also could satisfy this guideline.[235] With a weak association, the plaintiff should probably show strength in other epidemiological causation guidelines and, in addition, similarity between the plaintiff and those in the study population.

III. CONCLUSION

Analyzing causation when epidemiology and the law converge is complex, often requiring an interpretation of research at the very frontiers of modern science. An epidemiologist makes population-level causation determinations, relying on both statistic-based analyses and qualitative criteria such as temporality, biological plausibility, and consistency. The courts, on the other hand, have a greater, more difficult task—determining causation not only at a population level, but also at the level of the individual. Courts are inconsistent in their treatment of epidemiological evidence. If courts are to apply epidemiology,[236] it is

232. Professor E. Donald Elliott has proposed that courts rely on strong causal associations. E. Donald Elliott, *Goal Analysis Versus Institutional Analysis for Toxic Compensation Schemes*, 73 GEO. L.J. 1357, 1368 (1985).

233. *See supra* Table I, text accompanying notes 37-40. As stated by Sherman, "If this statistical correlation is large, plaintiffs could point to it as evidence of causation in the sense that the product was a substantial factor in producing their claimed injury." Sherman, *supra* note 95, at 381 (referring to the *Agent Orange* litigation).

234. *See supra* Table I, text accompanying notes 37-40.

235. Epidemiological causation guidelines, such as the Bradford Hill criteria, are described at *supra* notes 103-48 and accompanying text.

236. Proposals have been made in the context of toxic tort litigation to replace the traditional tort system with policy-based alternative liability schemes. Several commentators have advocated a model allowing plaintiffs to be compensated in proportion to the level of risk from a particular exposure. *See* Troyen A. Brennan & Robert F. Carter, *Legal and Scientific Probability of Causation of Cancer and Other Environmental Diseases in Individuals*, 10 J. HEALTH POL. POL'Y L. 33, 58 (1985) (recommending that legislatures adopt policy of compensation proportional to risk); Richard Delgado, *Beyond Sindell: Relaxation of Cause-In-Fact Rules for Indeterminate Plaintiffs*, 70 CAL. L. REV. 881, 908 (1982) (encouraging a proportionate recovery system); Rosenberg, *supra* note 192, at 859-60 (arguing that liability and compensation should be proportional to risk). Professor Farber discusses another risk-based compensation system, the "Most Likely Victim Model." Farber, *supra* note 61, at 1221. Under this model, "plaintiffs whose injuries were least likely to have been caused by the defendant receive nothing, while those with the highest causation probabilities get full compensa-

vitally important that they have some understanding of the applications and limitations of its basic tenets.

The guidelines proposed in this Comment recognize that statistic-based population data represent at best a rough approximation,[237] and for this reason should not be applied with certainty to an individual plaintiff in the absence of other evidence. Risk ratios and ARPs from an epidemiological study generally may not be extrapolated with precision to particular subjects within the study itself, largely due to the potential for study bias, lack of study precision, and the inherent variability in any human population.[238] Even if the results could be applied to individuals *within* the study, they may not be relevant to persons *outside* the study. External validity, the ability to generalize to persons outside the study, is a crucial determinant in the analysis of causation in the individual plaintiff.[239] Judge Weinstein, in the *Agent Orange* litigation, stated:

> While it may be possible to prove, through the use of such proof as . . . epidemiological evidence, that such harm—for example cancer—can be "caused" by a particular substance, it may be impossible to pinpoint which particular person's cancer would have occurred naturally and which would not have occurred but for exposure to the substance.[240]

The proposed guidelines respond to these difficulties and recommend that courts consider more than the statistic-based elements of epidemiological proof. Qualitative factors such as consistency among studies and external validity should also be addressed.

Professor Tribe urges careful scrutiny of statistics in the courtroom

---

tion." *Id.* Another proposed system creates a new cause of action based on group compensation. Kenneth A. Cohen, *Class Actions, Toxic Torts, and Legal Rules*, 67 B.U. L. REV. 581, 597 (1987). In this system, rather than proving individual-level causation, plaintiffs must prove they are members of a group for which causation has been established for a particular substance. *Id. But see* PETER H. SCHUCK, AGENT ORANGE ON TRIAL 268-72 (1986) (arguing that the group compensation model violates traditional tort doctrine by requiring defendants to compensate individuals they did not injure). Professor Sherman has suggested combining the class action suit with proportional recovery schemes, liability being delegated according to the ARP. Sherman, *supra* note 95, at 381. In a class action suit, the plaintiff could use epidemiological evidence to establish population-level causation. Plaintiffs able to show exposure to the substance and membership in the group for whom population-level causation has been established could receive damages "proportionate to the increased incidence of the condition" relative to the general population. *Id.* An in-depth discussion of alternative liability systems is beyond the scope of this Comment. For an overview, see generally 2 DORE, *supra* note 22, §§ 6.01-07.

237. *See supra* notes 42-45 and accompanying text.

238. *See supra* notes 52-78 and accompanying text.

239. *See supra* notes 79-84 and accompanying text.

240. *In re* Agent Orange Prod. Liab. Litig., 597 F. Supp. 740, 834 (E.D.N.Y. 1984), *aff'd*, 818 F.2d 145 (2d Cir. 1987).

and acknowledges that statistics have some utility "when properly combined with other, more conventional evidence."[241] The proposed guidelines are in accord with this view, recommending that courts require more than numerical epidemiological measures. They should not simply accept data from an epidemiological study as applicable to a given plaintiff. Regarding this matter, Professor Louis L. Jaffe has stated:

> [A]bstract probability may play a role in finding a fact, but what is referred to in the traditional formula is the greater probability in the case at hand. The 'probabilities' in the abstract or statistical sense is only a datum. The jury's quest for the fact can only be undertaken if there is evidence in addition to that upon which the mere abstraction is based which will enable the jury to make a reasoned choice between the competing possibilities. . . . There must be a *rational*, i.e., evidentiary basis on which the jury can choose the competing possibilities. If there is not, the finding will be based (in the words of the formula) on mere speculation and conjecture.[242]

In using epidemiology to prove legal cause, courts have found themselves in a statistic-based quandary. Can epidemiology be trusted? Should it be rejected? Or does its proper use lie somewhere in-between? Consider Lewis Carroll's Alice, trying to use the White Rabbit as her guide, but, ultimately, being forced to rely on her own reasoning. Similarly, although the courts may seek to use epidemiology, resolution of the issue of causation ultimately depends on a tempered application of epidemiology within the logic and reasoning of the law.

MELISSA MOORE THOMPSON

241. Tribe, *supra* note 99, at 1350. It should be noted that Professor Tribe's comments were made in the context of a criminal trial where presumably greater precautions are needed to protect the defendant.

242. Louis L. Jaffe, *Res Ipsa Loquitur Vindicated*, 1 BUFF. L. REV. 1, 4 (1951) (quoted by Rosenberg, *supra* note 192, at 857 n.38).