THE BRADFORD HILL CRITERIA: THE FORGOTTEN PREDICATE

Dr. Frank C. Woodside, III* & Allison G. Davis**

Table of Contents

---

* Frank C. Woodside, III is of counsel to the law firm of Dinsmore & Shohl. Dr. Woodside is a nationally-known trial lawyer representing manufacturers of pharmaceutical and medical devices, chemicals and flavorings, as well as producers of consumer products. Over a period of more than 40 years, he has tried 80-plus cases to verdict or judgment, serving as primary trial counsel in medical malpractice, product liability, and mass tort cases. Dr. Woodside's clients include Fortune 500 companies, health care providers, and hospitals. His background to practice Medicine and Surgery for nearly 40 years affords him added knowledge and insight he uses to his clients' advantage in preparing for and defending high-stakes litigation. Dr. Woodside, who received his medical degree from the University of Cincinnati in 1973, is the editor of the treatise, *Drug Product Liability*, published by Matthew Bender, and an emeritus clinical professor of Pediatrics at the University of Cincinnati College of Medicine. Dr. Woodside has been listed for the past 20 years in The Best Lawyers in America. He earned his J.D. and B.S. from The Ohio State University.**Allison Davis is a member of Dinsmore & Shohl's Litigation Department and Torts and Environmental Practice Group. Her work includes representing clients in matters involving claims of breach of contract, negligence, and fraud. Ms. Davis has also represented companies on matters related to regulatory requirements under the Clean Water Act and Clean Air Act. She has experience defending clients in a variety of industries, including pharmaceutical manufacturers, privately held coal companies and commercial real estate developers. Ms. Davis earned her J.D. from the University of Virginia School of Law and her B.A. from the University of Kentucky. Dr. Woodside and Ms. Davis would like to thank the following members of the Thomas Jefferson Law Review for their support and assistance with the following article: Justin Abbate (Chief Articles Editor), Jessica Mulvaney (Chief Articles Editor), Travis Davis (Executive Editor), and Walter Araujo (Editor-in-Chief).

I. INTRODUCTION

In product liability litigation, the determinants of causation between an agent and a disease are frequently a critical issue. In 1965 Sir Austin Bradford Hill, a British epidemiologist[1] and statistician, delivered a speech to the Royal Society of Medicine in which he presented a list of what are often referred to as the Bradford Hill criteria.[2] With increasing frequency, state and federal courts mention these criteria when discussing the admissibility of epidemiological evidence.[3] Courts referring to the Bradford Hill criteria, however, provide limited analysis of them. Furthermore, counsel seeking to base the admissibility of scientific evidence on satisfaction of the criteria generally does so without significant analysis of the criteria or relevant case law. In view of this limited analysis by counsel and the courts, it seems appropriate to explore the significance of the Bradford Hill criteria—if any—in consideration of the admissibility of epidemiological evidence on the issue of causation.[4]

Bradford Hill's list was an expansion of one offered previously in the landmark U.S. Surgeon General's Report on Smoking and Health (1964).[5] Hill himself never labeled the *criteria* as such.

1. *Epidemiology* is a field of public health and medicine that studies the incidence, distribution, and etiology (causation or origin) of disease in human populations. *See infra* Part II.A.

2. For reasons discussed later, Bradford Hill's list can be more aptly described as "considerations" as opposed to "criteria"; however, throughout this article we will refer to them as "criteria" in order to be consistent with prior literature on the matter. *See* A. Bradford Hill, *The Environment and Disease: Association or Causation*, 58 PROC. ROYAL SOC'Y MED. 295 (1965).

3. *See, e.g.*, *In re* Fosamax Prods. Liab. Litig., 645 F. Supp. 2d 164 (S.D.N.Y. 2009); Dunn v. Sandoz Pharms. Corp., 275 F. Supp. 2d 672 (M.D.N.C. 2003).

4. Although discussed in more detail, the analysis by the commentators is also frequently limited. *See, e.g.*, Terence M. Davidson & Christopher P. Guzelian, *Evidence-Based Medicine (EBM): The (Only) Means for Distinguishing Knowledge of Medical Causation from Expert Opinion in the Courtroom*, 47 TORT TRIAL & INS. PRAC. L.J. 741 (2012).

5. *See* KENNETH J. ROTHMAN ET AL., MODERN EPIDEMIOLOGY 297 (3d ed. 2008); *see also* U.S. DEP'T OF HEALTH, EDUC., & WELFARE, SMOKING & HEALTH: REPORT OF

Rather, Hill discussed nine "aspects of [] association"[6] that one examining causation should "especially consider" once an association has been established as "perfectly clear-cut"[7] and "beyond what we would care to attribute to the play of chance."[8] In other words, no examination of the criteria should be undertaken unless an association has first been concretely established. Even then, Hill's list constitutes more *considerations* than criteria.[9]

Authors of scientific literature and the courts often forget that an association must first be established as perfectly clear-cut. However,

---

THE ADVISORY COMMITTEE TO THE SURGEON GENERAL OF THE PUBLIC HEALTH SERVICE 19 (1964).

6. In epidemiology there can be an association that is not causal. An association is seemingly established when, for example, an outcome occurs more frequently following an exposure than chance would predict. However, this can be a result of study bias, study design, or an innumerable amount of other confounding factors. *See* Michael D. Green, D. Mical Freedman & Leon Gordis, *Reference Guide on Epidemiology*, *in* REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 552–53 (3d ed. 2011) [hereinafter *Reference*].

7. Although *perfectly clear-cut* is never defined by Bradford Hill, one might presume that he meant it to mean (1) a properly performed study (of sufficient power and without bias or confounding that would produce results that depart from the true value) and (2) a study that yields a "statistically valid" finding. However, Bradford Hill never mentions the first, and seemingly negates the second, when he discounts the value of significance testing. Hill, *supra* note 2, at 299. Since there are no scientific or legal definitions for these terms, consulting a thesaurus to get a more complete grasp of their meaning would seem to be appropriate. Synonyms for *perfectly* include "completely," "entirely," and "wholly." *See Perfectly*, THESAURUS.COM, http://thesaurus.com/browse/perfectly?s=t (last visited Jan. 20, 2013). Synonyms for *clear-cut* include "definitive," "obvious," "unambiguous," "undoubted," and "unequivocal." *See Clear-Cut*, THESAURUS.COM, http://thesaurus.com/browse/clear-cut?s=t (last visited Jan. 20, 2013).

8. *See* Hill, *supra* note 2. For an association to be beyond the play of chance, the increase in the odds ratio or relative risk must be statistically significant and must have resulted from (1) a properly performed study (of sufficient power and without bias or confounding that would produce results that depart from the true value) and (2) a study that yields a "statistically valid" finding. *See infra* Part II.B. That is, the increase in the relative risk or the odds ratio must be statistically significant, which means that the confidence interval does not include the number one. *See* Miller v. Pfizer, Inc., 196 F. Supp. 2d 1062, 1080 (D. Kan. 2002); s*ee also infra* Part II.B. The expert must have statistically significant studies to serve as basis of opinion on causation. *Miller*, 196 F. Supp. 2d 1062. Since Bradford Hill (in the 1965 presentation that is the subject of this paper) disputes the value of significance testing, one can only conclude that he would urge that an impermissible experiential or subjective assessment of the data be undertaken prior to developing the conclusion that an association is "perfectly clear cut" or "beyond what we would care to attribute to the play of chance."

9. Bradford Hill never uses the word *criteria*: instead, on two occasions referring to them as "nine viewpoints." At various other points the so-called criteria are also referred to as "aspects," "features," "evidence," "requirements," and "characteristics." *See* Hill, *supra* note 2, at 295–99.

it is key that this predicate be met before undertaking an assessment of Hill's criteria. If the predicate is not established, analysis of the criteria cannot serve as the basis for an opinion on causation.

Bradford Hill suggested that the following criteria be examined when considering whether a perfectly clear-cut association is causally related to the exposure (chemical, pathogen, or other substance) being studied: (1) strength of association, (2) consistency, (3) specificity, (4) temporality, (5) biological gradient, (6) plausibility, (7) coherence, (8) experiment, and (9) analogy.[10] These factors are commonly taught in epidemiological lectures today and used throughout the field in assessing causation.[11] While epidemiologists do not consider it necessary that all these criteria be met before drawing inferences about causation, the current scientific community generally accepts consideration of the criteria as sound methodology.[12]

The criteria are preserved in the hard copy of Bradford Hill's speech. To date, there is no indication that the content of the presentation was peer reviewed, and no literature supporting the validity of the criteria is cited or referenced. As Kenneth Rothman—a well known author, scholar, and researcher in the field of epidemiology—would agree, this lack of examination in conjunction with the "misguided but popular view that his considerations should be used as criteria for causal inference[,] makes it necessary to examine them in detail."[13] Whereas Rothman inspected these

10. *Id.* While it may have been the position of Bradford Hill that a subjective assessment would constitute the basis for such a conclusion, under existing case law it would have to be a statistically significant increase in the relative risk. *See Miller*, 196 F. Supp. 2d at 1080.

11. *See* Carl V. Phillips & Karen J. Goodman, *The Missed Lessons of Sir Austin Bradford Hill*, EPIDEMIOLOGIC PERSP. & INNOVATIONS (Oct. 4, 2004), http://archive.biomedcentral.com/1742-5573/content/1/1/3.

12. "Sound methodology" and general acceptance by the scientific community are not equivalent. It is possible that the "accepted methodology" may not be sound. *See* Morgan v. Sheppard, 188 N.E.2d 808 (Ohio Ct. App. 1963) ("Customary conduct or methods of treatment which are generally employed by physicians and surgeons in the diagnosis, care, and treatment of a patient do not furnish a test which is controlling on the question of negligence, or fix a standard by which negligence can be gauged."). It should also be noted that, at one point in time, the world was considered to be flat. *See* Martin v. Commissioner, 649 F.2d 1133, 1144 (5th Cir. 1981) ("The conclusion that the world was flat led to certainty in navigation, albeit almost at the price of discovering the New World. While the certainty of an erroneous solution may represent a temporary comfort for those who embrace it, it may well represent an injustice for those who are subject to it."); *see also* Merrell Dow Pharms., Inc. v. Havner, 953 S.W.2d 706, 718 (Tex. 1997).

13. ROTHMAN ET AL., *supra* note 5, at 26. Kenneth Rothman holds D.M.D. (Dental Medicine), Ph.D. (Epidemiology), and M.P.H. (Epidemiology) degrees. *See*

considerations from the perspective of an epidemiologist, it also seems appropriate to assess the viability of Hill's various criteria under *Daubert* and its progeny—a type of a legal peer review.[14]

Part II of this article will give a brief background on the field of epidemiology and its role in the court system, as well as offer a discussion of *Daubert* and its progeny. Part III will discuss each of the nine criteria by examining what Bradford Hill stated about each from an epidemiological perspective and ascertaining what role—if any—each of these criteria plays in assessing causation under the law of *Daubert* and its progeny.

II. EPIDEMIOLOGY, *DAUBERT*, AND ESTABLISHING ASSOCIATION

In order to consider the application of the Bradford Hill criteria in a legal setting, it is first necessary to fully understand the field of epidemiology, Bradford Hill's statements that proceed the delineation of his criteria in his 1965 presentation, and the standards for the admissibility of expert testimony. As laid out below, a predicate to the examination of scientific evidence in light of the Bradford Hill criteria is that a critical analysis has been performed to determine if (1) an association has been clearly established, (2) the studies suggesting an association are valid, and (3) the *Daubert* requirements (relating to the admissibility of expert testimony) have been satisfied.

A. Background on the Field of Epidemiology

The court system recognizes the field of epidemiology and the utility of epidemiological studies.[15] Epidemiology is a field of public health and medicine that studies the incidence, distribution, and etiology (causation or origin) of disease in human populations.[16] It can be described as a two-step process that begins with a statistical analysis of collected data in order to determine if any scientifically valid associations exist. Then, if an association exists, the epidemiological process requires a determination of the information or biologic conclusions that can be derived from such data.[17]

Kenneth J. Rothman, B.U. SCH. PUB. HEALTH, http://sph.bu.edu/index.php?option=com_sphdir&INDEX=668&Itemid=340 (last visited Feb. 16, 2013).

14. *See* Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579 (1993).

15. *Reference*, *supra* note 6, at 551.

16. *Id.*

17. *See* DAVID E. LILIENFELD & PAUL D. STOLLEY, FOUNDATIONS OF EPIDEMIOLOGY 12 (3d ed. 1994) ("[B]asically, the Epidemiologist uses a two-stage sequence of reasoning: 1) The determination of a statistical association between a

Epidemiology focuses on general causation as opposed to specific causation.[18] In other words, the field of epidemiology is not intended to utilize the results of a group study to demonstrate causation for any individual plaintiff.[19] Instead, the studies are performed or undertaken to first determine if a statistically significant association exists between an exposure and an outcome. If such an association is revealed and the studies are determined to be free of confounding,[20] bias,[21] or other error, then an association can be established. At this point, epidemiologists and others interpreting the epidemiologic data can make an inference vis-à-vis the existence of a causal relationship or the lack thereof.[22]

B. Establishing Association

Association is a term of art in epidemiology that is defined as "the degree of statistical dependence between two or more events or variables."[23] Bradford Hill was clear that the criteria he described should only be examined once an association is established as "perfectly clear-cut" and "beyond what we would care to attribute to the play of chance."[24] The U.S. Surgeon General's Report on

characteristic and a disease; 2) The derivation of biological information from such a pattern of statistical associations.").

18. *Reference*, *supra* note 6, at 552.

19. *Id.* at 553.

20. 1 F.C. WOODSIDE, DRUG PRODUCT LIABILITY § 5.03[3] (2012) ("Confounding is a situation where a distortion of the effect of an exposure upon the risk of developing a condition or disease is created as a result of an association of the condition or disease with other factors than the exposure under study.").

21. *Bias*, in the scientific realm, is the systematic (nonrandom) error in a study that compromises its validity. *Reference*, *supra* note 6, at 583. For example, a selection bias can occur from an error in the method of selecting cases and controls. *Id.*; *see also In re* "Agent Orange" Prod. Liab. Litig., 597 F. Supp. 740, 783 (E.D.N.Y. 1985), *aff'd*, 818 F.2d 145 (2d Cir. 1987) (comparing the mortality rate of an exposed cohort of young, healthy men who had been in the military with a control group of civilians might have resulted in error that was a result of selection bias due to failure to account for health status as an independent variable).

22. *Reference*, *supra* note 6, at 552–53 ("Causation is used to describe the association between two events when one event is a necessary link in a chain of events that results in the effect. Of course, alternative causal chains may exist that do not include the agent but that result in the same effect. For general treatment of causation in tort law and that for factual causation to exist an agent must be a necessary link in a causal chain sufficient for the outcome.") (citing RESTATEMENT (THIRD) OF TORTS: LIABILITY FOR PHYSICAL HARM § 26 (2010)).

23. *In re* TMI Litig., 193 F.3d 613, 711 (3d Cir. 1999). Statistical significance can only be generated in a cohort, case control, or other epidemiological study. It cannot be calculated in case reports as they have no denominator.

24. *Id.*

Smoking and Health (on which the Bradford Hill criteria were based) bolsters this, mandating that one must first determine whether an association exists: only then can causal significance be examined.[25] In summary, before assessing any of the individual Bradford Hill criteria (such as strength of association), it must be determined whether there is in fact an association that is "perfectly clear-cut" and not likely the product of chance.

Determining whether an association exists is a seemingly straightforward issue. The U.S. Surgeon General described the process in its report on smoking and health:

> [E]vents are said to be associated when they occur more or less frequently together than one would expect by chance. . . . Events are said not to have an association when the agent (or independent variable) has no apparent effect on the incidence of a disease (the dependent variable).[26]

In modern epidemiology, this principle has been carried forward through determinations of statistical analyses, including the *relative risk*.[27] The relative risk measures how much more likely an exposed person is to contract the disease under consideration than is an unexposed person.[28] It is calculated by comparing the proportion of diseased persons in an exposed group to the proportion of diseased persons in an unexposed group.[29] As logically inferred, strong

---

25. *See* U.S. DEP'T OF HEALTH, EDUC., & WELFARE, *supra* note 5, at 19.

26. *Id.* To further complicate this analysis, it is not unusual for there to be multiple studies with conflicting results. While a statistically significant association (without suspect or impaired methodology) may be shown to exist in a given study, other studies may not support this conclusion. Therefore, while an association must be examined for a single, isolated study, it is also critical to examine the consistency of findings in different studies. In the presence of conflicting studies, it is difficult to see how an association in a single study would be "perfectly clear cut." Indeed, in pointing out that different studies frequently yield conflicting results, Taubes notes that the variation "seems almost constitutionally contradictory." Gary Taubes, *Epidemiology Faces Its Limits*, 269 SCIENCE 164, 164 (July 14, 1995), *available at* http://geography.ssc.uwo.ca/faculty/baxter/readings/Taubes_limits_epidemiology_Science_1995.pdf .269.

27. An *odds ratio* is another statistical analysis used. Odds ratio is defined as the "ratios of the odds of an adverse outcome, which reflect the relative likelihood of a particular result. . . . For example, if the chances of an outcome are 50% (one in two) with treatment and 33a% (one in three) without treatment, the odds ratio for the treated group would be 1/2 divided by a, or 1.5 (signifying a 50% greater chance of recovery in the treated group). The odds ratio is, therefore, a metric that provides insight only on relative benefit or relative risk." Samaan v. St. Joseph Hosp., 670 F.3d 21, 33 (1st Cir. 2012).

28. Melissa Moore Thompson, *Toxic Tort Litigation*, 71 N.C. L. REV. 247, 250–51 (1992).

29. *Id.*

associations are preferred because they are less likely to be due to errors such as hidden bias. Error alone, by contrast, may easily generate a weak association.[30]

An important corollary to relative risk is an examination of *confidence intervals*.[31] A confidence interval provides both the relative risk found in the study and a range (interval) within which the risk would likely fall if the study were repeated numerous times.[32] If the confidence interval includes the number one, the increased risk is not statistically significant.[33] For an association to be beyond the play of chance and "perfectly clear-cut," it must be statistically significant.[34] As such, risk measures used in conjunction with confidence intervals are critical in establishing a perfectly clear-cut association when it comes to examining the results of a single study.[35]

Additionally, before concluding that a valid association exists, it must be determined whether the association could have resulted from limitations of the particular study such as bias, confounding, or sampling error.[36] In other words, the quality of a study must also be

30. *Id.* at 269. The mere fact that an association is strong cannot create a valid study if it was poorly performed with biases, confounders, or other errors.

31. *Reference*, *supra* note 6, at 621 ("[*Confidence intervals* are] a range of values calculated from the results of a study within which the true value is likely to fall; the width of the interval reflects random error. Thus, if a confidence level of .95 is selected for a study, 95% of similar studies would result in the true relative risk falling within the confidence interval.").

32. *Id.* at 573.

33. *See* Kelley v. Am. Heyer-Schulte Corp., 957 F. Supp. 873, 878 (W.D. Tex. 1997) (stating that the lower end of the confidence interval must be above 1.0—equivalent to requiring that a study be statistically significant—before a study may be relied upon by an expert).

34. Miller v. Pfizer, Inc., 196 F. Supp. 2d 1062, 1080 (D. Kan. 2002).

35. It is also important to note the critical nature of statistical significance, as demonstrated by current case law, is at odds with Bradford Hill's analysis as he stated that tests of significance "contribute nothing to the 'proof' of our hypothesis." Hill, *supra* note 2, at 299.

36. There are many types of biases that may render an association invalid. One example is *selection bias*, which results from the method of selection of the cases and controls. *See Reference*, *supra* note 6, at 583. Another example is *information bias*, which is a result of inaccurate information about either the disease or the exposure status of the study participants or a result of confounding. *Id.* at 585. *Misclassification bias*, a final example, is a consequence of information bias in which, because of problems with the information available, individuals in the study may be misclassified with regard to exposure status or disease status. *Id.* at 589; *see also* D.L. Sackett, *Bias in Analytic Research*, 32 J. CHRONIC DISEASES 51, 51–63 (1979) (containing a more complete list of biases). *Sampling error*, also known as random error, is the probability the results are "due to chance" and "when the result obtained in the sample differs from the result that would be obtained if the entire

examined to determine if methodological error had the potential for producing a number that would erroneously suggest an association. The strength or validity of an association cannot be considered if errors in the study methodology render the results spurious. Only if an association is clearly established pursuant to the foregoing requirements, may an expert (who is considering utilizing the association in opining on the issue of causation) proceed to a consideration of the Bradford Hill criteria.[37]

C. Daubert *and Its Progeny*

Whether the conclusion of experts interpreting the data is admissible evidence hinges on the holdings of *Daubert v. Merrell Dow Pharmaceuticals, Inc.* and its progeny. In *Daubert*, the Supreme Court assigned to the trial judge a "gatekeeping responsibility" to make "a preliminary assessment of whether the reasoning or methodology underlying the testimony is scientifically valid and of whether that reasoning or methodology properly can be applied to the facts in issue."[38] The Court provided four nonexclusive factors that trial courts should consider in making this determination under Rule 702 of the Federal Rules of Evidence (governing testimony of expert witnesses).[39] First, the court must evaluate whether the theory or technique can be and has been tested.[40] Second, the court must determine whether the theory or technique has been subjected to peer review and publication.[41] Third, the court must consider the known or potential rate of error.[42] Finally, the court must evaluate the general acceptance of the theory in the scientific community.[43] Various courts' evaluation of the individual Bradford Hill criteria, in essence, constitutes a determination of their validity.

Since physicians, scientists, or other potential experts can generally accept a methodology without it being scientifically valid,

---

population (universe) were studied." *In re* TMI Litig., 193 F.3d 613, 708 (3d Cir. 1999).

37. *See* Hill, *supra* note 2, at 295; *see also* Soldo v. Sandoz Pharms. Corp., 244 F. Supp. 2d 434, 514 (W.D. Pa. 2003); RESTATEMENT (THIRD) OF TORTS: PHYSICAL & EMOTIONAL HARM § 28 (2010) ("If an association is found, epidemiologists use a number of factors (commonly known as the 'Hill guidelines') for evaluating whether that association is causal or spurious.").

38. Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579, 589–93 (1993).

39. FED. R. EVID. 702.

40. *Daubert*, 509 U.S. at 593.

41. *Id.*

42. *Id.* at 594.

43. Broekelschen v. Sec'y of HHS, 89 Fed. Cl. 336, 343 (2009).

the application of *Daubert* (and how the courts treat the Bradford Hill criteria) is crucial in determining the admissibility of testimony on the issue of causation. Case reports (and case series reports), which are not proof of causation, may lead experts to conclude there is causation when, in fact, case reports only generate hypotheses.[44]

Before discussing each of the nine criteria, it is informative to review a question Bradford Hill asked when he introduced them: "What aspects of that association should we especially consider before deciding that the most likely interpretation of it is causation?"[45] In posing his question, Hill himself is acknowledging that the mere presence of a valid association does not, without the consideration of additional factors, allow one to conclude that there is a causal relationship.

III. CAUSATION: THE NINE CRITERIA

As stated at the beginning of this article, when courts refer to the Bradford Hill criteria and when counsel cites to them, they generally do so with limited analysis. Understanding the specifics of the nine criteria and the law applicable to each provides a foundation for understanding the legal and factual issues involved in the admissibility of epidemiological evidence on the issue of causation. The nine criteria are discussed below in the order set forth in Hill's paper.[46]

---

44. McClain v. Metabolife Int'l, Inc., 401 F.3d 1233, 1253 (11th Cir. 2005); *Case Series*, NAT'L CANCER INST., http://www.cancer.gov/dictionary?CdrID=44006 (last visited April 3, 2013) ("[Case series are a] group or series of case reports involving patients who were given similar treatment. Reports of case series usually contain detailed information about the individual patients. This includes demographic information (for example, age, gender, ethnic origin) and information on diagnosis, treatment, response to treatment, and follow-up after treatment.").

45. Hill, *supra* note 2, at 295. In delineating the "aspects" that "we should especially consider," Hill implies that they are a subset of a larger group of aspects that should be considered before deciding the causation issue. Hill never identifies the totality of the aspects to be considered—only those to be "especially considered."

46. Notably, although Bradford Hill listed nine criteria in his presentation, he did not disclose therein (1) the methodology he used in selecting these nine criteria; (2) any ranking or hierarchy of the nine criteria vis-à-vis their relative importance in a consideration of causation; (3) the number of the criteria (or which ones) that had to be satisfied before one could conclude that causation had been established; (4) the basis for concluding that one or more of the criteria need not be satisfied; (5) the consideration to be given to a finding—relating to one or more of the criteria—that would tend to refute a causal relationship; or (6) the process to be utilized in fusing into an opinion the information gleaned from an evaluation of the nine criteria.

A. Strength of Association

Bradford Hill listed *strength of association* as the first criteria to be considered once the predicate of having a perfectly clear-cut association is established.[47] In examining the strength of the association, Hill considered the magnitude of the increased risk of disease that the epidemiological study shows. Whether an epidemiological study yields an increased risk that is statistically significant (which would indicate an association), however, is a separate and distinct analysis from consideration of the implications of the increased risk's magnitude. Statistical significance deals not with the magnitude of any increase in the exposed group's relative risk, but rather with the comfort or confidence one has in the results of the study. More specifically, it is the confidence one has in the point estimate (relative risk) as a summary statistic of the underlying data.[48]

When discussing strength of association, Hill intimated that with regard to increased risks that are found to be statistically significant, the larger the increase in the relative risk, the more likely it is that a causal relationship exists.[49] As an example, he noted that the death rate in smokers from lung cancer was nine to ten times that seen in nonsmokers; whereas the death rate from coronary thrombosis in smokers was no more than twice that seen in nonsmokers.[50] Hill concluded that the prevalence of confounding factors (lack of exercise, nature of diet, etc.) more readily explained the increase in the incidence of coronary thrombosis than the increased incidence of lung cancer in smokers (a population in which a confounding factor should have been more easily detectable).[51]

Considering the strength-of-association issue, various courts (performing their *Daubert* gatekeeping function) have considered whether there must be a certain increase in the relative risk for a causation opinion based thereon to be admissible. For example, in *Allison v. McGhan Medical Corp.*, the court found a prospective study with a relative risk ratio of only 1:24 had a finding so close to

47. *See* Hill, *supra* note 2, at 295.

48. *Point estimate* is also known as relative risk. *Reference*, *supra* note 6, at 621.

49. *Id.* at 295–96.

50. *Id.* at 296.

51. *Id.* Bradford Hill is actually referencing the concept of *attributable risk.* Attributable risk describes the "proportion of disease in exposed individuals that can be attributed to exposure to an agent, as distinguished from the proportion of disease attributed to all other causes." *Reference*, *supra* note 6, at 619.

1.0 that the risk was too minimal for assessing causation.[52] As the court aptly noted, "showing association is far removed from proving causation."[53] It further stated that the threshold for concluding that an agent more likely than not caused a disease is 2.0,[54] and that a 2.0 relative risk implies a fifty percent likelihood (also known as an attributable risk)[55] that the agent caused the disease.[56] The court did not hold, however, that a relative risk of more than 2.0 is a litmus test, or that a single epidemiological test is legally sufficient evidence of causation.[57] Indeed, some scholars suggest that a relative risk of 3.0 is needed to demonstrate causation.[58] The fact that the legal community has held that a single study is not legally sufficient demonstrates the necessity of conducting more than one study to replicate the initial results.

B. Consistency

The next criterion to be considered is *consistency*, which Bradford Hill phrased as a question: "Has [the association] been repeatedly observed by different persons, in different places, circumstances and times?"[59] As an example, Hill again examined the link between lung cancer and smoking. In that case—where the association was established in twenty-nine retrospective and seven prospective studies, using a variety of techniques in a variety of situations—Hill stated, "we can justifiably infer that the association is not due to some constant error or fallacy that permeates every inquiry. And we have indeed to be on our guard against that." [60]

52. Allison v. McGhan Med. Corp., 184 F.3d 1300, 1315 (11th Cir. 1999).

53. *Id.*

54. A level above 1.0 can be probative of general causation or statistical significance; however, can only be probative of specific causation if the relative risk is greater than 2.0. Henricksen v. Conoco Phillips Co., 605 F. Supp. 2d 1142, 1158 (E.D. Wash. 2009).

55. As a relative risk of 2.0 constitutes an attributable risk of fifty percent, an attributable risk of greater than fifty percent is required to admit testimony on the issue of causation. *Id.*

56. *Id.*; *see also* Merrell Dow Pharms., Inc. v. Havner, 953 S.W.2d 706, 718 (Tex. 1997) ("The use of scientifically reliable epidemiological studies and the requirement of more than a doubling of the risk strikes a balance between the needs of our legal system and the limits of science.").

57. *Merrell*, 953 S.W.2d at 718.

58. *See* Taubes, *supra* note 26, at 268.

59. Hill, *supra* note 2, at 296.

60. *Id.*

Consistent is defined as "marked by harmony, regularity, or steady continuity: free from variation or contradiction."[61] The rationale for this consideration is that consistent findings, observed by different persons in different places, and utilizing different samples, increase the likelihood that a causal effect can be inferred. Bradford Hill's consistency requirement thus mirrors the fourth *Daubert* factor: general acceptance in the scientific community.

In the field of epidemiology, the first study that yields positive results is considered a hypothesis-generating study.[62] One study is an insufficient indicator of causation.[63] Following the first study, confirmatory studies are required to meet the "generally accepted methodology" requirement set forth in *Daubert*.[64]

Reduced to an elementary level, consistency demonstrates that the results of a particular study are not an outlier result. Consistency indicates that the results are generally concurrent with the results of other studies—not that they are generally accepted. For instance, multiple studies with low power (the study is not large enough to detect associations with rare diseases or adverse effects) may have consistent findings, but general acceptance is withheld until a more powerful, randomized study is done yielding the same results.

Furthermore, neglecting to discredit contrary research constitutes a failure to utilize accepted methodology as required by *Daubert*.[65] Indeed, the presence of contradictory research negates consistency

61. *See Consistent*, MERRIAM-WEBSTER.COM, http://www.merriam-webster.com/dictionary/consistent (last visited Jan. 20, 2013).

62. HARVEY CHECKOWAY, NEIL PEARCE & DAVID KRIEBEL, RESEARCH METHODS IN OCCUPATIONAL EPIDEMIOLOGY 10 (2d ed. 2004). However, the first study, if significantly positive and a cohort or case/control study is the first test of an association. It is not an hypothesis-generating study unless the hypotheses had not been previously generated and this was an incidental finding. The hierarchy is that case reports and case series reports are hypothesis-generating reports. The cohort studies or case/control studies that follow are hypothesis-testing studies—which may support the hypothesis, refute the hypothesis or (for a number of reasons) be indeterminate.

63. *See* Merrell Dow Pharms., Inc. v. Havner, 953 S.W.2d 706, 727 (Tex. 1997) ("As we have already observed, an isolated study finding a statistically significant association between Bendectin and limb reduction defects would not be legally sufficient evidence of causation.").

64. *See* Miller v. Pfizer, Inc., 196 F. Supp. 2d 1062, 1081 (D. Kan. 2002) ("The Court therefore concludes that in failing to discuss the consistency of his hypothesis with other research, Dr. Healy has not used generally accepted methodology.").

65. *Id.*

and underscores the need for further studies.[66] Criticism of the contrary study's methodology, however, does not convert it to a study supporting an association.

C. Specificity of the Association

The third criterion is described as the *specificity of the association*.[67] Bradford Hill stated that "if specificity exists[,] we may be able to draw conclusions without hesitation; if it is not apparent, we are not thereby necessarily left sitting irresolutely on the fence."[68] The crux of the specificity consideration is that causation is likely if a very specific population at a specific site develops a disease with no other likely explanation.[69] More specifically, well performed studies demonstrating an association between a specific exposure and a clearly defined disease or condition—otherwise known as the *case definition*—are of more value in inferring the existence of a causal relationship than studies with poorly defined exposures and/or loosely defined diseases or conditions.[70]

A review of the case law reveals that specificity of association requires a consistent case definition (of the specific condition being investigated) when initially performing an epidemiology study.[71] Vagueness in a case definition makes it impossible to gather data that would permit a meaningful epidemiological study.[72] Also, without a specific case definition, the theory cannot be refuted.[73]

---

66. Conde v. Velsicol Chem. Corp., 24 F.3d 809, 814 (6th Cir. 1994) ("[T]he critiques [of contradictory studies] only underscore the need for further studies, and do not, as the district court noted, establish causation.").

67. Hill, *supra* note 2, at 297.

68. *Id.*

69. *Id.*

70. *Specificity* means that the exposure-outcome relationship is to a specific (narrowly defined) outcome. That is, specificity does not exist if a particular exposure leads to one specific cancer in one study and to a different specific cancer in another study.

71. *See In re* Denture Cream Prods. Liab. Litig., 795 F. Supp. 2d 1345, 1361 (S.D. Fla. 2011) (stating that inconsistencies in a case definition limit the evidentiary value of the studies and that lack of a case definition is a deficiency); *see also* Young v. Burton, 567 F. Supp. 2d 121, 131 (D.D.C. 2008) (holding that a case definition for mold illness was unreliable).

72. Joseph Sanders & D.H. Kaye, *Expert Advice on Silicone Implants: Hall v. Baxter Healthcare Corp.*, 37 JURIMETRICS J. 113, 120 (1997) ("The features of a unique connective-tissue syndrome have not been put into a coherent, valid, or reproducible case definition, which severely limits scientific study." (citing Matthew H. Liang et al., *Letter to the Editor*, 333 NEW ENG. J. MED. 1424, 1424 (1995))).

73. *Id.* Every epidemiological study is done to disprove a hypothesis. *See* ROTHMAN ET AL., *supra* note 5, at 25 ("[E]pidemiologists usually focus on testing the

Case law further reveals that specific exposures cause limited (or specific) conditions.[74] When advocates attempt to expand the results of a study by providing evidence of a causal relationship between an agent and a specific disease to prove causation between the agent and a different disease or condition, courts have required the proponent of such evidence to demonstrate that the biological mechanism (whereby the exposure causes the disease or condition) is the same.[75] Likewise, when trying to extrapolate from animal data to human data, grounds for extrapolation must be demonstrated.[76]

Extrapolation between animal studies to humans is problematic in court.[77] Differences in absorption, metabolism, and other factors may result in interspecies variation in responses.[78] For example, in *Brock v. Merrell Dow Pharmaceuticals, Inc.*, the court noted the "very limited usefulness of animal studies when confronted with questions of toxicity."[79] Conversely, some courts have found animal

---

negation of the causal hypothesis, that is, the null hypothesis that the exposure does not have a causal relation to disease. Then, any observed association can potentially refute the hypothesis, subject to the assumption (auxiliary hypothesis) that biases and chance fluctuations are not solely responsible for the observation.").

74. *See, e.g.*, Nelson v. Am. Sterilizer Co., 566 N.W.2d 671, 676–77 (Mich. Ct. App. 1997) (affirming dismissal of plaintiff's claims that chemical exposure caused her liver disorder, but recognizing that evidence supported claims for neuropathy and other illnesses); *see also Young*, 567 F. Supp. 2d at 138 ("[There is a] need to identify specific toxins and connect them to specific symptoms."); Rider v. Sandoz Pharm. Corp., 295 F.3d 1194, 1201 (11th Cir. 2002) ("[E]ven minor deviations in chemical structure can radically change a particular substance's properties and propensities.").

75. *See, e.g.*, Austin v. Kerr-McGee Refining Corp., 25 S.W.3d 280 (Tex. App. 2000) (indicating where studies demonstrated a causal relationship between benzene and all leukemias, but there was no evidence on the relationship between benzene and the specific form of leukemia from which plaintiff suffered, the court required that plaintiff's expert demonstrate the similarity of the biological mechanism among leukemias); *see also* Hollander v. Sandoz Pharms. Corp., 95 F. Supp. 2d 1230, 1238 (W.D. Okla. 2000) ("Causation also cannot be shown by the fact that other ergot alkaloids, which are in the same class as bromocriptine, cause hypertension. The plaintiffs have failed to demonstrate that bromocriptine and the other ergots have sufficiently similar physiological effects to warrant comparison.").

76. *See supra* Part III.A.

77. Soldo v. Sandoz Pharms. Corp., 244 F. Supp. 2d 434, 466 (holding that the district court did not abuse its discretion in excluding expert testimony on causation based on expert's failure to explain how animal studies supported expert's opinion that agent caused disease in humans).

78. *Reference*, *supra* note 6, at 563.

79. Brock v. Merrell Dow Pharms., Inc., 874 F.2d 307, 313 (5th Cir. 1989); *see also Soldo*, 244 F. Supp. 2d at 466 ("The use of animal studies . . . 'are almost always fraught with considerable, and currently unresolvable, uncertainty.'").

studies to be reliable evidence.[80] However, "in order for animal studies to be admissible to prove causation in humans, there must be good grounds to extrapolate from animals to humans, just as the methodology of the studies must constitute good grounds to reach conclusions about the animals themselves."[81]

D. Temporality

Temporality, according to Hill, raises the question: "[W]hich is the cart and which the horse?"[82] In epidemiology, the temporality requirement can be simply described as requiring that the exposure occur prior to development of the disease.[83] Bradford Hill claimed that "[n]one of [his] nine viewpoints can bring indisputable evidence for or against the cause-and-effect hypothesis and none can be required as a sine qua non."[84] With regard to temporality, however, case law demonstrates the inaccuracy of this statement.[85] For example, *Carroll v. Litton Systems, Inc.* involved claims that chemicals from a plant entered certain homes and workplaces through groundwater, causing illness to occupants.[86] The court, when discussing expert testimony on the timing of when the chemicals entered the groundwater, stated that "it is essential for . . . [the plaintiffs' medical experts opining on causation] to know that exposure preceded plaintiffs' alleged symptoms in order for the exposure to be considered as a possible cause of those symptoms . . . ."[87]

Indeed, temporality is necessary in determining causation; however, establishing temporality will not prove causation. In

80. *In re* Heparin Prods. Liab. Litig., 2011 WL 2971918 (N.D. Ohio July 21, 2011) (holding that animal toxicology in conjunction with other non-epidemiologic evidence can be sufficient to prove causation); Ruff v. Ensign-Bickford Indus., Inc., 168 F. Supp. 2d 1271, 1281 (D. Utah 2001) (affirming animal studies as a sufficient basis for opinion on general causation).

81. *In re* Paoli R.R. Yard PCB Litig., 35 F.3d 717, 743 (3d Cir. 1994).

82. *See* Hill, *supra* note 2, at 297.

83. *See Reference*, *supra* note 6, at 601.

84. *See* Hill, *supra* note 2, at 299; *see also Sine qua non*, MERRIAM-WEBSTER.COM, http://www.merriam-webster.com/dictionary/ sine%20qua%20non (last visited Feb. 21, 2013) ("[Sine qua non is] something absolutely indispensible or essential.").

85. Henricksen v. ConocoPhillips Co., 605 F. Supp. 2d 1142, 1156 (E.D. Wash. 2009) ("[T]he chronological relationship between exposure and effect must be biologically plausible.").

86. Carroll v. Litton Sys., Inc., No. B-C-88-253, 1990 U.S. Dist. LEXIS 16833 (W.D.N.C. 1990).

87. *Id*. at *29.

McClain v. Metabolife International, Inc., the Eleventh Circuit explained:

> [P]roving a temporal relationship . . . does not establish a causal relationship. In other words, simply because a person takes drugs and then suffers an injury does not show causation. Drawing such a conclusion from temporal relationships leads to the blunder of the post hoc ergo propter hoc fallacy. The post hoc ergo propter hoc fallacy assumes causality from temporal sequence. It literally means "after this, because of this."[88]

Not only must the exposure precede the development of the alleged symptoms, but the period of time between the alleged exposure and the onset of symptoms for which compensation is sought must be consistent with the known *latency period* for the exposure in question.[89] The latency period is the period of time between exposure to an agent and manifestation of disease symptoms.[90] Knowing the latency period between exposure and outcome is critical when considering temporality.[91] An exposure and outcome temporally consistent with the known latency period can support a causal relationship.[92] Conversely, exposure outside a known latency period is evidence—perhaps conclusive evidence—against the existence of causation.[93]

---

88. *See* McClain v. Metabolife Int'l, Inc., 401 F.3d 1233, 1243 (11th Cir. 2005).

89. *Research*, *supra* note 6, at 601.

90. *See In re* TMI Litig., 193 F.3d 613, 643 (3d Cir. 1999); *see also* Robinson v. McNeil Consumer Healthcare, 615 F.3d 861, 866 (7th Cir. 2010) ("[Latency period is the] interval between the infection or other trauma and when the first symptoms appear.").

91. *In re TMI Litig.*, 193 F.3d at 643 ("[T]he ability to make a plausible association between the accident and a diagnosis . . . depends upon the length of the latency period.").

92. *See* Bonner v. ISP Techs., Inc., 259 F.3d 924, 930–31 (8th Cir. 2001) (explaining expert's opinion on causation was supported more for acute response based on temporal relationship than for chronic disease that plaintiff also developed); Alder v. Bayer Corp., 61 P.3d 1068, 1090 (Utah 2002) ("If a bicyclist falls and breaks his arm, causation is assumed without argument because of the temporal relationship between the accident and the injury [and, the court might have added, the absence of any plausible competing causes that might instead be responsible for the broken arm].").

93. *See In re* Phenylpropanolamine (PPA) Prods. Liab. Litig., 289 F. Supp. 2d 1230, 1238 (W.D. Wash. 2003) (explaining expert testimony on causation for injuries occurring more than three days (the known latency period) after ingestion of PPA was inadmissible); *see also* Burleson v. Glass, 268 F. Supp. 2d 699, 707 (W.D. Tex. 2003) (granting defendant's motion to exclude plaintiff's expert testimony in part because two-year latency period from alleged exposure to onset of cancer was unusually short given the scientific literature indicating typical latency of ten to fifteen years for the tumor type); Nat'l Bank of Commerce v. Associated Milk

E. Biological Gradient

For the fifth item, Bradford Hill stated, "if the association is one which can reveal a biological gradient, or dose-response curve, then we should look most carefully for such evidence."[94] As an example, he noted that the fact the death rate from cancer rises linearly with the number of cigarettes smoked daily "adds a great deal to the simpler evidence that cigarette smokers have a higher death rate than nonsmokers."[95] He recognized the occasional difficulty in securing dose-response evidence, but stated that "we should invariably seek it."[96]

Bradford Hill was not the first to cite to this concept. In the 1500s, Auroleus Phillipus Theophrastus Bombastus von Hohenheim (known to the world as Paracelsus) stated that "[a]ll substances are poisonous—there is none which is not; the dose differentiates a poison from a remedy."[97] Substances thought to be poisonous may not be poisonous at low doses; conversely, substances that are generally regarded as safe (like water and oxygen) can be poisonous or toxic in high doses.[98]

Case law delineates the necessity of plaintiffs establishing not only the dose-response relationship (which is generally shown graphically as a dose-response curve), but also the level of exposure required to produce the disease or symptoms in dispute.[99] The point on the dose-response curve above which further exposure causes disease or symptoms is referred to as the *threshold*.[100] Initially, for

---

Producers, 22 F. Supp. 2d 942, 975 (E.D. Ark. 1998) (excluding expert testimony because an unusually short latency period "creates one more negative for the plaintiffs").

94. *See* Hill, *supra* note 2, at 298. *Dose-response curve* is defined as "A graphic representation of the relationship between the dose of a chemical administered and the effect produced." *Reference*, *supra* note 6, at 681.

95. *Dose-response relationship* is defined as the "characteristics of exposure and the spectrum of toxic effects [that] come together in a correlative relationship." CASARETT AND DOULL'S TOXICOLOGY 19 (Curtis D. Klaassen, Ph.D., 7th ed. 2008). Moreover "the relationship between the degree of response of the biological system and the amount of toxicant administered assumes a form that occurs so consistently as to be considered the most fundamental and pervasive concept in toxicology." *Id.*

96. Hill, *supra* note 2, at 298.

97. *In re* Denture Cream Prods. Liab. Litig., 795 F. Supp. 2d 1345, 1351 (S.D. Fla. 2011).

98. *Id.* at 1352.

99. *See, e.g.*, *In re Denture Cream*, 795 F. Supp. at 1352; Newman v. Motorola, Inc., 78 Fed. Appx. 292, 294 (4th Cir. 2003).

100. The threshold level means that an exposure may not cause disease until the exposure exceeds a certain dose. *See Reference*, *supra* note 6, at 603.

expert testimony to be admissible, it must demonstrate that a certain threshold level of exposure has been exceeded.[101]

Demonstration of a dose-response relationship is critical in establishing causation.[102] Conversely, a lack of an attempt to show a dose-response can reflect a defect in methodology.[103] As such, the case law generally requires proof of a dose-response relationship in order to demonstrate causation.[104]

The shape of the dose-response curve can vary with the substance under consideration.[105] The classic shape is "S"-shaped.[106] On the other hand, essential metals have "U"-shaped curves because toxicity can result from exposure to either insufficient levels or excess levels of the essential metals.[107] Other substances that are necessary for survival, such as water, may also become poisonous at high enough levels.[108]

---

101. *In re Denture Cream*, 795 F. Supp. 2d at 1352 ("'[F]or most types of dose-response relationships following chronic (repeated) exposure, thresholds exist, such that there is some dose below which even repeated, long-term exposure would not cause an effect in any individual.' Often 'low dose exposures—even for many years—will have no consequence at all, since the body is often able to completely detoxify low doses before they do any damage.'"); Mitchell v. Gencorp, Inc., 165 F.3d 778, 781 (10th Cir. 1999) ("It is well-established that a plaintiff in a toxic tort case must prove that he or she was exposed to and injured by a harmful substance manufactured by the defendant. . . . In order to carry this burden, a plaintiff must demonstrate the levels of exposure that are hazardous to human beings generally as well as the plaintiff's actual level of exposure to the defendant's toxic substance before he or she may recover."); *see also* Johnston v. United States, 597 F. Supp. 374, 393 (D. Kan. 1984) (stating that standard toxicological science assumes that there is a threshold for all harmful chemicals).

102. *See Newman*, 78 Fed. Appx. at 294 ("Showing a dose-response relationship is . . . an important factor in establishing causation.").

103. *See* McClain v. Metabolife Int'l, Inc., 401 F.3d 1233, 1241 (11th Cir. 2005); *see also Mitchell*, 165 F.3d at 781.

104. *In re Denture Cream*, 795 F. Supp. 2d at 1351 (S.D. Fla. 2011).

105. Eli Lilly & Co. v. Teva Pharms. U.S., Inc., 657 F. Supp. 2d 967, 991 (S.D. Ind. 2009).

106. *Id*. at 991 n.19 ("Not all drugs have the same dose response curve. The typical dose response curve has an 'S' shape when depicted on a graph plotting dose of the drug versus response. At very low doses, most drugs do not provide a significant response. As more and more of a drug is administered, the response grows. At some point, the response reaches an upper plateau at which the response no longer increases, regardless of how much of the drug is administered.").

107. An essential nutrient can be defined as one whose absence from the diet will lead to growth impairment, organ dysfunction, or failure to maintain nitrogen balance on an adequate intake of all other nutrients. *See* George K. Grimble, *Essential and Conditionally Essential Nutrients in Clinical Nutrition*, 6 NUTRITION RES. REVS. 1, 97 (1997); *see also* HANDBOOK ON THE TOXICOLOGY OF METALS 107–08 (Gunnar F. Nordberg et al. eds., 3rd ed. 2007).

108. *In re Denture Cream*, 795 F. Supp. 2d at 1352 (S.D. Fla. 2011).

One should not conclude from this analysis, however, that to pass *Daubert* muster an expert must give precise numbers about a dose-response relationship. Indeed, case law demonstrates that some ambiguity about individual responses is expected. Nevertheless, the link between an expert's opinions and the dose-response relationship is a key element of reliability in toxic tort cases.[109]

F. Plausibility

As his sixth "feature," Bradford Hill stated that "[i]t will be helpful if the causation we suspect is biologically plausible. But this is a feature I am convinced we cannot demand. What [is] biologically plausible depends upon the biological knowledge of the day."[110] This is a challenging point because of the subjectivity of this feature and its dependence on the knowledge existent at the time of the inquiry, which could certainly be later proven to be incorrect. Some have interpreted this feature to indicate that a relationship predicted prospectively is much more convincing than one provided retrospectively because after observing an association, it is often easy to give a plausible explanation.[111]

In examining the plausibility criterion, it is important to note the potential fallibility of current knowledge. What is known at a particular point in time may be built upon a foundation of incorrect, inaccurate, or erroneous science: thus limiting its affirming potential.

Furthermore, in dealing with the role of plausibility, it is important to provide evidence of the mechanism supporting the biologic plausibility. As one court explained, "a biological explanation without evidence of the mechanism by which it works is merely an unproven hypothesis, a theory."[112]

G. Coherence

Bradford Hill introduced *coherence* as follows: "cause-and-effect interpretation of our data should not seriously conflict with the generally known facts of the natural history and biology of the

109. McClain v. Metabolife Int'l, Inc., 401 F.3d 1233, 1241 (11th Cir. 2005).

110. *See* Hill, *supra* note 2, at 298.

111. *See* Michael Höffler, *The Bradford Hill Considerations on Causality: A Counterfactual Perspective*, 2 EMERGING THEMES EPIDEMIOLOGY 11 (2005) (citing D. R. COX & NANNY WERMUTH, MULTIVARIATE DEPENDENCIES: MODELS, ANALYSES AND INTERPRETATION (1996)).

112. *In re* Accutane Prods. Liab., 511 F. Supp. 2d 1288, 1295 (M.D. Fla. 2007).

disease . . . ."[113] The difference between coherence and plausibility would seem, in part, to be one of semantics. While plausibility is worded positively (an association should be in line with substantive knowledge), coherence is presented negatively (an association should not seriously conflict with substantive knowledge).[114] Consideration of coherence would reject an observed result as *non-causal* if it contradicted a predominant theory; while plausibility leaves the researcher more room regarding which particular piece of substantive knowledge to evaluate the results against.[115]

As with plausibility, when examining coherence it is important to note the potential fallibility of current knowledge (e.g., that it is incorrect, inaccurate, or erroneous). Additionally, this criterion is related to consistency. If a proposed relationship is in conflict with generally established scientific data, it would not only be incoherent, but it would also be inconsistent.

H. Experiment

To Bradford Hill, *experiment* meant the evidence obtained from reducing or eliminating a putatively harmful exposure and ascertaining if the frequency of the disease subsequently declines.[116] Hill thought of this as the strongest evidence of causality that could be obtained. The most notable type of experimental evidence supporting this viewpoint is cigarette smoking and lung cancer.[117] He recognized, however, the limitations of the experimental evidence consideration:

> It can be faulty . . . as the "semi experimental" approach is nothing more than a "before and after" time trend analysis which can be confounded or otherwise biased by a host of concomitant secular changes. Moreover, even if the removal of exposure does causally reduce the frequency of disease, it might not be for the etiological reason hypothesized.[118]

113. *See* Hill, *supra* note 2, at 298.

114. *See* Höffler, *supra* note 111.

115. *Id.*

116. ROTHMAN ET AL., *supra* note 5, at 29.

117. *Reference*, *supra* note 6, at 605.

118. *Id.* The reticence to use this type of evidence can also be seen in *In re* Denture Cream Prods. Liab. Litig., 795 F. Supp. 2d 1345, 1364 (S.D. Fla. 2011). As that case noted, evidence of removal of exposure to a hypothesized agent, in this case denture cream (also known as a "de-challenge"), in order to show a change in outcome was not reliable enough evidence of causation. There has to be a "re-challenge" thereafter to see if the disease or condition again develops. *Id.*

From a scientific standpoint, it is unfortunate that this type of evidence is generally not available. When an agent's effects are suspected to be harmful, researchers cannot knowingly expose people to the agent.[119] It is difficult to design these types of studies due to the ethical implications of experimentation on humans.

Instead, any experimental evidence used to support a causal relationship (which would generally not involve studies on humans) should be derived utilizing the *scientific method*.[120] The scientific method is "a method of research in which a problem is identified, relevant data are gathered, a hypothesis is formulated from these data, and the hypothesis is empirically tested."[121] Key aspects of the scientific method include the ability (1) to test or verify a scientific experiment by parallel experiment or other standard of comparison (control) and (2) to replicate the experiment to expose or reduce error.[122]

I. Analogy

Bradford Hill, in his ninth "viewpoint," states that "[i]n some circumstances it would be fair to judge by analogy."[123] As an example, he states that "with the effects of thalidomide and rubella before us we would surely be ready to accept slighter but similar evidence with another drug or another viral disease in pregnancy."[124] The reference to thalidomide and rubella is undoubtedly based on their ability to produce birth defects.[125] In essence, it is the position of Hill that since these two exposures have been proven to cause birth defects, it is plausible that another drug or viral disease may cause birth defects.

Recent case law has cast caution upon the extent to which evidence of analogy may be considered in developing opinions on causation. Courts have warned that a reliable methodology must still be utilized in drawing analogies.[126] A reliable methodology includes

119. Glastetter v. Novartis Pharms. Corp., 252 F.3d 986, 992 (8th Cir. 2001).

120. *See* Trach v. Fellin, 817 A.2d 1102 (Pa. Super. Ct. 2003).

121. *Id.* at 1113 (internal citations omitted).

122. *Id.*

123. *See* Hill, *supra* note 2, at 299.

124. *Id.*

125. DANIEL BERGSMA, NATIONAL FOUNDATION, BIRTH DEFECTS COMPENDIUM 435–38 (2d ed. 1979).

126. *See* McClain v. Metabolife Int'l, Inc., 401 F.3d 1233, 1241 (11th Cir. 2005) (finding testimony inadmissible where an expert drew an analogy between ephedrine

showing that the differences in chemical structure make no difference in effect.[127] As *McClain* explained, "even small differences in chemical structure can sometimes make very large differences in the type of toxic response that is produced."[128]

Thus, when drawing an analogy, all the other rules relating to acceptable scientific methodology must be employed. In other words, an analogy may support the testing of a hypothesis. As Judge Posner explained: "[T]he courtroom is not the place for scientific guesswork, even of the inspired sort. Law lags science; it does not lead it."[129]

IV. CONCLUSION

If the Bradford Hill criteria are to be considered vis-à-vis the admissibility of epidemiologic evidence on the issue of causation, several points must be taken into consideration. First, the predicate must be fulfilled before any consideration of the criteria can be undertaken—to wit, an association must first have been established that is "perfectly clear-cut" and "beyond what we would attribute to the play of chance."[130] This mandatory predicate, which was established by Hill himself, must not be forgotten. To consider the criteria without the predicate first having been demonstrated would yield results that are meaningless—if not clearly erroneous. Second, assessment of the individual criteria must consider the applicable case law. Third, as there is no established methodology for conflating the results of a consideration of the multiple criteria into an expert opinion, the rationale and methodology for doing so must take into account the rulings of *Daubert* and its progeny.

---

or PPA, but did not show the reliability of each of his steps, which was a "fatal defect under Daubert").

127. *Id.*

128. *Id.* (internal citations omitted); *see also* Rider v. Sandoz Pharms. Corp., 295 F.3d 1194, 1201 (11th Cir. 2002) ("[E]ven minor deviations in chemical structure can radically change a particular substance's properties and propensities.").

129. Rosen v. Ciba-Geigy Corp., 78 F.3d 316, 319 (7th Cir. 1996).

130. Hill, *supra* note 2, at 295.