EPIDEMIOLOGY FOR LAWYERS

TAA Annual Meeting – Chicago, Illinois

June 7 - 9, 2017

Chad R. Hutchinson
Butler Snow LLP
1020 Highland Colony Parkway, Suite 1400
Ridgeland, MS 39157

INTRODUCTION

In 1854, a deadly cholera outbreak struck London. The prevailing belief at the time was that cholera was spread by "miasma"—a poisonous form of "bad air" arising from decomposing matter on the ground. But not everyone agreed. Dr. John Snow, an English physician, believed that cholera could be spread by contaminated water. To prove his theory, Dr. Snow mapped the locations of deaths from cholera and observed that the majority of deaths occurred near a water pump on Broad Street in Soho. He then used his findings to convince local officials to remove the handle of the pump, which was apparently contaminated, and helped end the epidemic. Though the worst of the outbreak had passed by the time the Broad Street pump was disabled, Dr. Snow's research was foundational to the ultimate elimination of cholera in London. He is now considered the father of modern epidemiology.[1]

WHAT IS EPIDEMIOLOGY?

"Epidemiology is the field of public health and medicine that studies the incidence, distribution, and etiology of disease in human populations."[2] Its purpose is to better understand what causes disease and how disease can be prevented.[3]

Epidemiology is relevant to questions of general causation, i.e. whether an agent can cause a particular disease or health outcome. For lawyers, epidemiology may also be relevant to prove the safety and efficacy of a client's product, to explain a client's actions (e.g. the reason for a labeling change on a prescription drug), to establish alternative explanations of a plaintiff's alleged injury, or to mount a successful Daubert/Frey challenge against an expert.

But it has limitations. Epidemiology does not answer questions about specific causation, i.e. whether an agent caused a *particular* individual's disease or health outcome.[4] In other words, epidemiology addresses whether an agent *can* cause a disease, not whether an agent *did* cause a disease. And epidemiology only permits an inference that general causation exists. Thus, standing alone, epidemiological studies are generally insufficient to prove causation.

While there is no absolute requirement that epidemiological evidence be used to establish causation,[5] relevant studies may be admitted as government records or learned treatises under hearsay exceptions Fed. R. Evid. 803(8) or 803(18), respectively, or as the basis for an expert's opinion under Fed. R. Evid. 703.[6]

---

[1] Peter Daniel et al., *Dr. John Snow and Reverend Whitehead*, CHOLERA AND THE THAMES, http://www.choleraandthethames.co.uk/cholera-in-london/cholera-in-soho/ (last visited Nov. 30, 2016).

[2] Michael D. Green et al, *Reference Guide on Epidemiology, in* REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 549, 551 (3d ed. 2011) ("Reference Guide on Epidemiology").

[3] *Id.* at *551.*

[4] *Id*. at 608-09.

[5] *See In re Meridia Prod. Liab. Litig.,* 328 F. Supp. 2d 791, 800 (N.D. Ohio 2004) ("No requirement exists that a party *must* offer epidemiological evidence to establish causation."), *aff'd sub nom. Meridia Prod. Liab. Litig. v. Abbott Labs.,* 447 F.3d 861 (6th Cir. 2006);*see also In re Phenylpropanolamine (PPA) Prods. Liab. Litigation*, 289 F. Supp. 2d 1230, 1242 (W.D. Wash. 2003) ("Non-epidemiological sources are frequently utilized by experts . . . .").

[6] *See, e.g., Ellis v. International Playtex, Inc*., 745 F.2d 292, 300-01 (4th Cir. 1984) (admitting epidemiological studies of toxic shock syndrome under Rule 803(3).; *see also In re Agent Orange Prod. Liab. Litig.*, 611 F. Supp. 1223, 1241 (E.D.N.Y. 1985) (noting that epidemiological studies may be admitted as "learned articles" or as the basis for an expert opinion).

CATEGORIES OF EPIDEMIOLOGIC STUDIES

There are two categories of epidemiologic studies: experimental and observational. In an experimental study, scientists assign subjects to one of two groups: (1) a group that will be exposed to an agent, and (2) a group that will not be exposed to an agent. Scientists later evaluate and compare both groups for development of the disease.

Experimental studies include randomized controlled trials and are often double-blinded or placebo-controlled, making them the "gold standard" of epidemiological evidence. Subjects of the studies may include animals or humans. However, experimental studies in humans are ethically prohibited when an agent is known to be potentially harmful. As a result, most epidemiologic studies for humans are observational.[7]

There are two main types of observational studies: cohort and case-control. Both types of studies have a comparison group and determine if there is an association between exposure to an agent and a disease. If an association is present, the study then measures the strength of that association. The studies differ, however, in their use of exposure or disease as the independent variable.

1. Cohort Studies

Cohort studies, generally known as "prospective studies," begin with a defined group who have been *exposed* and then look *forward* to compare their experiences (i.e. development of a disease) with another group. In a cohort study, scientists select two groups: (1) participants who have been exposed to a particular agent and (2) participants who have not been exposed to the agent. Scientists then follow the groups and compare their development of the disease.[8]

For example, suppose a scientist wanted to study whether exposure to elevators causes the flu. To conduct the study, the scientist would select two equally-sized groups: (1) people who ride the elevator (exposed group) and (2) people who do not ride the elevator (control group). The scientist would then follow the groups and compare how many people in each group contract the flu.

Cohort studies are often used in occupational studies, and are useful in establishing temporal relationships between an agent and disease (i.e. whether exposure to the agent occurred before development of the disease).[9] Behind randomized controlled trials, cohort studies are the strongest form of scientific evidence.[10]

2. Case-Control Studies

Case-control studies, also known as "retrospective studies," differ in that they begin with a group of people with the disease (i.e. cases) and then look *backward* to determine if the subjects were ever exposed to the agent in question. In a case-control study, scientists identify two groups: (1) participants who have a disease and (2) participants who do not. Scientists then compare each group's past history of exposures to the

---

[7] *Reference Guide on Epidemiology* at 555-56.
[8] *Id.* at 557.
[9] *Id*. at 558.
[10] Hassan Murad et al., *New evidence pyramid*, 0 Evidence Based Medicine 2 (2016), http://ebm.bmj.com/content/early/2016/06/23/ebmed-2016-110401.full

agent in question. If the agent causes the disease, the scientists should find a higher proportion of past exposures among those with the disease.[11]

Using our previous example, a scientist conducting a case-control study would select two equally-sized groups of participants: (1) those who have the flu and (2) and those who did not. The scientist would then look backward to determine whether the persons ever rode the elevator (i.e. whether they were "exposed" to the agent). If riding the evaluator causes the flu, the scientist should find that people who contracted the flu rode the elevator more frequently than those who never got sick.

Because case-control studies do not require that scientists track subjects for development of a disease, they are generally less expensive than cohort studies. They fall just below cohort studies in the hierarchy of scientific evidence.[12]

Case reports, which describe clinical events in a single patient, are among the weakest forms of scientific evidence.[13] Case reports cannot establish a causal link between an agent and a disease because they lack a comparison group and fail to exclude alternative causes. As a result, they are generally viewed as unreliable scientific evidence of causation.[14]

STRENGTH OF ASSOCIATION

An association exists when events occur together more frequently than if by chance. Epidemiologists commonly measure the strength of an association with a relative risk or odds ratio number. Both numbers compare the likelihood of an event occurring between two groups. In most cases, a cohort study quantifies risk using a relative risk number; a case-control study uses an odds ratio number.

1. Relative Risk

With the understanding that risk is always present, relative risk quantifies the amount of *excess* risk that is associated with an agent. In other words, the relative risk is a number that shows how much more likely a person is to get a disease if exposed to an agent. Relative risk is calculated by dividing the incidence rate of disease in the exposed group by the incidence rate of disease in the unexposed group.

A relative risk of 1.0 means that there is no association between an agent and a disease, and that the risk of contracting the disease is the same in both exposed and unexposed individuals. A relative risk of less than 1.0 means there is a negative association between the exposure and disease. A relative risk greater than 1.0 means there is a positive association between an agent and disease, which could be causal. The higher the relative risk, the stronger the association.

Staying with our elevator example, suppose a cohort study found that 66 out of 100 people who rode the elevator contracted the flu, compared to only 22 out of 100 people who did not ride the elevator. To calculate the relative risk, the scientists would divide the incidence rate of flu in elevator-riding people (66/100 = 0.66) by the incidence rate of flu of non-elevator-riding people (22/100 = 0.22). This equals a relative risk of 3.0

---

[11] *Reference Guide on Epidemiology* at *559.*
[12] Murad, *supra* note 9, at 2.
[13] *Id.*
[14] *See Casey v. Ohio Med. Prod.*, 877 F. Supp. 1380, 1385 (N.D. Cal. 1995) ("case reports are not reliable scientific evidence of causation").

(0.66/0.22). This relative risk not only shows a positive association between riding the elevator and contracting the flu (because it is over 1.0), but also implies that people who ride the elevator are three times more likely to get the flu than people who do not.

Many jurisdictions will only admit epidemiologic evidence if the relative risk is *greater than* 2.0—a level that permits an inference that the disease was more likely than not caused by the agent in question.[15] In essence, a relative risk of 2.0 is a doubling of the risk at issue. Others courts reject this reasoning and will admit epidemiologic studies with a relative of *2.0 or less* as evidence of causation, thereby leaving the sufficiency of the evidence for the jury to decide.[16]

2. Odds Ratio

Epidemiologists also measure the strength of an association with an odds ratio number. An "odds ratio" is the ratio of the odds of an event occurring in one group to the odds of it occurring in another group. An odds ratio is particularly useful when the disease in question is rare.[17]

For example, suppose a researcher wanted to compare the relative odds of death between men and women on the Titanic. While the exact numbers of passengers and crew aboard the Titanic are unknown, an estimated 1,690 men and 425 women were on the ship that fateful night. Of the 1,690 men, 338 survived and 1,352 died, and of the 425 women, 316 survived and 109 died. For men, the odds were 4 to 1 in favor of death (1,352/338=4.0). For women, the odds were 3 to 1 against dying (109/316=0.344). The odds ratio calculates to 11.627 (4.0/0.344). Thus, the odds of a man dying on Titanic were more than 11 times greater than that of a woman.

RANDOM ERROR

Epidemiological studies are often based on relatively small sample groups. As a result, a study may erroneously find an association where one does not actually exist, or not find an association where one does exist, simply due to "chance" or "random error."[18] Scientists can assess the potential for random error by using a study's *p*-value and/or confidence interval to determine if a study is "statistically significant." A study result that is not statistically significant cannot be used to infer causation.

1. The *p*-Value

"A *p*-value represents the probability that an observed positive association could result from random error even if no association were in fact present."[19] For example, a *p*-value of 0.05—which is the most common significance level—means that there is a 5% chance that will study will erroneously find an association where no true association exists. Thus, is a p-value of 0.05 is used for a study, the scientist can be 95% sure that the observed association is true. As long as the *p*-value for the study is less than 5%,

---

[15] *See, e.g., Daubert v. Merrell Dow Pharms, Inc.*, 43 F.3d 1311, 1321 (9th Cir. 1995) ("the relative risk . . . will, at a minimum, have to exceed 2") (internal quotations omitted).
[16] *Reference Guide on Epidemiology* at 566-67, 612, 616.
[17] *Id.* at 568.
[18] *Id*. at 572-73.
[19] *Id.* at 576.

then the relative risk or odds ratio can potentially be "statistically significant." If the *p*-value is greater than 5%, then the results cannot be "statistically significant."

2. The Confidence Interval

A second way to assess random error is by using a confidence interval. A confidence interval is a range of values within which the true value is likely to fall. Suppose a study finds a relative risk of 2.66 with a 95% confidence interval of 2.14 to 3.36. The confidence interval tells scientists that they can be 95% sure that the true relative risk is somewhere between 2.14 and 3.36. Stated differently, the true relative risk will be between 2.14 and 3.36 if the study was repeated numerous times. The more narrow the confidence interval, the more precise the result. A relative risk cannot be considered statistically significant if the confidence interval includes 1.0.[20] The reason is because, as stated above, a relative risk of 1.0 means no association.

If a study fails to find a statistically significant association, scientists should consider the "power" of the study in determining what weight to give the results. "Power" is the ability of a study to detect an association. The power of a study depends on the sample size, the statistical significance specified, the background incidence of disease, and the specified relative risk selected by the scientist. The greater the power of the study, the more reliable the results.[21]

OTHER FACTORS TO CONSIDER

Bias and confounding factors are two other sources of error capable of producing a false result in an epidemiological study and should be taken into account when interpreting the results.

1. Bias

Bias refers to anything "that results in a systematic (nonrandom) error in a study result and thereby compromises its validity."[22] The presence of significant bias may invalidate the results of a study altogether, but lesser bias may be overcome by other factors that limit its effect. The two primary types of bias in epidemiological studies are selection bias and information bias.

Selection bias results from the method by which study participants are chosen. Groups of subjects should ideally be drawn from the same population, and should not be selected on the basis of diagnosis of a disease, referral, or volunteerism.[23]

Information bias results from "inaccurate information about either the disease or the exposure status of the study participants or a result of confounding."[24] Inaccurate information can come from incomplete or unreliable sources, differences in data collection protocol, or from misdiagnosis of disease status. Case-control studies are particularly susceptible to information bias because they often rely on interviews with

[20] *Id.* at 576-581.
[21] *Id.* at 582.
[22] *Id.* at 583.
[23] *Id.* at 583-85.
[24] *Id.* at 585.

the study participants or family members about past exposures which test memory and may be speculative.[25]

2. Confounding Factors

A "confounding factor" is an "extra" factor in a study group which independently increases both the risk of disease and exposure.[26] If not properly accounted for, confounding factors can skew the results of a study by producing an observed association when no true association exists.

A classic example of a confounding factor is found in the association between murder rates and ice cream sales. Statistics show that when ice cream sales increase, so do murder rates. We hope that eating ice cream does not cause an increase in violent crime! The simple explanation for this association is that warmer weather acts as a confounding factor, causing an increase in both ice cream sales and murder rates. If left unaccounted for, this confounding factor skews the results and shows an association that does not in fact exist.

ASSOCIATION VS. CAUSATION

A basic tenant of epidemiology is that association and causation are different. After determining that an observed association is statistically significant – and then minimizing chance, bias, and confounding factors – scientists must then consider scientifically recognized guidelines to determine if the association is truly causal.

The Bradford Hill criteria provides a number of factors for scientists to consider in assessing causation, including: (1) the existence of a temporal relationship; (2) strength of the association; (3) dose-response relationship; (4) replication of the findings; (5) biological plausibility (coherence with existing knowledge); (6) consideration of alternative explanations; (7) cessation of exposure; (8) specificity of the association; and (9) consistency with other knowledge. While not all factors must be present for a causal relationship to exist, an assessment of causation requires this analysis.[27]

CONCLUSION

Epidemiological evidence plays a significant role in evaluating general causation. However, no epidemiological study is perfect. So the next time you are presented with an epidemiological study, practitioners attempting to attack or support it should ask:

1. Was the type of study appropriate to the research question?
2. Was an appropriate sample size used?
3. How were the participants/controls recruited?
4. Were confounding factors considered and appropriately accounted for?
5. How strong is the association between exposure and disease?
6. How wide is the confidence interval?
7. Does the relative risk meet the jurisdictional requirement for admissibility?
8. Is the association consistent with other research or scientific literature?
9. How many Bradford Hill criteria are satisfied?
10. Do the numbers suggest causation?

---

[25] *Id.* at 585-90.
[26] *Id.* at 591.
[27] *Id.* at 552, 598-600.

Practitioners should also consider holding a “Science Day” to educate the judge on epidemiology in cases where such evidence plays a crucial role. “Science Days” have been used in various courts to allow parties to explain the history and background of products and to present relevant medical and scientific literature.[28] Among other things, a well-executed “Science Day” lays the groundwork for later motions to bar expert testimony based on unreliable epidemiologic studies.

[28] *See, e.g., In re Depakote,* No. 14-CV-847-NJR-SCW, 2015 WL 4775868, at *3 n.2 (S.D. Ill. Feb. 13, 2015).