http://www.fjc.gov/public/pdf.nsf/lookup/SciMan3D01.pdf/$file/SciMan3D01.pdf

Reference Manual on Scientific Evidence
Third Edition

Committee on the Development of the Third Edition of the
Reference Manual on Scientific Evidence
Committee on Science, Technology, and Law
Policy and Global Affairs

FEDERAL JUDICIAL CENTER

THE NATIONAL ACADEMIES PRESS 500 Fifth Street, N.W. Washington, DC 20001

II. Interpreting *Daubert*

Although almost 20 years have passed since *Daubert* was decided, a number of basic interpretive issues remain.

A. Atomization

When there is a ***Daubert* challenge to an expert**, should the court look at all the studies on which the expert relies for their collective effect or should the court examine the reliability of each study independently? **The issue arises with proof of causation in toxic tort cases when plaintiff's expert relies on studies from different scientific disciplines, or studies within a discipline that present different strengths and weaknesses, in concluding that defendant's product caused plaintiff's adverse health effects**. Courts rarely discuss this issue explicitly, but some appear to look at each study separately and give no consideration to those studies that cannot alone prove causation.

Although some use the language in *Joiner* as the basis for this slicing-and-dicing approach,50 **scientific inference typically requires consideration of numerous findings, which, when considered alone, may not individually prove the contention.**51 **It appears that many of the most well-respected and prestigious scientific bodies (such as the International Agency for Research on Cancer (IARC), the Institute of Medicine, the National Research Council, and the National Institute for Environmental Health Sciences) consider all the**

relevant available scientific evidence, taken as a whole, to determine which conclusion or hypothesis regarding a causal claim is best supported by the body of evidence. In applying the scientific method, scientists do not review each scientific study individually for whether by itself it reliably supports the causal claim being advocated or opposed. Rather, as the Institute of Medicine and National Research Council noted, "summing, or synthesizing, data addressing different linkages [between kinds of data] forms a more complete causal evidence model and can provide the biological plausibility needed to establish the association" being advocated or opposed.52 **The IARC has concluded that "[t]he final overall evaluation is a matter of scientific judgment *reflecting the weight of the evidence* derived from studies in humans, studies in experimental animals, and mechanistic and other relevant data."53**

B. Conflating Admissibility with Sufficiency

In *Daubert,* Justice Blackmun's opinion explicitly acknowledges that in some cases admissible evidence may not suffice to support a verdict in favor of plaintiffs. In other words, it seems to recognize that the admissibility determination comes first and is separate from the sufficiency determination. But in *Joiner* the Court pays little attention to this distinction and suggests that plaintiff's expert testimony may be excluded if the evidence on which he seeks to rely is itself deemed insufficient.

But what difference does it make if sufficiency is conflated with admissibility?54 After all, the case's final outcome will be the same. **As *Daubert* recognizes, the trial judge's authority to decide whether the plaintiff has produced sufficient evidence to withstand a dispositive motion under Rule 56 or 50 is indisputable; a one-step process that considers sufficiency when adjudicating a *Daubert* motion is arguably more efficient than a two-step process that requires the district judge to analyze admissibility before it can turn to sufficiency**.

There are, however, consequences to conflating admissibility and sufficiency. The de novo standard of review that ordinarily applies to judgments as a matter of law following a determination of insufficient evidence is converted into the lower abuse-of-discretion standard that governs evidentiary rulings on admissibility, and thereby undermines the jury trial mandate of the Seventh Amendment. **Science proceeds by cumulating and synthesizing evidence until there is enough for a new paradigm. That does not mean that every study meets the most rigorous scientific standards. Judgment is required in determining which inferences are appropriate, but an approach that encourages looking at studies sequentially rather than holistically has costs that must be considered**.

---

51. *See e.g.*, Susan Haack, *An Epistemologist in the Bramble-Bush: At the Supreme Court with Mr. Joiner*, 26 J. Health Pol. Pol'y & L. 217–37 (1999) (discussing the individual studies that lead to the compelling inference of a double-helical structure of a DNA molecule, which, when considered separately, fail to compel that inference). ***See also* Milward v. Acuity Specialty Products Group, Inc., __ F.3d __, 2011 WL 982385, *10 639 F.3d 11, 26 (1st Cir. 2011) (reversing the district court's exclusion of expert testimony based on an assessment of the direct causal effect of the individual studies, finding that the "weight of the evidence" properly supported the expert's opinion that exposure to benzene can cause acute promyelocytic leukemia).**
(pp. 19-21)

…III. Applying *Daubert*

B. Assessing the Scientific Foundation of Studies from Different Disciplines

Expert opinion is typically based on multiple studies, and those studies may come from different scientific disciplines. Some courts have explicitly stated that certain types of evidence proffered to prove causation have no probative value and therefore cannot be reliable.59 Opinions based on animal studies have been rejected because of reservations about extrapolating from animals to humans or because the plaintiff's extrapolated dose was lower than the animals'—which is invariably the case because one would have to study unmanageable, gigantic numbers of animals to see results if animals were not given high doses. **The field of toxicology, which, unlike epidemiology, is an experimental science, is rapidly evolving, and prior case law regarding such studies may not take into account important new developments**.

But even when there are epidemiological studies, a court may conclude that they cannot prove causation because they are not conclusive and therefore unreliable. And if they are unreliable, they cannot be combined with other evidence.60

Experts will often rely on multiple studies, each of which has some probative value but, when considered separately, cannot prove general causation.

As noted above, trial judges have great discretion under *Daubert* and a court is free to choose an atomistic approach that evaluates the available studies one by one. **Some judges have found this practice contrary to that of scientists who look at knowledge incrementally.61** But there are no hard-and-fast scientific rules for synthesizing evidence, and **most research can be critiqued on a variety of grounds. Few studies are flawless. Epidemiology is vulnerable to attack because of problems with confounders and bias. Furthermore, epidemiological studies are grounded in statistical models**. What role should statistical significance play in

assessing the value of a study? Epidemiological studies that are not conclusive but show some increased risk do not prove a lack of causation. **Some courts find that they therefore have some probative value,62 at least in proving general causation.63**

Even, however, if plaintiffs convince the trial judge that their experts relied on reliable and relevant evidence in establishing general causation, that is, in opining that the defendant's product can cause the adverse effects for which plaintiffs seek compensation, **plaintiffs must *also* present admissible expert testimony that the defendant's product caused their specific injuries**.

---

59. *See, e.g., In re* Rezulin, 2004 WL 2884327, at *3 (S.D.N.Y. 2004); Cloud v. Pfizer Inc., 198 F. Supp. 2d 1118, 1133 (D. Ariz. 2001) (stating that case reports were merely compilations of occurrences and have been rejected as reliable scientific evidence supporting an expert opinion that *Daubert* requires); Haggerty v. Upjohn Co., 950 F. Supp. 1160, 1164 (S.D. Fla. 1996), *aff'd,* 158 F.3d 588 (11th Cir. 1998) ("scientifically valid cause and effect determinations depend on controlled clinical trials and epidemiological studies"); Wade-Greaux v. Whitehall Labs., Inc., 874 F. Supp. 1441, 1454 (D.V.I. 1994), *aff'd,* 46 F.3d 1120 (3d Cir. 1994) (stating there is a need for consistent epidemiological studies showing statistically significant increased risks).

60. *See* Hollander v. Sandoz Pharm. Corp., 289 F.3d 1193, 1216 n.21 (10th Cir. 2002) ("To suggest that those individual categories of evidence deemed unreliable by the district court may be added to form a reliable theory would be to abandon 'the level of intellectual rigor of the expert in the field.'").

61. *See, e.g., In re* Ephedra, 393 F. Supp. 2d 181, 190 (S.D.N.Y. 2005) (allowing scientific expert testimony regarding "a confluence of suggestive, though non-definitive, scientific studies [that] make[s] it more-probable-than-not that a particular substance . . . contributed to a particular result. . . ."; after a two-week *Daubert* hearing in a case in which there would never be epidemiological evidence, the court concluded that some of plaintiffs' experts could testify on the basis of animal studies, analogous human studies, plausible theories of the mechanisms involved, etc.); **Milward v. Acuity Specialty Prods. Group, Inc., 639 F.3d 11 (1st Cir. 2011)**.

(pp. 23-24)

…*Reference Guide on Epidemiology*

II. What Different Kinds of Epidemiologic Studies Exist?

C. Epidemiologic and Toxicologic Studies

In addition to observational epidemiology, toxicology models based on live animal studies (in vivo) may be used to determine toxicity in humans.39 Animal studies have a number of advantages. They can be conducted as true experiments, and researchers control all aspects of the animals' lives. Thus, they can avoid the problem of confounding,40 which epidemiology often confronts. Exposure can be carefully controlled and measured. Refusals to participate in a study are not an issue, and loss to followup very often is minimal. Ethical limitations are

diminished, and animals can be sacrificed and their tissues examined, which may improve the accuracy of disease assessment. Animal studies often provide useful information about pathological mechanisms and play a complementary role to epidemiology by assisting researchers in framing hypotheses and in developing study designs for epidemiologic studies.

Animal studies have two significant disadvantages, however. First, animal study results must be extrapolated to another species—human beings—and differences in absorption, metabolism, and other factors may result in interspecies variation in responses. For example, one powerful human teratogen, thalidomide, does not cause birth defects in most rodent species.41 Similarly, some known teratogens in animals are not believed to be human teratogens. In general, it is often difficult to confirm that an agent known to be toxic in animals is safe for human beings.42 **The second difficulty with inferring human causation from animal studies is that the high doses customarily used in animal studies require consideration of the dose–response relationship and whether a threshold no-effect dose exists.43** Those matters are almost always fraught with considerable, and currently unresolvable, uncertainty.44

Toxicologists also use in vitro methods, in which human or animal tissue or cells are grown in laboratories and are exposed to certain substances. The problem with this approach is also extrapolation—whether one can generalize the findings from the artificial setting of tissues in laboratories to whole human beings.45

Often toxicologic studies are the only or best available evidence of toxicity.46 **Epidemiologic studies are difficult, time-consuming, expensive, and sometimes, because of limited exposure or the infrequency of disease, virtually impossible to perform.47 Consequently, they do not exist for a large array of environmental agents**. **Where both animal toxicologic and epidemiologic studies are available, no universal rules exist for how to interpret or reconcile them.48**

---

39. For an in-depth discussion of toxicology, see Bernard D. Goldstein & Mary Sue Henifin, Reference Guide on Toxicology, in this manual.

40. *See infra* Section IV.C.

41. Phillip Knightley et al., Suffer the Children: The Story of Thalidomide 271–72 (1979).

42. *See* Ian C.T. Nesbit & Nathan J. Karch, Chemical Hazards to Human Reproduction 98–106 (1983); Int'l Agency for Research on Cancer (IARC), Interpretation of Negative Epidemiologic Evidence for Carcinogenicity (N.J. Wald & Richard Doll eds., 1985) [hereafter IARC].

43. *See infra* Section V.C & note 119.

44. *See* Soldo v. Sandoz Pharms. Corp., 244 F. Supp. 2d 434, 466 (W.D. Pa. 2003) (quoting this reference guide in the first edition of the Reference Manual); ***see also*** **General Elec. Co. v. Joiner, 522 U.S. 136, 143–45 (1997) (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).**

45. For a further discussion of these issues, see Bernard D. Goldstein & Mary Sue Henifin, Reference Guide on Toxicology, Section III.A, in this manual.

46. IARC, a well-regarded international public health agency, evaluates the human carcinogenicity of various agents. In doing so, IARC obtains all of the relevant evidence, including animal studies as well as any human studies. On the basis of a synthesis and evaluation of that evidence, IARC publishes a monograph containing that evidence and its analysis of the evidence and provides a categorical assessment of the likelihood the agent is carcinogenic. In a preamble to each of its monographs, IARC explains what each of the categorical assessments means. Solely on the basis of the strength of animal studies, IARC may classify a substance as "probably carcinogenic to humans." International Agency for Research on Cancer, *Human Papillomaviruses*, 90 Monographs on the Evaluation of Carcinogenic Risks to Humans 9–10 (2007), *available at* http://monographs.iarc.fr/ENG/Monographs/vol90/index.php; *see also* Magistrini v. One Hour Martinizing Dry Cleaning, 180 F. Supp. 2d 584, 600 n.18 (D.N.J. 2002). When IARC monographs are available, they are generally recognized as authoritative. Unfortunately, IARC has conducted evaluations of only a fraction of potentially carcinogenic agents, and many suspected toxic agents cause effects other than cancer.

47. Thus, in a series of cases involving Parlodel, a lactation suppressant for mothers of newborns, efforts to conduct an epidemiologic study of its effect on causing strokes were stymied by the infrequency of such strokes in women of child-bearing age. *See, e.g.*, Brasher v. Sandoz Pharms. Corp., 160 F. Supp. 2d 1291, 1297 (N.D. Ala. 2001). In other cases, a plaintiff's exposure to an overdose of a drug may be unique or nearly so. *See* Zuchowicz v. United States, 140 F.3d 381 (2d Cir. 1998).

48. *See* IARC, *supra* note 41 (identifying a number of substances and comparing animal toxicology evidence with epidemiologic evidence); Michele Carbone et al., *Modern Criteria to Establish Human Cancer Etiology*, 64 Cancer Res. 5518, 5522 (2004) (National Cancer Institute symposium concluding that "There should be no hierarchy [among different types of scientific methods to determine cancer causation]. Epidemiology, animal, tissue culture and molecular pathology should be seen as integrating evidences in the determination of human carcinogenicity.")

A number of courts have grappled with the role of animal studies in proving causation in a toxic substance case. One line of cases takes a very dim view of their probative value. For example, in *Brock v. Merrell Dow Pharmaceuticals, Inc.*, 874 F.2d 307, 313 (5th Cir. 1989), the court noted the "very limited usefulness of animal studies when confronted with questions of toxicity." A similar view is reflected in *Richardson v. Richardson-Merrell, Inc.*, 857 F.2d 823, 830 (D.C. Cir. 1988), Bell v. Swift Adhesives, Inc., 804 F. Supp. 1577, 1579–80 (S.D. Ga. 1992), and *Cadarian v. Merrell Dow Pharmaceuticals, Inc.*, 745 F. Supp. 409, 412 (E.D. Mich. 1989).

Other courts have been more amenable to the use of animal toxicology in proving causation. Thus, in *Marder v. G.D. Searle & Co.*, 630 F. Supp. 1087, 1094 (D. Md. 1986), *aff'd sub nom. Wheelahan v. G.D. Searle & Co.*, 814 F.2d 655 (4th Cir. 1987), the court observed: "There is a range of scientific 50methods for investigating questions of causation—for example, toxicology and animal studies, clinical research, and epidemiology—which all have distinct advantages and disadvantages." **In *Milward v. Acuity Specialty Products Group, Inc.*, 639 F.3d 11, 17-19 (1st Cir. 2011), the court endorsed an expert's use of a "weight-of-the-evidence" methodology, holding that the district court abused its discretion in ruling inadmissible an expert's testimony about causation based on that methodology. As a corollary to recognizing weight of the evidence as a valid scientific technique, the court also noted the role of judgment in making an appropriate inference from the evidence. While recognizing the legitimacy of the methodology, the court also acknowledged that, as with any scientific technique, it can be improperly applied.** *See also* Metabolife Int'l, Inc. v. Wornick, 264 F.3d 832, 842 (9th Cir. 2001) (holding that the lower court erred in per se dismissing animal studies, which must be examined to determine whether they are appropriate as a basis for causation

determination); *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 sufficient basis for opinion on general causation.); *cf. In re* Paoli R.R. Yard PCB Litig., 916 F.2d 829, 853–54 (3d Cir. 1990) (questioning the exclusion of animal studies by the lower court). The Third Circuit in a subsequent opinion in *Paoli* observed:

> [I]n 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. Thus, the requirement of reliability, or "good grounds," extends to each step in an expert's analysis all the way through the step that connects the work of the expert to the particular case.

In re Paoli R.R. Yard PCB Litig., 35 F.3d 717, 743 (3d Cir. 1994); *see also* Cavallo v. Star Enter., 892 F. Supp. 756, 761–63 (E.D. Va. 1995) (courts must examine each of the steps that lead to an expert's opinion), *aff'd in part and rev'd in part,* 100 F.3d 1150 (4th Cir. 1996).
(pp. 563-565)