ADMISSIBILITY VERSUS SUFFICIENCY: CONTROLLING THE QUALITY OF EXPERT WITNESS TESTIMONY

Michael D. Green[*] *& Joseph Sanders*[**]

INTRODUCTION

In *Daubert v. Merrell Dow Pharmaceuticals, Inc.*,[1] the United States Supreme Court established a new method of assessing the admissibility of expert testimony in federal courts.[2] The opinion and two subsequent opinions, *General Electric Co. v. Joiner*[3] and *Kumho Tire Co. v. Carmichael*,[4] have had a profound impact on civil litigation in the United States.[5] Parties now routinely challenge the admissibility of the opposing expert's testimony. This has led to the exclusion of experts in hundreds, possibly thousands, of cases. Given this substantial impact, it is perhaps surprising that considerable confusion still persists concerning exactly how trial courts should think about the decision to admit or exclude expert testimony. In this Article, we argue that much of the confusion exists because of the unresolved relationship between the admissibility of expert testimony on the one hand and the sufficiency of the scientific evidence to support a plaintiff verdict on the other hand. We argue that the best way to clarify the issue is to

---

* Williams Professor of Law, Wake Forest University School of Law. This Article is based in part on a presentation made by the authors at a conference on causation at the University of Girona (Spain) in 2013, recently published in a collection entitled UNCERTAIN CAUSATION IN TORT LAW (Miguel Martín-Casals & Diego M. Papayannis eds., 2015).

** A.A. White Professor of Law, University of Houston Law Center.

1. 509 U.S. 579 (1993).

2. *Id.* at 587–95, 597. In relatively short order, most state courts followed suit. *See* Joëlle Anne Moreno, *What Happens When Dirty Harry Becomes an (Expert) Witness for the Prosecution?*, 79 TUL. L. REV. 1, 15 n.66 (2004) ("Nearly thirty states have adopted *Daubert*, in whole or in part.").

3. 522 U.S. 136 (1997).

4. 526 U.S. 137 (1999).

5. The impact on criminal cases has been much less pronounced, primarily for reasons that go beyond the scope of this Article. For a discussion of the differences between civil law and criminal law in the application of expert admissibility rules, see Joseph Sanders, *Applying* Daubert *Inconsistently?: Proof of Individual Causation in Toxic Tort and Forensic Cases*, 75 BROOK. L. REV. 1367, 1367–69 (2010); Julie A. Seaman, *A Tale of Two* Dauberts, 47 GA. L. REV. 889, 890–93 (2013).

appreciate that most admissibility decisions regarding expert testimony are best thought of as sufficiency judgments about the scientific evidence supporting the expert's testimony. We not only believe this is a better approach, we also believe that a close reading of opinions reveals that, in fact, many courts do adopt a sufficiency approach when making admissibility rulings.[6]

An important tenet that is embedded in our discussion requires identification before we proceed. Any assessment of causation involves an inferential process from evidence to the causation conclusion.[7] Causation, unlike, for example, the presence of another human being, is not something we observe directly but rather is an inference from what may be very strong or very weak circumstantial evidence. This principle is an important foundation for our claim that *Daubert* assessments frequently entail consideration of the circumstantial evidence—the science proffered by the plaintiff's expert—and whether that evidence is sufficient to justify an inference of causation.

I. BEFORE *DAUBERT*[8]

The conventional understanding of *Daubert* is that it replaced the general acceptance standard established by *Frye v. United States*[9] for the admissibility of expert scientific testimony. This understanding neglects to account for the fact that in the pre-*Daubert* period, there was little or no screening of expert witnesses in civil cases save on the basis of their qualifications to testify as an expert.[10] Serious screening of expert witnesses did not occur until

---

6. *See, e.g.*, Chapman v. Procter & Gamble Distrib., LLC, 766 F.3d 1296, 1304–06 (11th Cir. 2014); United States v. Frazier, 387 F.3d 1244, 1260–63 (11th Cir. 2004); Allison v. McGhan Med. Corp., 184 F.3d 1300, 1309–13 (11th Cir. 1999).

7. *See* 1 DAVID HUME, TREATISE ON HUMAN NATURE, pt. 3, §§ 14–15, at 155–76 (L.A. Selby-Bigge ed., Oxford Univ. Press 2d ed. 1978) (1888); Alexander Rosenberg, *Hume and the Philosophy of Science, in* THE CAMBRIDGE COMPANION TO HUME 64, 71–77 (David Fate Norton ed., 1993).

8. Portions of this section draw on Michael D. Green, *The Road Less Well Traveled (and Seen): Contemporary Lawmaking in Products Liability*, 49 DEPAUL L. REV. 377 (1999).

9. 293 F. 1013 (D.C. Cir. 1923).

10. Professor Paul Giannelli reports: "The civil cases, spurred by toxic tort litigation, also came later. *Frye* had been applied almost exclusively to criminal cases and was not applied in a federal civil case until 1984." Paul C. Giannelli, Daubert: *Interpreting the Federal Rules of Evidence*, 15 CARDOZO L. REV. 1999, 2008 (1994) (footnotes omitted); *see also* MICHAEL H. GRAHAM, HANDBOOK OF FEDERAL EVIDENCE § 703.2, at 651 (3d ed. 1991) ("The *Frye* test has been applied most frequently over the years in criminal cases . . . ."); 1 DAVID W. LOUISELL & CHRISTOPHER B. MUELLER, FEDERAL EVIDENCE § 105, at 853 (1977) ("The *Frye* standard . . . is rarely applied in civil litigation; *Frye* itself has been cited only in a very few civil cases, principally in state courts in connection with blood tests to determine paternity."); FAUST F. ROSSI, EXPERT WITNESSES 36 (1991) ("The *Frye* standard traditionally has been applied almost exclusively in criminal

the early 1980s with the emergence of large-scale toxic substances litigation—notably, the *Agent Orange*[11] litigation and the Bendectin case congregation.[12]

In the *Agent Orange* litigation, which occurred in parallel with some of the early Bendectin cases, Judge Weinstein granted defendants summary judgment in individual suits brought by veterans who opted out of the class action settlement.[13] The sticking point for the opt-out plaintiffs was proving factual causation, and they had several experts who proposed to testify in support of causation. In the course of his opinion, Judge Weinstein discounted the value of animal studies and emphasized the importance of human epidemiology.[14] He surveyed a wide range of epidemiologic studies that tended to exonerate Agent Orange and resurrected the concerns behind *Frye*. In the end, however, he based his opinion on Federal Rule of Evidence 703, concluding that the facts and data—for example, scientific evidence—relied on by plaintiff's experts did not suffice as a basis for their opinions: "If the underlying data are so lacking in probative force and reliability that no reasonable expert could base an opinion on them, an opinion which rests entirely upon them must be excluded."[15]

---

cases."); Kenneth J. Chesebro, *Galileo's Retort: Peter Huber's Junk Scholarship*, 42 AM. U. L. REV. 1637, 1691–96 (1993); Jennifer L. Mnookin, *Expert Evidence, Partisanship, and Epistemic Competence*, 73 BROOK. L. REV. 1009, 1016 (2008) ("Through 1970, [*Frye*] was cited only fifty-eight times, and the bulk of those cases involved the lie detector, the same technology at issue in *Frye*.").

A Westlaw search for all noncriminal cases that cited *Frye* and were decided before 1990 produced twenty-five cases. Of those twenty-five cases, eight were habeas corpus cases, seven employed the narrow holding of *Frye* to decide that lie detector evidence is inadmissible, four cases cited to *Frye* generally without employing it as a precedent in the case, and five cases employed *Frye* to deal with traditional criminal forensic evidence (e.g., fingerprinting) whose admissibility was at issue in a civil case. Only two of these twenty-five cases were toxic substances cases. In *Brock v. Merrell Dow Pharmaceuticals, Inc.*, 884 F.2d 167 (5th Cir. 1989), Judge Higginbotham cited *Frye* in his dissent to the denial of a rehearing en banc. *Id.* at 168–70, 169 n.2. The final case, *Ellis v. International Playtex, Inc.*, 745 F.2d 292 (4th Cir. 1984), cited *Frye* critically and concluded that the dispute over the validity of the methodology in certain epidemiologic studies should be submitted to the jury, rather than decided by the judge. *Id.* at 303–05.

11. *In re* "Agent Orange" Prod. Liab. Litig., 611 F. Supp. 1223 (E.D.N.Y. 1985), *aff'd*, 818 F.2d 187 (2d Cir. 1987).

12. *See generally* MICHAEL D. GREEN, BENDECTIN AND BIRTH DEFECTS: THE CHALLENGES OF MASS TOXIC SUBSTANCES LITIGATION (1996) (discussing the fifteen-year Bendectin product liabilities litigation, which involved over 2,000 claimants alleging that the drug had caused birth defects); JOSEPH SANDERS, BENDECTIN ON TRIAL: A STUDY OF MASS TORT LITIGATION (1998) (same).

13. *Agent Orange*, 611 F. Supp. at 1264.

14. *Id.* at 1238–41.

15. *Id.* at 1245; *see also* FED. R. EVID. 703 (providing the bases upon which an expert may make an opinion).

Alternatively, he criticized the plaintiffs and their experts for failure to assess the individual plaintiffs, to submit their medical records, or to have treating physicians testify without acknowledging that the evidence about the individual plaintiffs bore on specific causation while epidemiology, in the first instance, addresses general causation.[16] Judge Weinstein sounded a clarion call for careful examination of expert witness testimony in toxic substances cases:

> Such careful scrutiny of proposed evidence is especially appropriate in the toxic tort area. The uncertainty of the evidence in such cases, dependent as it is upon speculative scientific hypotheses and epidemiological studies, creates a special need for robust screening of experts and gatekeeping under Rules 403 and 703 by the court.[17]

By contrast, Judge Thomas Penfield Jackson pursued a different approach in *Richardson v. Richardson-Merrell, Inc.*,[18] an early Bendectin case. As with *Agent Orange*, the critical issue in all of the Bendectin cases was whether Bendectin caused the birth defects of children who were exposed to it in utero.[19] Carita Richardson, born after her mother took Bendectin during her pregnancy, had multiple limb deformities, including no lower right leg.[20] As in *Agent Orange*, the defendant moved for summary judgment.[21] Unlike Judge Weinstein in *Agent Orange*, Judge Jackson denied the motion and tried the case.[22] Plaintiffs' experts testified to Bendectin's chemical structure and its similarity to other substances known to be animal teratogens, described in vivo examination of Bendectin, and criticized and reanalyzed animal studies performed by Merrell and epidemiologic studies of Bendectin that tended to exonerate Bendectin as a teratogen.[23] Dr. Alan Done, a professor of pediatrics, pharmacology, and toxicology at Wayne State University, testified based on this evidence that Bendectin was a human teratogen and that it caused Carita Richardson's birth

---

16. *Agent Orange*, 611 F. Supp. at 1235, 1247–48.

17. *Id.* at 1260. Peter Schuck has chronicled the *Agent Orange* litigation and explained why Judge Weinstein, in order to protect a global class-wide settlement that he had largely masterminded through his appointed special master, Kenneth Feinberg, felt the necessity to dismiss all of the opt-out claims by veterans. *See* PETER H. SCHUCK, AGENT ORANGE ON TRIAL: MASS TOXIC DISASTERS IN THE COURTS 226–44 (enlarged ed. 1987); *see also* Charles Nesson, *Agent Orange Meets the Blue Bus: Factfinding at the Frontier of Knowledge*, 66 B.U. L. REV. 521, 535–37 (1986).

18. 649 F. Supp. 799 (D.D.C. 1986), *aff'd*, 857 F.2d 823 (D.C. Cir. 1988).

19. *Id.*

20. *Id.* at 800.

21. *Id.*

22. *Id.*

23. *Id.* at 801.

defects.[24] Defendant's expert, an academic epidemiologist, summarized the epidemiologic studies that had been performed and their results, which found no statistically significant increase in limb defects.[25]

The jury found for the plaintiffs, awarding $1.16 million, and the defendant moved for judgment notwithstanding the verdict.[26] Referring to the published epidemiologic studies introduced by the defendant, Judge Jackson concluded that "the literature on Bendectin, individually and in the aggregate, fails to demonstrate Bendectin's teratogenicity to a scientifically acceptable degree of accuracy."[27] Judge Jackson's confidence in that assessment was no doubt bolstered by the outcome of an advisory committee convened by the FDA to examine the evidence on Bendectin's safety and make a recommendation about whether the drug was sufficiently safe to remain on the market. After a two-day hearing, the committee had concluded that there was no evidence that Bendectin caused birth defects. Thus, Judge Jackson concluded that "there is now nearly universal scientific consensus that Bendectin has not been shown to be a teratogen, and . . . reasonable jurors could not reject that consensus without indulging in . . . speculation and conjecture . . . ."[28]

---

24. *Id.*
25. *Id.* at 802.
26. *Id.* at 799–800.
27. *Id.* at 802.
28. *Id.* at 803. Two other courts relied on sufficiency of the evidence to decide as a matter of law for Bendectin's manufacturer. *See* Turpin v. Merrell Dow Pharm., Inc., 959 F.2d 1349, 1353 (6th Cir. 1992); Merrell Dow Pharm., Inc. v. Havner, 953 S.W.2d 706, 711–14 (Tex. 1997). Rather than relying on the issue of whether Bendectin was a teratogen, the Sixth Circuit in *Turpin* took a subtly different course. The court recognized that although there was a significant body of exonerative scientific evidence for Bendectin, it was inadequate to rule out the possibility that Bendectin could cause birth defects at a low incident rate—one small enough that the existing studies could not have identified. 959 F.2d at 1358–59. The *Turpin* court concluded that the plaintiff's evidence of Bendectin's teratogenicity simply could not suffice to permit a reasonable jury to conclude that Bendectin more likely than not caused the plaintiff's birth defects—what we refer to today as specific causation. *Id.* at 1350, 1360–61. What is notable is not only that the *Turpin* court, like Judge Jackson in *Richardson,* ruled on sufficiency of the evidence grounds, but that the court engaged in its extensive analysis reviewing a ruling granting summary judgment, without the benefit of a trial record.

As the Supreme Court explained in *Anderson v. Liberty Lobby, Inc.*, 477 U.S. 242 (1986), the standards for judgment as a matter of law and for summary judgment, with regard to sufficiency of the evidence, are the same even if the records on which the two motions are based are different. *Id.* at 250–52.

In addition to *Turpin*, the Texas Supreme Court in *Havner* nominally relied on the insufficiency of the evidence to overturn a jury verdict of $3.75 million in compensatory damages in a late-stage Bendectin case. *Havner*, 953 S.W.2d at 709, 730. But, because of the Texas "no-evidence" requirement for

On appeal, the D.C. Circuit Court of Appeals was faced with a troubling conflict between one of its recent decisions, *Ferebee v. Chevron Chemical Co.*,[29] and the lower court's *Richardson* decision. In *Ferebee*, the court had ruled that epidemiologic evidence was not necessary for a toxic substance—paraquat, an herbicide—in order for the plaintiff to meet the burden of production on cause in fact.[30] Rather, the court set forth the traditional laissez-faire treatment of expert witnesses—passive acceptance of scientific testimony, leaving to the jury resolution of conflicting expert opinions. Thus, the *Ferebee* court concluded that the plaintiff's experts—two treating pulmonary specialists who testified that the fatal lung disease was a result of exposure to paraquat—provided sufficient evidence to support a jury verdict for the plaintiff.[31] As long as "experts are willing to testify" about complex and technical matters at the boundaries of scientific knowledge, then the questions are for the jury, according to the *Ferebee* court.[32]

The court of appeals in *Richardson* adopted a different tack from Judge Jackson in affirming his decision.[33] The court ruled that the plaintiffs' experts' testimony was inadmissible.[34] After devoting almost three pages of its opinion to a review of Dr. Done's testimony and the scientific data that supported it, the court concluded that the opinion lacked an adequate foundation.[35] The lower court was affirmed, but the court of appeals tilted its approach in the direction of Judge Weinstein in *Agent Orange* and his use of Rule 703 to assess the bases for an expert's opinion. The inadmissibility of plaintiffs' expert's testimony both affirmed the lower court decision and made summary judgment more available in toxic substances cases, an effect the court noted with approval.[36] The *Richardson*

---

determining the sufficiency of the evidence, the court was forced to confront the experts' opinions for the plaintiff and assess those opinions as unreliable under the court's precedent for the admissibility of expert testimony. *Id.* at 711–14. Thus, the court stated that reviewing the reliability of an expert's testimony is perfectly appropriate as an aspect of evaluating the sufficiency of the evidence. *Id.* at 713–14. The court concluded that if the "foundational data" relied on by the expert are unreliable, the expert's testimony is unreliable and provides "no evidence" with regard to the sufficiency of the evidence on causation. *Id.* at 714. That sounds a great deal like saying that when the scientific evidence is inadequate, there is insufficient evidence of causation.

29. 736 F.2d 1529 (D.C. Cir. 1984).

30. *Id.* at 1534–36.

31. *Id.* at 1535–46.

32. *Id.* at 1534.

33. Richardson v. Richardson-Merrell, Inc., 857 F.2d 823 (D.C. Cir. 1988).

34. *Id.* at 831.

35. *Id.* at 829–32.

36. *Id.* at 832–33. By contrast, *Ferebee* had permitted experts who did not even purport to rely on traditional toxicological methods to testify about causation. Rather, the plaintiff's experts were treating physicians who relied on their clinical examinations of the patient and three similar cases. 736 F.2d at 1535. In what sounded very much like Merrell's claims in *Richardson*, the

court distinguished *Ferebee* on the ground that Bendectin involved a drug for which the scientific evidence was mature, unlike paraquat, about which there had been little scientific study.[37]

While the court of appeals in *Richardson* pursued the inadmissibility direction developed in *Agent Orange,* we want to return briefly to Judge Jackson's sufficiency decision to highlight different types of sufficiency analyses. Judge Jackson cited *Merit Motors, Inc. v. Chrysler Corp.*[38] in support of his conclusion that reasonable jurors could not find causation without running afoul of his "speculation" observation.[39] The *Merit Motors* court affirmed a grant of summary judgment against an antitrust plaintiff who had no evidence and only the speculation of an expert witness to demonstrate that the defendant's selling programs caused the plaintiff-dealers' harm.[40]

We want to suggest a subtle difference between the *Richardson* "speculation" assessment and that in *Merit Motors.* The latter involved the classic situation in which the plaintiff's evidence simply is not sufficient to satisfy the burden of production. The burden of production requires a party with the burden of proof to "introduce[] enough evidence to permit the jury to find in her favor."[41] Thus, if a plaintiff begins a trial by resting, the ground for dismissal would be that the plaintiff failed in her burden of production. Similarly, if the plaintiff's evidence is circumstantial and requires too great an inference to justify a finding on the issue in dispute, the court will dismiss for failure to satisfy the burden of production. In *Gordon v. American Museum of Natural History,*[42] the plaintiff fell on a piece of slippery paper on the steps of the defendant museum.[43] On the issue of whether the defendant had sufficient notice of the paper's existence–essential for foreseeability of risk and negligence–the court held that the plaintiff introduced insufficient evidence about

---

defendant argued that the plaintiff's experts' opinions were outside the mainstream of scientific thought and therefore inadmissible. *Id.* Nevertheless, the *Ferebee* court held that neither epidemiologic nor animal toxicology studies were required to support an expert opinion in a toxic substance case:

> As long as the basic methodology employed to reach such a conclusion is sound, . . . products liability law does not preclude recovery until a "statistically significant" number of people have been injured or until science has had the time and resources to complete sophisticated laboratory studies of the chemical.

Id. at 1535–36.

37. *Richardson*, 857 F.2d at 831–32.

38. 569 F.2d 666 (D.C. Cir. 1977).

39. Richardson v. Richardson-Merrell, Inc., 649 F. Supp. 799, 803 (D.D.C. 1986).

40. 569 F.2d at 673.

41. JACK H. FRIEDENTHAL ET AL., CIVIL PROCEDURE: CASES AND MATERIALS 16 (10th ed. 2009).

42. 492 N.E.2d 774 (N.Y. 1986).

43. *Id.* at 774.

notice.[44] No reasonable inference of notice could be drawn from the fact that there was paper on the defendant's steps without something further to indicate that the paper had been there long enough that the defendant knew or should have known of its existence.[45]

A somewhat different basis for judgment as a matter of law is revealed in *O'Connor v. Pennsylvania Railroad*.[46] In *O'Connor*, there was insufficient evidence to permit a jury to find that the plaintiff slipped on old ice on the terrace of the defendant's railroad station.[47] The issue in the case was whether the ice was leftover from an older storm (in which case the defendant could be found negligent) or whether the condition of the terrace was the product of a current storm (in which case the defendant would not be liable).[48] The plaintiff testified that he slipped on ice that was dirty and gray, not on fresh snow.[49] That testimony, if credited by the jury, would surely be sufficient evidence on which to rest a finding against the defendant. Against that evidence were reports from the United States Weather Bureau that virtually ruled out the possibility that any product of the prior storm could have remained on the terrace when the plaintiff fell.[50] The unchallenged objective evidence of the Weather Bureau overwhelmed the self-serving testimony of the plaintiff, reasoned the court, "so as to render the inference sought to be drawn [that the plaintiff slipped on old snow] unreasonable."[51]

Thus, *Gordon* and *O'Connor* reveal two different ways in which the sufficiency of the evidence can be a basis for judgment as a matter of law (or summary judgment).[52] The first, reflected in *Gordon*, involves a situation in which the evidence introduced is insufficient to permit a reasonable inference of the necessary facts. That does not mean no evidence was introduced, nor that some of that evidence fails to make the necessary facts more likely. Surely, the evidence that there was a slippery piece of paper on the steps of the Museum of Natural History made it more likely

---

44. *Id.* at 775.

45. *Id.* For a similar result, see Wal-Mart Stores, Inc. v. Reece, 81 S.W.3d 812, 813, 816–17 (Tex. 2002).

46. 308 F.2d 911 (2d Cir. 1962).

47. *Id.* at 912.

48. *Id.* at 913.

49. *Id.*

50. *Id.* at 913–14.

51. *Id.* at 915.

52. There is a third way when the evidence only reveals a fifty-fifty probability that the plaintiff or the defendant's version is what occurred and there is no basis for an inference that would permit the factfinder to find otherwise. *See* Denman v. Spain, 135 So. 2d 195, 197 (Miss. 1961). We need not be concerned with this third type, as it has not revealed itself in toxic substances litigation.

that there was a piece of slippery paper on the steps for long enough to be attended to. But, at some level of weak circumstantial evidence, courts decline to permit a jury to address the issue and instead rule as a matter of law on sufficiency grounds.[53] *O'Connor*, on the other hand, was not about circumstantial evidence that was too weak; the plaintiff's testimony was awfully close to direct evidence, and the inference from "dirty gray" ice to it being leftover from a prior storm was quite strong. Without contradictory evidence, the case surely would be one for the jury. But the evidence contradicting the plaintiff's testimony was too powerful to permit the jury to make a contrary finding.

II. THE EMERGENCE OF *DAUBERT*

By the late 1980s, the small coterie of plaintiffs' lawyers who had pursued Bendectin litigation were back on their heels having suffered defeats not only in *Richardson* and *Turpin v. Merrell Dow Pharmaceuticals, Inc.*,[54] but also in several other federal courts of appeals,[55] and in a multidistrict trial involving several hundred plaintiffs conducted in Cincinnati in 1985.[56] At this late stage, the Bendectin litigation, which had been ongoing for a decade, looked as if it might peter out and disappear.

Two Bendectin cases filed in federal court in southern California completed discovery and had reached the dispositive motion stage. The defendant filed a motion for summary judgment.[57] The plaintiffs' lawyers mustered not only the two primary experts they had previously used but six additional experts, including those with expertise in biostatistics, epidemiology, pediatrics, pharmacology, and toxicology.[58] However, by this point in time, the trial judge had the benefit of numerous other opinions addressing causation in Bendectin cases. Rather than providing a coherent explanation of the basis for granting summary judgment, the court employed a pastiche from those decisions to rationalize the judge's conviction that the case should not go forward.[59] One strand,

---

53. See the discussion of *Wolf v. Kaufman*, 237 N.Y.S. 550 (N.Y. App. Div. 1929), *infra* notes 212–14 and accompanying text.

54. 959 F.2d 1349 (6th Cir. 1992).

55. *See, e.g.*, Lynch v. Merrell-Nat'l Labs., Div. of Richardson-Merrell, Inc., 830 F.2d 1190 (2d Cir. 1987); Oxendine v. Merrell Dow Pharm., Inc., 506 A.2d 1100 (D.C. Cir. 1986).

56. *In re* Richardson-Merrell, Inc. "Bendectin" Prods. Liab. Litig., 624 F. Supp. 1212 (S.D. Ohio 1985).

57. Daubert v. Merrell Dow Pharm., Inc., 727 F. Supp. 570, 571 (S.D. Cal. 1989), *vacated*, 509 U.S. 579 (1993). The separate cases involved Jason Daubert and Eric Schuller. They were consolidated for purposes of the summary judgment motion and subsequent proceedings.

58. *Id.* at 573–75.

59. *Id.* at 575–76.

however, is important to what emerged in the Supreme Court in *Daubert* and was new to Bendectin litigation.

The court wrote about the limitations imposed by Rule 703 of the Federal Rules of Evidence: "A necessary predicate to the admission of scientific evidence is that the principle upon which it is based 'must be sufficiently established to have general acceptance in the field to which it belongs.'"[60] The language quoted came from a case in the Ninth Circuit.[61] While the *Daubert* court did not cite *Frye*, its general acceptance standard is undisguised. Not surprisingly, the case from which the quoted language came was a criminal case (which did cite *Frye*),[62] rather than a prior Bendectin case or another civil case.

The Ninth Circuit Court of Appeals, picking up on the general acceptance suggestion from the lower court, wrote a brief opinion that relied on *Frye*'s general acceptance standard.[63] Here, the scientific methodology or technique employed by plaintiffs' experts must be recognized by others in the field as reliable, and if it is not, the expert's opinion is inadmissible.[64]

Thus, the Ninth Circuit's opinion starkly presented the following question that the Supreme Court agreed to resolve: Did *Frye* and its general acceptance standard survive the adoption of the Federal Rules of Evidence? That the issue reached the Supreme Court in that form is critical not only to the Supreme Court's agreeing to enter the fray,[65] but also to the form that the Court's opinion ultimately took: how shall we cabin the excesses produced by liberalized expert witness rules contained in the Federal Rules of

60. *Id.* at 572 (quoting United States v. Kilgus, 571 F.2d 508, 510 (9th Cir. 1978) (quoting United States v. Brown, 557 F.2d 541, 556 (6th Cir. 1977) (quoting Frye v. United States, 293 F. 1013, 1014 (D.C. Cir. 1923)))).

61. *Kilgus*, 571 F.2d at 510.

62. *Brown*, 557 F.2d at 556.

63. Daubert v. Merrell Dow Pharm., Inc., 951 F.2d 1128, 1129–30 (9th Cir. 1991).

64. *Id.* at 1130. Focusing on the epidemiologic record for Bendectin, the court reasoned that plaintiffs' experts who relied on reanalyzing the published studies and found a sufficient association to support an opinion that causation existed had not employed a proper methodology. *Id.* at 1130–31. The plaintiffs' reanalysis of epidemiological studies fell short because of a lack of peer review and publication. *Id.* at 1131. With all of their experts' opinions excluded, plaintiffs could not demonstrate causation and therefore summary judgment was affirmed. *Id.*

65. Absent the *Frye* question, it seems unlikely the Supreme Court would have granted certiorari. The Court had refused to grant certiorari in several earlier Bendectin cases and refused to review two other court of appeals decisions after its granting certiorari in *Daubert*. *See* GREEN, *supra* note 12, at 309. In general, the Court expressed no interest in the specifics of the evidence involved in its *Daubert* opinion. *See id.*

Evidence[66] and the growing conviction that adversarial expert witnessing was spiraling out of control?[67]

As is well known, the Supreme Court rose to the admissibility bait and provided standards for determining when an expert's testimony is of sufficient reliability to justify its admission. Instead of adopting *Frye* or relying on Rule 703 of the Federal Rules of Evidence, which addresses the legitimate bases on which an expert may base an opinion, the *Daubert* court found its standards grounded in language in Rule 702, which refers to experts with "scientific . . . knowledge."[68] To be considered scientific knowledge, the Court explained that valid reasoning and methodology must be employed.[69] Four nonexclusive factors were provided: (1) peer review and publication; (2) the known or potential rate of error; (3) general acceptance; and (4) testing a theory by attempting to find evidence to disprove it (falsification).[70]

A great deal of thought and ink has been devoted to understanding *Daubert,* its implications, and how lower courts should apply it. With the benefit of more than twenty years of experience with *Daubert* in toxic substances litigation, we have a modestly different perspective about what *Daubert* has wrought, yet one that, immodestly, is quite illuminating and reveals that *Daubert* inquiries are rather similar to a task that courts have confronted alongside the civil jury throughout modern legal history. The primary role for *Daubert's* four-factor test is to provide amorphous standards that can be pressed into the service of evaluating the sufficiency of the scientific evidence proffered by plaintiffs' experts or the entire body of scientific evidence in the record, including that proffered by defendants to undertake the sort of sufficiency analysis

---

66. Common law treatment of experts and the restrictions that impeded their admissibility and usefulness are described in Jack B. Weinstein, *Improving Expert Testimony*, 20 U. RICH. L. REV. 473, 475–76 (1986).

67. Peter Huber, if not single-handedly responsible for concerns about the quality of expert testimony, surely beat the loudest drum on this subject. *See* PETER HUBER, GALILEO'S REVENGE: JUNK SCIENCE IN THE COURTROOM 192–213 (1991).

68. Daubert v. Merrell Dow Pharm., Inc., 509 U.S. 579, 580, 589–90 (1993); *see also* FED. R. EVID. 702 (providing the qualifications for the admission of expert witness testimony).

69. *Id.* at 592–93. Finding its screening authority in the phrase "scientific knowledge" is a bit of a stretch. Science is less about knowledge than about accepted processes for conditionally accepting theories or explanations of natural phenomena as true. And, as we explain in the text, those processes are rarely invoked to answer the kinds of questions that the law requires in a tort case: whether the plaintiff's exposure to the defendant's toxic substance caused the disease for which the plaintiff seeks to recover. While some scientific disciplines address other matters that bear circumstantially on that question, none inquire into the ultimate factual issue of individual causation.

70. *Id.* at 593–94.

conducted by Judge Thomas Jackson in *Richardson v. Richardson-Merrell, Inc.*[71]

We believe our insight in describing what courts actually do in *Daubert* hearings without always appreciating what they are doing could be quite helpful. *Daubert's* factor approach is not well suited for evaluating scientific evidence or expert testimony drawn from that evidence. Recognition that what courts are doing in *Daubert* determinations is assessing the sufficiency of the scientific evidence could have a variety of implications, several of which we sketch out in the conclusion. Our focus in this Article is descriptive: we seek to explain how the *Daubert* inquiry in toxic tort litigation has evolved from *Daubert's* conception of an assessment of the expert's methodology and reasoning to an assessment of whether the scientific record is sufficient to permit an inference of negligence. We confess, however, that normatively we applaud this reconceptualization of what *Daubert* is about, as it will provide greater coherency and illumination to the process of deciding which cases should be submitted to the jury and which should not.

III. POST-*DAUBERT*

Initially, we should explain that we think the factors identified in *Daubert* are not well designed to assess whether expert testimony of the sort employed in toxic substances cases should be admitted.[72] Recall that the four factors are whether a causal hypothesis has been tested, the error rate associated with a given technique or test, whether results have been subjected to peer review and publication, and whether the methods employed by the expert have achieved general acceptance. Note that the latter two factors are variations on the *Frye* theme. They ask the judge to look to the opinions of others as an aid to determine the admissibility of expert testimony. If *Daubert* does anything, however, it shifts the courts' emphasis away from deference to extrajudicial authorities and toward judicial assessment of the evidence.[73] Thus, greater emphasis has been placed on the first two factors—whether a causal

71. *See supra* text accompanying notes 18–28.

72. For a similar criticism of the *Daubert* factors for determining the admissibility of expert economic testimony in antitrust cases, see Andrew I. Gavil, *After* Daubert: *Discerning the Increasingly Fine Line Between the Admissibility and Sufficiency of Expert Testimony in Antitrust Litigation*, 65 ANTITRUST L.J. 663, 673–75 (1997).

73. Michael J. Saks, *Merlin and Solomon: Lessons from the Law's Formative Encounters with Forensic Identification Science*, 49 HASTINGS L.J. 1069, 1139 (1998) ("But perhaps the purpose of the rules is simply to hold up a target to the courts; call one the *Frye* target and the other the *Daubert* target. The *Frye* ideal says: do whatever the experts tell you to do. The *Daubert* ideal says: figure out the science yourself.").

argument has been tested and the error rate associated with a technique.[74]

The first factor, whether a theory has been tested, is the factor most relied on by courts and indeed is useful in assessing questions of general causation—does the substance harm anyone and, if so, at what dose?[75] Courts look to see if experts have employed accepted methodologies for determining general causation.[76] This inquiry begins with the scientific evidence relied on by the expert: epidemiology and in vivo and in vitro toxicologic studies. Even chemical structure analysis and adverse case reports, while often consigned to the bottom of the scientific probity barrel, are marginally relevant to the question of general causation. Ultimately, of course, judgment and inference is required as to whether the available research sufficiently supports a finding of causation.[77]

---

74. *See* John B. Meixner & Shari Seidman Diamond, *The Hidden* Daubert *Factor: How Judges Use Error Rates in Assessing Scientific Evidence*, 2014 WIS. L. REV. 1063, 1120–21. References to lack of publication and peer review are often added as a makeweight after a court concludes that there is no scientific evidence to support the challenged expert's opinion. *E.g.*, McMunn v. Babcock & Wilcox Power Generation Grp., Inc., Nos. 10-143, 10-368, 10-650, 10-728, 10-744, 10-908, 10-1736, 11-898, 11-1381, 12-1221, 12-1459, 2013 WL 3487560, at *28 (W.D. Pa. July 12, 2013) (discussing and criticizing an expert's lack of epidemiologic evidence and methodology in conducting a differential etiology, and observing that the expert's methodology had not been published nor generally accepted).

75. Of course, it is not that a theory is tested, but that the results of that testing process support the validity of the theory. Professor Susan Haack explains an egregious misunderstanding of this criterion, in which a court concluded that although evidence was proffered to show that a forensic test was unreliable, the court concluded that the fact that the validity of the forensic test had been subjected to testing was grounds for its admissibility. *See* Susan Haack, *Federal Philosophy of Science: A Deconstruction—and a Reconstruction*, 5 N.Y.U. J.L. & LIBERTY 394, 419 (2010). Haack documents other serious misuses of the testing factor in this article. *See id.* at 424, 424–25 n.130.

76. *E.g.*, Pritchard v. Dow Agro Scis., 705 F. Supp. 2d 471, 483–89 (W.D. Pa. 2010); Ranes v. Adams Labs., Inc., 778 N.W.2d. 677, 690–95 (Iowa 2010); Betz v. Pneumo Abex, LLC, 44 A.3d 27, 52–54 (Pa. 2012).

77. Milward v. Acuity Specialty Prods. Grp., Inc., 639 F.3d 11, 18 (1st Cir. 2011); RESTATEMENT (THIRD) OF TORTS: LIABILITY FOR PHYSICAL AND EMOTIONAL HARM § 28 cmt. c(1) (AM. LAW INST. 2010); *see* Sheldon Krimsky, *The Weight of Scientific Evidence in Policy and Law*, 95 AM. J. PUB. HEALTH (SUPP. 1) S129, S134 (2005). We do not mean to suggest that all evidence is of equal probative value or that an expert could reasonably rely on case report data alone as the basis of a positive causal opinion. Stating that one has followed a weight-of-the-evidence approach in arriving at a conclusion that causation exists is not by itself a sufficient response to a sufficiency challenge. *See* Michael D. Green, *Pessimism About* Milward, 3 WAKE FOREST J.L. & POL'Y 41, 63–64 (2013); Joseph Sanders, Milward v. Acuity Specialty Products Group: *Constructing and Deconstructing Science and Law in Judicial Opinions*, 3 WAKE FOREST J.L. & POL'Y 141, 172–76 (2012).

Error rate, on the other hand, has played a very limited role in judicial discussions of general causation. This is not because studies used to assess general causation do not contain error rates.[78] All research, indeed all decisions, has error rates. For example, epidemiologic studies are generally conducted on a sample of individuals and then the results are used to make inferences about the population at large. Statistical tests of significance are one aspect of this inferential process and are designed to assess the possibility that the results of the study are the result of random error. However, random error is only one aspect of a study's potential error rate. Even more important are systematic errors that are the result of imperfect study design. Many of these systematic errors are difficult to control and indeed may not even be known to the researcher.[79] They are largely unquantifiable. There is no realistic means to determine what the error rate is for epidemiology in general or a given study in particular.[80] It is not surprising, therefore, that error rate has played such a limited role in assessing questions of general causation, and that in many cases, its only mention is as one of the four factors with no further discussion or analysis of it by the court.[81] In other cases, courts throw in the observation that the expert has not identified an error rate as a makeweight to a decision already made to exclude the expert's opinion.[82]

---

78. *See* NATE SILVER, THE SIGNAL AND THE NOISE: WHY SO MANY PREDICTIONS FAIL—BUT SOME DON'T 451–52 (2012).

79. Given this problem, few scientists would be confident in reaching conclusions based on one or two studies. Only through replication, using various designs and methods, do scientists gain confidence that a hypothesis has been confirmed or disproven.

80. *Cf.* Douglas L. Weed, *Evidence Synthesis and General Causation: Key Methods and an Assessment of Reliability*, 54 DRAKE L. REV. 639, 649 (2006) ("Turning to testability and rates of error, the overall reliability of any scientific method can be assessed by comparing its results to those of a better method known as a 'gold standard.' When such a standard methodology exists, then rates of error can, in principle, be calculated.").

81. *See, e.g.*, Perlman v. Universal Restoration Sys., Inc., No. 09-4215, 2013 WL 5278211, at *7–8 (E.D. Pa. Sept. 19, 2013) (identifying error rate as a *Daubert* factor but failing to apply it to the challenged expert's testimony); Mallozzi v. EcoSMART Techs., Inc., No. 11-CV-2884, 2013 WL 2415677, at *4–8 (E.D.N.Y. May 31, 2013) (same); Johns v. Bayer Corp., No. 09CV1935, 2013 WL 1498965, at *6 (S.D. Cal. Apr. 10, 2013) (same); Junk v. Obrecht, No. 12-1987, 2013 WL 47969433, at *3 (Iowa Ct. App. Sept. 5, 2013).

82. *See, e.g.*, Anderson v. Ford Motor Co., 950 F. Supp. 2d 1217, 1222, 1224 (D. Utah 2013); Zellars v. NexTech Ne., LLC, 895 F. Supp. 2d 734, 738, 745 (E.D. Va. 2012). These experts cannot testify to an error rate—there simply is no gold standard by which to test the validity of the studies that address general causation. Error rate is neither known nor knowable. We should note that there are a few rather extreme cases in which the court does exclude an expert's argument based on potential error rates. *See* Heller v. Shaw Indus., Inc., 167 F.3d 146, 162–63 (3d Cir. 1999) (involving an error rate in a "back-extrapolation methodology").

Of course, proving general causation is not sufficient for a plaintiff to prevail. A plaintiff must also establish specific causation—the substance in question caused that plaintiff's injury. Here, the *Daubert* factors are of very limited use. Scientists generally devote themselves to the search for the effects of causes—does Bendectin cause birth defects?—but not the causes of effects—was Jason Daubert's birth defect caused by Bendectin?[83] An inquiry as to whether a plaintiff's disease was caused by exposure to a toxic agent is one in which scientists rarely engage,[84] and therefore, the first *Daubert* criteria is of limited use.

And what of error rate with regard to opinions on specific causation? This factor is often cited in criminal cases in which the issue is similar to the specific causation issue in toxic tort cases. For example, one might ask about the error rate associated with the assertion of a handwriting expert that a specific signature is a forgery. This assertion is akin to a physician testifying that a particular patient's disease is caused by a particular substance. Unfortunately, outside the context of litigation, this is an inquiry to which most doctors devote very little time. It is true that they frequently serve as expert witnesses in such cases because the law demands evidence on this issue, but in most areas of medicine, there is no accepted scientific methodology for determining the cause of an individual's disease and, therefore, the error rate is simply unknown and unquantifiable.[85]

---

83. For useful discussions of this distinction, see A. Philip Dawid, David L. Faigman & Stephen E. Feinberg, *Fitting Science Into Legal Contexts: Assessing Effects of Causes or Causes of Effects?*, 43 SOC. METHODS & RES. 359, 360–61 (2014); A. Philip Dawid, *The Role of Scientific and Statistical Evidence in Assessing Causality*, *in* PERSPECTIVES ON CAUSATION 133–48 (Richard Goldberg ed., 2011); David L. Faigman, John Monahan & Christopher Slobogin, *Group to Individual (G2i) Inferences in Scientific Expert Testimony*, 81 U. CHI. L. REV. 417, 424 (2014).

84. As Professor Ed Cheng explained, scientific facts are "general truths, they ideally should apply consistently from one case to another." Edward K. Cheng, *Scientific Evidence as Foreign Law*, 75 BROOK. L. REV. 1095, 1099 (2010).

85. There are exceptions. Allergists do devote considerable effort to devising methods to detect the causes of allergic reactions. *See* A.D.A.M., Inc., Nat'l Insts. of Health, *Allergy Testing—Skin*, MEDLINEPLUS, http://www.nlm.nih.gov/medlineplus/ency/article/003519.htm (last updated May 18, 2014). However, most oncologists are not interested in what substance caused a patient's lung cancer because the cause frequently is irrelevant to clinical treatment.

The process by which medical experts assess causes is often described by courts as a "differential diagnosis," but this is a misnomer. *See* Sanders, *supra* note 5, at 1375. The correct term is differential etiology, which simply reflects the logical proposition that if there are multiple alternative causes for a phenomenon, the more of the potential alternatives that can be ruled out, the more likely that one of the remaining ones explains the phenomenon. *Id.* at 1375 n.34. This is not science or medicine, just basic logic.

The underlying point is that the criteria crafted by the *Daubert* court for assessing the validity of an expert's methodology (or conclusions, after *General Electric Co. v. Joiner*) are of little help. At least with regard to the issue of a substance's toxicity—general causation—the generally accepted approach is to examine all of the scientific evidence and then make a judgment, appropriately weighting the individual studies about toxicity.[86] Testing and error rate are little more than heuristics that capture only part of an expert's overall assessment of the support for a general causation assertion and appear to be even less useful in answering questions of specific causation.[87] In sum, the *Daubert* factors are of limited assistance in resolving questions of specific causation.

The nature of this problem is captured in the remand opinion in *Daubert* itself. The Ninth Circuit proceeded to apply the new *Daubert* framework and concluded that the plaintiffs' experts' testimony was inadmissible.[88] The court employed but one of the factors from the Supreme Court—peer review and publication—to conclude that opinions based on reanalyses of published studies that themselves were not published were inadmissible.[89] With regard to other experts who proposed to testify to causation, the court took a different tack. Even if Bendectin were a teratogen, plaintiffs would have to show that it more than doubled the risk of their birth defects in order to satisfy the preponderance-of-the-evidence standard.[90] The best way to interpret this approach is to say that

---

86. *See* Weed, *supra* note 80, at 644–46.

87. In making that judgment, it "makes little sense to ask whether the technique employed 'can be (and has been) tested' . . . or what its 'known or potential rate of error' might be . . . ." Daubert v. Merrell Dow Pharm., Inc., 43 F.3d 1311, 1317 (9th Cir. 1995) (citations omitted) (quoting Daubert v. Merrell Dow Pharm., Inc., 509 U.S. 579, 580 (1993)).

One should note that when a body of epidemiology demonstrates a very strong relationship between a substance and an injury, the epidemiology itself may provide telling data with respect to both general and specific causation. The example of asbestos and the asbestos "signature diseases" of asbestosis and mesothelioma is a case in point. *See* Joseph Sanders, Michael D. Green & William C. Powers, Jr., *The Insubstantiality of the "Substantial Factor" Test for Causation*, 73 MO. L. REV. 399, 401–02 (2008).

88. *Daubert*, 43 F.3d at 1319, 1322.

89. *Id.* at 1318–19.

90. *Id.* at 1321. The idea behind this point of view is as follows. If a plaintiff can show more than a doubling of the risk associated with some exposure, then the plaintiff has provided enough evidence to support a finding of specific causation. For example, if the background risk of a disease among people who were not exposed is 10 in 100,000 and the risk of the disease among those exposed to the suspected cause is 25 in 100,000, we still do not know which of these 25 would have gotten sick anyway, even had they not been exposed. However, for each of these 25, it is more likely than not that their injury was the result of an exposure because 15 of the 25 injuries—more than half—is associated with exposure.

the court concluded that even if there was evidence that Bendectin increased the risk of birth defects some modest amount, plaintiffs still could not prevail. Although such evidence supports the argument that general causation exists, it fails to provide *sufficient* evidence of specific causation.

We could understand a court ruling that that evidence was insufficient for the jury to find in the plaintiff's favor, as did the court in *Turpin v. Merrell Dow Pharmaceuticals, Inc.*[91] But we find a ruling denying admissibility, even if the evidence is wrapped within an expert's opinion, incongruous. We think that the best explanation for ruling as a matter for law against plaintiffs is that a substantial body of epidemiologic evidence tended to exonerate Bendectin as a teratogen,[92] and the contrary evidence that the plaintiffs' experts were able to muster simply was not enough for a jury to find for the plaintiff. Thus, we think that with a mature body of epidemiologic evidence, Bendectin became like the *O'Connor* case described above: the defendant's evidence was sufficiently powerful that the meager evidence that the plaintiffs relied on was insufficient to support a jury verdict in their favor.[93]

This, of course, is the approach adopted by the *Turpin* opinion referred to above. It too specifically addressed the testimony of a medical doctor who was the only one prepared to state that the plaintiff's birth defect was caused by Bendectin. After surveying the scientific evidence, including epidemiology and toxicology, and explaining why it was inadequate, the court turned to the medical doctor's proposed testimony and observed that there was "no understandable scientific basis . . . stated" by him to support his

---

Note that not all courts require a doubling of the risk in order to prove specific causation, but few courts would conclude that substantial epidemiologic evidence of more than a doubling of risk is insufficient to prove specific causation. For a discussion of this issue, see Jon Todd Powell, Comment, *How to* Tell the Truth *With Statistics: A New Statistical Approach to Analyzing the Bendectin Epidemiological Data in the Aftermath of* Daubert v. Merrell Dow Pharmaceuticals, 31 HOUS. L. REV. 1241, 1289 (1994).

91. 959 F.2d 1349, 1349–50 (6th Cir. 1992).

92. That scientific evidence was not quite as strong as some claimed and certainly did not prove that Bendectin was "safe," as its manufacturer claimed, if by safe we mean that there is evidence that eliminates all possibility that it causes birth defects, however infrequently. Although early Bendectin litigation involved many different types of birth defects, most of the individual cases, including the *Daubert* case, involved limb-reduction defects. These injuries are difficult to study because they are quite rare. The meta-analyses performed on the Bendectin limb-reduction studies that exist indicate a relative risk of approximately 1.1, which is a slight increase over background risk. *See* SANDERS, *supra* note 12, at 76 (1998). On the other hand, the animal studies that do exist fail to find any relationship between Bendectin and this type of injury. *See id.* at 61–89.

93. Another case congregation similar to Bendectin was silicone gel breast implants, at least in the later stages of that litigation. *See infra* text accompanying notes 112–21.

opinion.[94] Then, the court concluded, "The analytical gap between the evidence presented and the inferences to be drawn on the ultimate issue of human birth defects is too wide. Under such circumstances, a jury should not be asked to speculate on the issue of causation."[95] The proposition is simple and the one we urge: the scientific evidence, entirely apart from the expert's testimony, is insufficient to justify a reasonable inference of causation. At least in cases where the exclusion of experts results in a summary judgment, we believe that this is a preferable approach.

We are not alone in this opinion. The next two Supreme Court cases in the so-called *Daubert* trilogy, *General Electric Co. v. Joiner* and *Kumho Tire Co. v. Carmichael,* tend to support this perspective by moving away from an emphasis on the *Daubert* factors and toward an emphasis on the concept of "fit."[96]

In *Joiner*, the Court concluded that the same "abuse of discretion" standard should be employed by appellate courts regardless of whether the trial court admitted or excluded expert testimony.[97] More importantly for our analysis, the Court revisited a statement in *Daubert* that drew a bright line between methodology and conclusions:

> Respondent points to Daubert's language that the "focus, of course, must be solely on principles and methodology, not on the conclusions that they generate." He claims that because the District Court's disagreement was with the conclusion that the experts drew from the studies, the District Court committed legal error and was properly reversed by the Court of Appeals. But conclusions and methodology are not entirely distinct from one another. Trained experts commonly extrapolate from existing data. But nothing in either Daubert or the Federal Rules of Evidence requires a district court to admit opinion evidence that is connected to existing data, only by the ipse dixit of the expert. A court may conclude that there is simply too great an analytical gap between the data and the opinion proffered. See Turpin v. Merrell Dow Pharmaceuticals, Inc., 959 F.2d 1349, 1360 (C.A.6), cert. denied, 506 U.S. 826, 113 S. Ct. 84, 121 L.E.2d 47

---

94. *Turpin*, 959 F.2d at 1360.

95. *Id.* at 1360–61.

96. Kumho Tire Co. v. Carmichael, 526 U.S. 137, 153–54 (1999); Gen. Elec. Co. v. Joiner, 522 U.S. 136, 146 (1997).

97. *Joiner*, 522 U.S. at 146.

> A court of appeals applying "abuse-of-discretion" review to such rulings may not categorically distinguish between rulings allowing expert testimony and rulings disallowing it. We likewise reject respondent's argument that because the granting of summary judgment in this case was "outcome determinative," it should have been subjected to a more searching standard of review.

Id. at 142–43 (citations omitted).

> (1992).[98] That is what the District Court did here, and we hold that it did not abuse its discretion in so doing.[99]

The *Joiner* opinion went beyond simply stating this as a general proposition. Chief Justice Rehnquist's opinion discusses and critiques animal studies and epidemiologic research cited by the plaintiff as supporting the position that exposure to polychlorinated biphenyls ("PCBs") either caused or "promoted"[100] the plaintiff's lung cancer.[101] According to Justice Rehnquist, the epidemiology studies did not support the plaintiff's position, and he ended the opinion by saying:

> We further hold that, because it was within the District Court's discretion to conclude that the studies upon which the experts relied were not sufficient, whether individually or in combination, to support their conclusions that Joiner's exposure to PCB's contributed to his cancer, the District Court did not abuse its discretion in excluding their testimony.[102]

The Supreme Court granted certiorari in *Kumho Tire Co. v. Carmichael* to make it clear that the *Daubert* admissibility standard

---

98. Following is the language in *Turpin* to which the *Joiner* court was referring:

> We cannot find, however, that [Dr. Palmer's] testimony is anything more than a personal belief or opinion. . . . Upon analysis, we conclude that Dr. Palmer's conclusions go far beyond the known facts that form the premise for the conclusion stated. This conclusion so overstates its predicate that we hold that it cannot legitimately form the basis for a jury verdict.

Turpin, 959 F.2d at 1360.

99. *Joiner*, 522 U.S. at 146 (citations omitted).

100. A cancer promoter is an agent that is not capable of inducing cellular mutagenesis but can either increase the capacity of another agent to induce mutagenesis or speed the cancer process along once mutagenesis has occurred. *See* ROBERT A. WEINBERG, ONE RENEGADE CELL: HOW CANCER BEGINS 61 (1998); Raymond Tennant, *What Is a Tumor Promoter?*, 107 ENVTL. HEALTH PERSP. A390, A390–91 (1990).

Something like this is what the *Joiner* expert had in mind. The promotion argument was introduced by one of the plaintiff's experts who testified:

> It [was] more likely than not, given Mr. Joiner's limited tobacco use, and also considering his second hand tobacco smoke exposure, and given his age at the onset of lung cancer, 37 years, that tobacco smoke served only as the initiator of the cancer and that some other agent served as the promoter of the initiated cells. It was the promotion of these initiated cells which caused Mr. Joiner to be harmed.

Joiner v. Gen. Elec. Co., 864 F. Supp. 1310, 1314 (N.D. Ga. 1994). But for the promotion effect of his PCB exposure, his cancer "would not have developed for many years, if at all." *Joiner*, 522 U.S. at 139–40. "Accelerator" (of the initiation of the disease) may be a more accurate term than promoter.

101. *Joiner*, 522 U.S. at 144–45.

102. *Id.* at 146–47.

applied to all expert testimony.[103] The plaintiffs in *Kumho Tire* were injured when the right rear tire on their minivan failed.[104] They claimed it did so because it was defectively manufactured.[105] Through a rather complex line of reasoning, the plaintiff's expert concluded that the tire failure occurred because of a manufacturing defect.[106] Here is the gist of the expert's argument:[107] A tire's carcass should stay bound to the inner side of the tread for a significant period of time after its tread depth has worn away. The tread of the tire at issue had separated from its inner steel-belted carcass prior to the accident. This "separation" caused the blowout. And what caused the separation? A separation may be caused by a type of tire misuse called "overdeflection," which consists of underinflating the tire or causing it to carry too much weight, thereby generating heat that can undo the chemical tread/carcass bond. If the tire has not been subjected to this type of misuse, then the cause of a separation is a tire defect. A visual and tactile inspection of the tire reveals the tire had not been overdeflected, and therefore, it must have been defective.

Without ever using the word "fit," the opinion reinforced the *Joiner* approach. Admissibility analyses should focus on "the case at hand," not on broad general principles and theories.[108] Mr. Carlson's testimony was properly excluded because although the general method he employed may be reliable, his application of the method was not. His application could not support his conclusion just as the studies cited in *Joiner* could not support the expert conclusions in that case.

As we read admissibility cases, especially the cases that come after *Joiner* and *Kumho Tire*, the *Daubert* factors seem to play less and less of a role, and are supplanted by an increasing focus on fit and the existence of large analytical gaps in reasoning.[109] Courts

---

103. Kumho Tire Co. v. Carmichael, 526 U.S. 137, 141, 147 (1999). The Eleventh Circuit had held that it applied only to "scientific principle." *Id.* at 151.

104. *Id.* at 142.

105. *Id.*

106. *Id.* at 142–45.

107. *Id.*

108. *Id.* at 153.

109. *See, e.g.*, McEwen v. Balt. Wash. Med. Ctr., Inc., No. 09-2141, 2010 WL 5129873, at *1 (4th Cir. 2010) (per curiam) ("[A]bsent any reliance on or support from the relevant medical literature, in effect the only basis for the experts' conclusions was 'a[n] ipse dixit statement of a clinician saying that I think causation has been proved, which is simply not sufficient as a matter of law.'"); Heller v. Shaw Indus., Inc., 167 F.3d 146, 156 (3d Cir. 1999) ("[R]eliable methods for making a diagnosis cannot sanitize an otherwise untrustworthy conclusion."); Milward v. Acuity Specialty Prods. Grp., Inc., 969 F. Supp. 2d 101, 114 (D. Mass. 2013) ("Here, it remains a mystery exactly what evidence Butler's 'clinical approach' weighs without the benefit of established, tested, and reliable methods of analyzing specific causation discussed above—including

that adopt this approach are asking the plaintiffs' experts to lay out the scientific bases for their opinion.[110] At the end of the day, the question is what scientific evidence has been proffered in support (and against) factual causation. This idea is at least partially enshrined in the revised Federal Rule of Evidence 702 that was adopted in the aftermath of the *Daubert* trilogy and was intended to reflect the rule in those cases. The Rule now reads:

> Rule 702. Testimony by Expert Witnesses
>
> A witness who is qualified as an expert by knowledge, skill, experience, training, or education may testify in the form of an opinion or otherwise if:
>
> **(a)** the expert's scientific, technical, or other specialized knowledge will help the trier of fact to understand the evidence or to determine a fact in issue;
>
> **(b)** the testimony is based on sufficient facts or data;
>
> **(c)** the testimony is the product of reliable principles and methods; and
>
> **(d)** the expert has reliably applied the principles and methods to the facts of the case.[111]

Especially worthy of note is the requirement that testimony is based on sufficient facts or data.

---

differential etiology, safe threshold, and relative risk analyses. Butler's 'clinical assessment' or 'qualitative approach' thus appears to do nothing more than connect Milward's estimated benzene exposure to his APL 'by the *ipse dixit* of the expert.'"); LeBlanc v. Chevron USA, Inc., 513 F. Supp. 2d 641 (E.D. La. 2007), *vacated and remanded*, No. 07-30599, 2008 WL 1805448 (5th Cir. 2008); Edwards v. Safety-Kleen Corp., 61 F. Supp. 2d 1354 (S.D. Fla. 1999); Zabilansky v. Am. Bldg. Restoration Prods., Inc., No. 200101985, 2004 WL 2550458 (Mass. Super. Ct. Oct. 13, 2004), *aff'd*, No. 05-P-948, 2006 WL 2520228 (Mass. Ct. App. 2006).

110. *See, e.g.*, Moore v. Ashland Chems., Inc., 151 F.3d 269, 275 (5th Cir. 1998); Rosen v. Ciba-Geigy Corp., 78 F.3d 316, 318 (7th Cir. 1996); Porter v. Whitehall Labs., Inc., 9 F.3d 607, 615–16 (7th Cir. 1993). As the Second Edition of the Reference Manual on Scientific Evidence counsels, "[T]he physician expert witness must determine if the medical and research literature supports a determination of environmental causation." Mary Sue Henifin et al., *Reference Guide on Medical Testimony*, *in* REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 439, 473 (2d ed. 2000). The guide goes on to list courts that require the expert to provide the court with the literature on which the expert relies. *Id.* at 473–74 n.126. A few cases are identified that do not require proffering scientific literature, but those are cases in which the courts determined that other evidence—including challenge, dechallenge, rechallenge, and biological mechanism evidence, apart from formal studies examining causation—was sufficient to permit testimony about causation and for the jury to find that causation exists. *Id.*

111. FED. R. EVID. 702.

This trend toward concern about the scientific evidence supporting causation and its sufficiency is reflected in another mass tort—the silicone gel breast implant litigation. By the time that litigation reached maturity, a substantial body of epidemiology had developed, and it revealed there was no connection to the autoimmune disorders plaintiffs were alleging. *Norris v. Baxter Healthcare Corp.*[112] demonstrates the point that we are trying to make. The plaintiff received implants manufactured by the defendant and then had them removed four years later.[113] After she developed a systemic autoimmune disease, she sued the manufacturer of her breast implants.[114] The district court granted summary judgment, and the plaintiff appealed.[115] The focus of the court of appeal's opinion is on the scientific evidence supporting the existence of general causation. One needs to read deep into the opinion before finding any discussion of the plaintiff's experts or the admissibility of their opinions. Indeed, the district court's grant of summary judgment was based on the scientific evidence of record, rather than on a *Daubert* hearing or even *Daubert* considerations other than that the court felt that an assessment of the scientific evidence was required as a preliminary matter under *Daubert*.[116]

Echoing the Bendectin cases, the Tenth Circuit explained that the large body of epidemiologic evidence (proffered by the defendant), in failing to find an association between silicone gel breast implants and systemic disease, suggested that general causation did not exist.[117] As in *Richardson*, there was the judgment of an independent expert body—the Institute of Medicine of the National Academy of Sciences—that the research revealed no association between implants and disease.[118] Distinguishing other cases in which substantial epidemiologic evidence did not exist, the court explained that, in the face of the epidemiologic record for breast implants, the plaintiff had a higher burden to produce supporting evidence.[119] Five pages into its opinion, the court of appeals turned to the admissibility of the plaintiff's expert's opinions

---

112. 397 F.3d 878 (10th Cir. 2005).

113. *Id.* at 880.

114. *Id.* She also suffered from and sued for localized injury due to leakage from her breast implants, but that issue was resolved on different grounds. *Id.*

115. *Id.*

116. *Id.* at 883–84 n.2.

117. *Id.* at 884–86.

118. *Id.* at 882. The lack of an association means that the incidence of systemic disease in women with breast implants is no different from the incidence in women without breast implants. An association is a necessary, but not sufficient, condition for an inference of causation. The research is discussed in INST. OF MED. & COMM. ON THE SAFETY OF SILICONE BREAST IMPLANTS, SAFETY OF SILICONE BREAST IMPLANTS (Stuart Bondurant, Virginia Ernster & Roger Herdman eds., 1999).

119. *Norris*, 397 F.3d at 882.

and ruled, as was preordained from the discussion of the scientific evidence, that the district court properly excluded the testimony, concluding, "We cannot allow the jury to speculate based on an expert's opinion which relies only on clinical experience in the absence of showing a consistent, statistically significant association between breast implants and systemic disease."[120] To us, *Norris* is transparently a decision that the evidence contrary to the plaintiff's was so overwhelming that no reasonable jury could find otherwise,[121] and reference to the *Daubert* factors—general acceptance and lack of peer review—was merely window dressing to support the conclusion based on the weight of the scientific evidence.

Both the Bendectin and the silicone gel breast implant litigation are atypical because they address an issue for which there was a very substantial body of epidemiologic evidence. Most toxic tort cases address causal questions about which there is little or no epidemiologic evidence. The cases that we discuss below comprise this genre, but we confess that they are not randomly selected. We chose them because they well illustrate the sufficiency-of-the-evidence approach that we believe is prevalent in deciding *Daubert* motions.[122] But the fact that some courts are examining the sufficiency of the scientific evidence rather than the methodology of the expert is useful, we believe, both for reform in the role of judge and jury in toxic substances litigation and for greater self-consciousness by courts confronting these issues.[123]

We begin with *Allen v. Pennsylvania Engineering Corp.*,[124] a case in which the plaintiffs' decedent contracted brain cancer after occupational exposure to ethylene oxide.[125] Plaintiffs' experts proposed to testify that the ethylene exposure caused decedent's

120. *Id.* at 887.

121. *See also, e.g.*, Grant v. Pharmative, LLC, 452 F. Supp. 2d 903, 908 (D. Neb. 2006) ("Dr. Corbett's failure to adequately address the body of contrary epidemiological evidence weighs heavily against admission of his testimony.").

122. In this Article, we have generally resisted critiquing the reasoning of courts in their assessments of the scientific evidence and its sufficiency. Instead, our focus is on the fact that their methodology was to examine the scientific evidence and determine if it was sufficient to permit admissibility of the proffered expert's testimony or, in our view, sufficient to permit the jury to find causation.

123. We set aside cases where the plaintiff's experts simply had not done their homework before the *Daubert* challenge to their testimony. These cases are equivalent to the plaintiff's lawyer who rests upon being given the opportunity to present her case. The result would be judgment for the defendant, with the reasoning that the plaintiff has failed to satisfy her burden of producing sufficient evidence. Our impression is that the incidence of those cases has declined as plaintiffs' lawyers have learned the perils of delaying their experts' preparation.

124. 102 F.3d 194 (5th Cir. 1996).

125. *Id.* at 195.

brain cancer.[126] They presented "suggestive" epidemiologic evidence—in vivo rat studies—and a bit of biologic mechanism theory to support their "weight of the evidence" testimony.[127] We find the court's summary of its views instructive: "On examination, none of the scientific data on which appellants' experts rely furnishes a scientifically valid basis for the conclusion they would draw. The paucity of epidemiological evidence, the unreliability of animal studies, and the inconclusiveness of cell biology combine to undercut the expert testimony."[128]

The "methodology" of the court in reviewing the admissibility of the expert's testimony was to consider the scientific evidence proffered by the experts, assess its probity, compare that evidence to other evidence that it appears was presented by the defendants (including two epidemiologic studies that failed to find an association between ethylene oxide and brain cancer), and conclude that the evidence was insufficient to support an inference, whether by the plaintiffs' experts or the jury, that causation existed.[129] The probity assessment included the rejection of "suggestive" epidemiology:[130] suggestiveness did not meet the threshold set forth previously in *Brock v. Merrell-Dow Pharmaceuticals, Inc.*,[131] which required statistically significant epidemiologic evidence.[132] Animal studies that were not replicated in another rodent species provided "at best speculative support for [the plaintiffs' experts'] causation theory."[133] We note that "speculation" describes the line between permissible inferences from circumstantial evidence and those that

126. *Id.* at 196.

127. *Id.* at 196–97.

128. *Id.* at 198.

129. *Id.* at 196–98. Thus, the court rejected studies that found brain tumors in rats exposed to ethylene oxide because the same effect was not replicated in mice studies, thereby raising questions about generalizing the rat study results to other species. The court did this without any reference to a methodology for extrapolating animal studies to humans. *See* Bernard D. Goldstein & Mary Sue Henifin, *Reference Guide on Toxicology, in* REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 633, 646–47 (3d ed. 2011).

130. *Allen*, 102 F.3d at 197.

131. 874 F.2d 307 (5th Cir. 1989), *modified*, 884 F.2d 166 (5th Cir. 1989), *cert. denied*, 494 U.S. 1046 (1990). Unfortunately, the court ignored a critical difference between *Brock* and *Allen*. By the time that *Brock* was decided, a substantial body of epidemiologic evidence existed that tended to exonerate Bendectin as a teratogen. Based on that evidence, the *Brock* court required the plaintiff to provide statistically significant epidemiology in order to proceed. *Id.* at 312. The two epidemiology studies identified by the court in *Allen* appear to be inconclusive rather than exonerative. *Allen*, 102 F.3d at 197 ("[T]he study's 'findings do not provide evidence for a positive association between exposure to [ethylene oxide] and cancers of the . . . brain . . . .'" (quoting L. Stayner et al., *Exposure-Response Analysis of Cancer Mortality in a Cohort of Workers Exposed to Ethylene Oxide*, 138 AM. J. EPID. 787, 797 (1993))).

132. *Brock*, 874 F.2d at 312.

133. *Allen*, 102 F.3d at 197.

are too weak and, therefore, may not be submitted to or drawn by the jury.

A somewhat different case involved a court that analyzed each of the scientific studies identified by the plaintiffs' expert and methodically explained why each study was insufficient to permit a determination of causation.[134] Studies were eliminated because they did not investigate the specific disease suffered by the plaintiffs but only addressed a more general class of disease; they did not find a statistically significant association; their measure of exposure included exposure to a broader class of substances (for example, all organic solvents) rather than the alleged agent in the plaintiff's case (benzene); or they failed to measure the extent of the dose to which the exposed cohort was exposed.[135] With all of the studies eliminated on these grounds, the court proceeded to conclude that *Daubert* required exclusion of the expert's testimony, invoking the general acceptance factor for the proposition that the expert's causation opinion was not generally accepted.[136] The Fifth Circuit Court of Appeals summarized the district court's decision: "The district court based its exclusion of Dr. Levy's testimony on several findings. The district court acknowledged that Dr. Levy's methodology was 'unassailable,' but found deficiencies in the underlying data, namely the various studies and articles Dr. Levy relied upon for his research conclusions."[137]

Another class of cases involves questions of dosage. In *Nelson v. Tennessee Gas Pipeline Co.*,[138] the court noted, "With respect to the question of dose, plaintiffs cannot dispute that [their expert] made no attempt to determine what amount of PCB exposure the Lobelville subjects had received and simply assumed that it was sufficient to make them ill."[139] In *Wills v. Amerada Hess Corp.*,[140] the plaintiff advanced an oncogene or no-threshold theory of cancer development.[141] The court rejected the plaintiff's expert's testimony, saying, "[E]ven though benzene and [polycyclic aromatic hydrocarbons] have been shown to cause some types of cancer, it is

134. Knight v. Kirby Inland Marine, Inc., 363 F. Supp. 2d 859, 862–67 (N.D. Miss. 2005), *aff'd*, 482 F.3d 347 (5th Cir. 2007).

135. *See id.* at 864–66.

136. *Id.* at 866–67. We read the court's observation that the expert's opinion was one that did not permit a determination of an error rate as acknowledging its inapplicability rather than asserting that factor counted against admissibility. *Id.* at 866.

137. *Knight*, 482 F.3d at 351.

138. 243 F.3d 244 (6th Cir. 2001).

139. *Id.* at 252; *see also* Johnson v. Arkema, Inc., 685 F.3d 452, 472 (5th Cir. 2012); Cunningham v. Masterwear, Inc., No. 1:04-cv-1616-JDT-WTL, 2007 WL 1164832, at *5–6 (S.D. Ind. 2007).

140. No. 98CIV.7126(RPP), 2002 WL 140542 (S.D.N.Y. 2002), *aff'd*, 379 F.3d 32 (2d Cir. 2004).

141. *See id.* at *7–9.

too difficult a leap to allow testimony that says any amount of exposure to these toxins caused squamous cell carcinoma of the head and neck in the Decedent."[142] In a similar vein, in *Burleson v. Texas Department of Criminal Justice*,[143] the plaintiff claimed that exposure to radioactive particles while engaging in welding operations caused the plaintiff's cancer.[144] His expert presented a "radiation hot spot" theory that the primary risk for cancer is the local microscopic dose of radiation that is received by the one cell that transforms into cancer, not the total dose to the individual.[145] The Fifth Circuit affirmed the exclusion of the expert opinion, in part because there was too great an analytical gap between the data available to the expert and his conclusion.[146]

Interesting and illuminating dosage cases arise in the combination hormone replacement therapy ("CHRT") litigation. The most important study involved in that litigation is a clinical trial mandated by the U.S. Food and Drug Administration ("FDA") to

---

142. *Id.* at *15. No-threshold arguments have been the topic of a substantial body of recent asbestos litigation. In *Anderson v. Ford Motor Co.*, 950 F. Supp. 2d 1217 (D. Utah 2013), the court excluded this testimony, primarily because it was not based on sufficient facts or data:

> Plaintiffs [sic] experts do not base their opinions on scientific evidence that every exposure to asbestos causes mesothelioma. Instead, their testimony is based on their lack of information sufficient to show the level of exposure which does not create a risk of mesothelioma. This is not reliable enough evidence for the Court to allow it in under the standards of *Daubert* and Rule 702.

Id. at 1224.

143. 393 F.3d 577 (5th Cir. 2004).

144. *Id.* at 581.

145. *Id.* at 582.

146. *Id.* at 587. The *Allen* case, discussed above, may also be seen as a dosage case. In another part of the opinion, the court stated, "The experts actually knew more about Allen's exposure to [ethylene oxide] through his smoking a pack of cigarettes a day than they did about his occupational exposure to the chemical." Allen v. Pennsylvania Eng'g Corp., 102 F.3d 194, 198 (5th Cir. 1996); *see also* Cano v. Everest Minerals Corp., 362 F. Supp. 2d 814, 848 (W.D. Tex. 2005).

The Fifth Circuit has repeatedly recognized that "[s]cientific knowledge of the harmful level of exposure to a chemical, plus knowledge that the plaintiff was exposed to such quantities, are minimal facts necessary to sustain the plaintiffs' burden in a toxic tort case." Curtis v. M & S Petroleum, Inc., 174 F.3d 661, 670 (5th Cir. 1999) (quoting *Allen*, 102 F.3d at 199). Further, if an expert's "causation opinion was not based on sufficient information of the level of [the agent] to which Plaintiffs were exposed, his methodology would not be reliable, rendering his causation opinion inadmissible." *Curtis*, 174 F.3d at 671; *see also* Moore v. Ashland Chem., Inc., 151 F.3d 269, 278 n.10 (5th Cir. 1998) ("Given the paucity of facts Dr. Jenkins had available about the level of Moore's exposure to the Toluene solution, his causation opinion would have been suspect even if he had scientific support for the position that the Toluene solution could cause [Reactive Airways Dysfunction Syndrome] in a worker exposed to some minor level of the solution.").

determine the efficacy and safety of CHRT (Prempro).[147] The study was terminated early because of serious adverse effects, including breast cancer, among those exposed.[148] The study, however, did not find an increased risk of breast cancer until women had completed three years of Prempro therapy.[149]

In *Kuhn v. Wyeth*,[150] one of the plaintiffs had taken Prempro for only one year and nine months.[151] The plaintiffs' expert, addressing general causation, claimed that short-term—less than three years—exposure increased the risk of breast cancer.[152] The defendant's *Daubert* motion was granted by the trial court, and the plaintiff appealed.[153]

The Eighth Circuit explained that the plaintiffs' expert based his opinion on three observational epidemiologic studies that found an increased risk of breast cancer with short-term exposure to CHRT.[154] Those studies had their identifiable flaws, aside from the fact that observational studies may be subject to unidentified biases and confounding factors, problems that clinical trials with randomization avoid.[155] The plaintiffs' expert also claimed that the clinical trial was underpowered to reveal an increased risk of breast cancer with less than three years of exposure.[156] Moreover, the plaintiffs' expert criticized the clinical trial in its study design

147. *See* Kuhn v. Wyeth, Inc., 686 F.3d 618, 621 (8th Cir. 2012).

148. *See id.*

149. *See id.* at 622. Although there was no increased risk combining women who were exposed for less than three years, the plaintiffs' expert testified that there was an increased risk among those who had taken the drug for between two and three years, but that increase washed out because those with less than two years of exposure had a lower rate of breast cancer than the unexposed cohort. *See id.* at 626.

150. *Id.*

151. *Id.* at 620.

152. *See id.*

153. *See id.*

154. *Id.* at 628–30.

155. Among those flaws identified by the court were that the studies did not limit the CHRT agent to Prempro and that they failed reliably to determine length of exposure. *Id.* at 630–31. In addition, the defendants claimed there were other observational studies, ignored by the plaintiffs' expert, that did not find a short-term increased risk. *Id.* at 633.

156. *Id.* at 624. Unfortunately, the *Kuhn* court did not provide any power calculations. A critical matter in assessing the power of a study is the magnitude of an effect that the researcher seeks to identify. If, hypothetically, the Prempro trial was underpowered to identify an increase of risk of, say, ten percent, its lack of power at that magnitude of effect would not be valid criticism because increases of risk of that magnitude could not support a finding of specific causation. *See* Michael D. Green et al., *Reference Guide on Epidemiology*, *in* REFERENCE MANUAL ON SCIENTIFIC EVIDENCE, *supra* note 129, at 549, 582–83.

because it was not designed to assess the risk of short-term exposures to Prempro.[157]

The court concluded that despite the flaws in the epidemiologic studies relied on by the plaintiffs' expert, they provided an adequate basis for the plaintiffs' expert's opinion.[158] A passage in the court's opinion explained:

> Dr. Austin's testimony is admissible because the studies upon which he relied were sufficient to support his opinion that short-term use of Prempro increases the risk of breast cancer. Taken together, the Calle study and the foreign studies constitute appropriate validation of and good grounds for Dr. Austin's opinion. The studies' limitations may be presented to the jury, and Dr. Austin's reliance on the studies may be tested through the traditional means of cross examination and presentation of contrary evidence.[159]

To us, that sounds like a court that has assessed the scientific evidence proffered by the plaintiff and found it sufficient to support expert testimony and, more significantly, a jury finding of causation. Our assessment of the scientific evidence is that reasonable judges could reach different results on the question of whether the scientific evidence was sufficient to justify a finding of general causation for those exposed less than two years, as was the plaintiff in *Kuhn*.[160] Just as in traumatic injury cases—which do not require expert testimony—in which the inference of causation is neither strong nor weak, judges will differ in their assessments of this issue of sufficiency.[161]

Kuhn also reveals an interesting difference between methodology and sufficiency in *Daubert* determinations. The court rejected the defendant's claims that the plaintiffs' expert had "cherry picked" the epidemiologic studies on which he relied and ignored others that did not find short-term increased risks.[162] The court concluded that it should not choose between different expert opinions when both had scientific evidence supporting

---

157. *Kuhn*, 686 F.3d at 624. This critique included that study participants were older than the general population that might take CHRT, the latter being more susceptible to carcinogens; that the loss rate of forty percent of participants in the clinical trial who dropped out of the study threatened its validity; and that some study participants were misclassified as to whether they were in the exposed or unexposed cohort. *Id.* at 623.

158. *See id.* at 632.

159. *Id.*

160. On balance, we are inclined to disagree with the Eighth Circuit, largely because it ignored conflicting epidemiologic evidence that the plaintiffs' expert did not address. Both as a matter of *Daubert's* methodological focus and as a matter of assessing the strength of an inference of causation, all of the evidence should be considered.

161. *See infra* text accompanying notes 211–18.

162. *Kuhn*, 686 F.3d at 633.

them.[163] From a methodological perspective, the court's position is unjustified. The methodology employed by scientists in determining causation is through a process known as weight of the evidence.[164] The key point of that method, for our purposes, is that *all* of the evidence is considered in determining whether an inference of causation is justified. The plaintiffs' expert did not use that methodology, and that alone should have been sufficient to rule his opinion inadmissible based on *Daubert*. On the other hand, from a sufficiency of the evidence perspective, the plaintiff is not required to address contrary evidence while making a prima facie case. While a court may rule, as in *O'Connor*,[165] that the contrary evidence so overwhelms plaintiff's evidence, the *Kuhn* court's decision did not recognize that principle.

Still another group of cases involves attempts to draw conclusions based on inappropriate analyses of data sets. *Blackwell v. Wyeth*[166] is an example of this type of case. *Blackwell* is one of the cases in the thimerosal controversy. Thimerosal, which contained mercury, was used as a preservative in childhood vaccines until early in 2001. Many claims were made that thimerosal caused a vaccinated child's autism, no doubt spurred by the fact that mercury was known to have neurotoxic effects, although not of the autism variety. Several doctors—including Mark Geier, the plaintiff's expert in *Blackwell*—championed the theory that thimerosal was responsible for children's autism. Geier, in fact, published six articles purporting to support that theory.[167] But Geier's data came from sources like the national Vaccine Adverse Event Reporting System ("VAERS"), which relies on physicians' and parents' reports of adverse vaccine reactions. There are several difficulties with using data like that as the basis for an epidemiologic study, but the predominant one is that it is impossible to determine an incidence rate of adverse events for those who are vaccinated because an incidence rate requires not only the number suffering the adverse effect, but also a denominator reflecting the number who were exposed (vaccinated). A comparison of incidence rates of disease between those who are exposed and those who are not exposed is a central tenet of epidemiologic methodology.[168]

---

163. *Id.*

164. *See* CARL CRANOR, TOXIC TORTS: SCIENCE, LAW, AND THE POSSIBILITY OF JUSTICE 75–79 (2006).

165. *See supra* text accompanying notes 46–53.

166. 971 A.2d 235 (Md. 2009).

167. *See generally* PAUL OFFIT, AUTISM'S FALSE PROPHETS (2008) (explaining the irresponsibility of Geier and others who contributed to the thimerosal-causes-cancer crusade).

168. *See* Kenneth J. Rothman & Sander Greenland, *Measures of Disease Frequency*, *in* MODERN EPIDEMIOLOGY 29, 30–32 (Kenneth J. Rothman & Sander Greenland eds., 2d ed. 1998). *See generally* DAVID E. LILIENFIELD & PAUL D. STOLLEY, FOUNDATION OF EPIDEMIOLOGY 13–19 (3d ed. 1996) (explaining the

Blackwell was decided in Maryland, a jurisdiction that has continued to rely on the *Frye* test of admissibility. Nevertheless, in this situation, the court turned to the language in *Joiner* to justify rejecting the Geier testimony.[169] The VAERS data set is inadequate as the basis for a causal inference, and Geier's application of epidemiologic methods to the data does not alter this underlying fact.

> Generally accepted methodology, therefore, must be coupled with generally accepted analysis in order to avoid the pitfalls of an "analytical gap." Dr. Geier's faulty extrapolation from VAERS data, a potentially reliable source, manifests the *ipse dixit* identified in the *Joiner* opinion . . . . None of Dr. Geier's research aimed at establishing a link between thimerosal and autism, moreover, is based upon sound methodology.[170]

Although our primary focus in this Article is on toxic tort cases, we find some indications that our general point applies to other situations. We have already referred to the Supreme Court's *Kumho Tire* case above.[171] *Smith v. Cangieter*[172] is another automobile accident case. In *Smith*, the plaintiffs claimed that a four-wheel drive rental car was defectively designed.[173] There were two steps to the plaintiffs' expert's reasoning: (1) part-time four-wheel drive cars result in different wheels covering different distances, which results in slippage and loss of traction; and (2) that effect led to the driver (of the rental vehicle) becoming confused and losing control of the car, which went across a median and into oncoming traffic.[174] The court noted that there was no dispute about the first proposition.[175] However, the court observed there was no evidence of the second proposition: "[T]he agreed-upon fact that a loss of traction can occur with part-time four-wheel drive was simply not linked to the conclusion that the Pathfinder's four-wheel drive system was therefore unsafe at highway speeds."[176] The court invoked the *ipse dixit* language from *Joiner* to exclude the expert's opinion.[177] The fact that four-wheel drives may cause slippage was

---

epidemiological approach of comparing incidence rates in the context of food poisoning).

169. *Blackwell*, 971 A.2d at 253–56.

170. *Id.* at 255.

171. *See supra* text accompanying notes 103–08.

172. 462 F.3d 920 (8th Cir. 2006).

173. *Id.* at 922.

174. *Id.*

175. *Id.* at 923.

176. *Id.* at 924.

177. *Id.* ("Where 'opinion evidence . . . is connected to existing data only by the *ipse dixit* of the expert,' a district court 'may conclude that there is simply too great an analytical gap between the data and the opinion proffered.'" (quoting Gen. Elec. Co. v. Joiner, 522 U.S. 136, 146 (1997))).

not sufficient to support a conclusion that this is what occurred in the specific instance of *Smith*.[178]

To this point, our examples have focused on opinions dealing with individual cases. Another source of evidence revealing the role of sufficiency in expert admissibility determinations can be found in multidistrict litigation ("MDL") rulings. MDL judges often seek to resolve admissibility issues prior to remanding cases to the transferor court where they will be tried. These are mass tort cases in which the parties frequently identify multiple experts to testify on the same subject. Such designations are driven by the need to have a sufficient supply of experts to testify in what may be a substantial number of individual trials that may occur across a significant period of time and throughout the country. Thus, the MDL judge may be faced with making an admissibility determination for many experts when the inevitable *Daubert* motion is made and heard.

In this regard, the opinion in *In re Phenylpropanolamine (PPA) Products Liability Litigation*[179] is particularly instructive. The case involved phenylpropanolamine ("PPA"), an appetite suppressant, and claims by those who had taken it that it caused hemorrhagic and ischemic strokes as well as seizures.[180] Although there was some preliminary investigation of the toxicity of PPA, a comprehensive study of the relationship between PPA and hemorrhagic strokes had been conducted by researchers at Yale University.

The plaintiffs offered the testimony of fourteen experts in pharmacology, epidemiology, neurology, toxicology, and pediatrics to testify to general causation, and the defendants challenged the admissibility of all fourteen.[181] Defense challenges related to the types of injury suffered by the plaintiffs (for example, hemorrhagic strokes, ischemic strokes, and cardiac injuries), the attributes of the plaintiffs (like gender and age), and the time between ingestion of PPA and the relevant adverse event.[182] Judge Barbara Rothstein, the MDL judge, held a *Daubert* hearing in order to rule on the defendants' motions.[183]

The court's opinion systematically addressed each of these arguments. It concluded that the plaintiffs' experts could not testify that events occurring more than three days after ingesting PPA

178. *Id.* at 924–25. Thus, *Smith* is analogous to a toxic tort case in which general causation has been established but no evidence exists to support specific causation.

179. 289 F. Supp. 2d 1230 (W.D. Wash. 2003).

180. *See id.* at 1234–36.

181. *Id.* at 1236.

182. *Id.*

183. *Id.* at 1236–38.

were causally related to the drug.[184] After a rather thorough review of the strengths and weaknesses of the epidemiologic and nonepidemiologic research, the court concluded that the plaintiffs' experts could testify as to a causal relationship between PPA and hemorrhagic stroke.[185] Although most of the existing research involved women between the ages of eighteen and forty-nine, the court permitted the plaintiffs' experts to extrapolate to older and younger individuals and to men by explaining why these were reasonable inferences to draw.[186] The court found the scientific evidence with respect to ischemic strokes to be a much closer call. Ischemic stroke had not been investigated in the Yale study, and one senses that it might have gone either way, but in the end the court concluded that experts could testify about a purported PPA-ischemic stroke association.[187] On the other hand, the court ruled that the evidence linking PPA and other cardiac injuries was too attenuated to permit the plaintiffs' expert to testify as to this relationship.[188]

What is most notable about Judge Rothstein's opinion is that she did not address the admissibility of the testimony of the individual experts proffered by the plaintiffs. Instead, she canvassed the scientific evidence to determine which causal relationships were sufficiently supported by the scientific evidence and which were not. Rather than addressing the methodology and reliability of any experts,[189] she analyzed the reliability of the Yale study and its probity in supporting the causal inferences necessary to the plaintiffs' claims. Her conclusion capsulized her approach: "The court finds expert testimony as to an association between PPA and hemorrhagic or ischemic stroke, in either gender and any age group, admissible. The court finds expert testimony associated with seizures, psychoses, injuries occurring more than three days after ingestion of a PPA-containing product, and cardiac injuries inadmissible."[190]

---

184. *Id.* at 1238.

185. *Id.* at 1238–46.

186. *Id.* at 1244–46.

187. *Id.* at 1248–49.

188. *Id.* at 1249–51 ("Dr. Goldenberg's scattershot expert testimony lacks both the cumulative evidentiary support and the thoroughness the court found reliable with respect to both hemorrhagic and ischemic stroke. Simply put, the evidence proffered by Dr. Goldenberg fails to reliably support his ultimate opinion." (citing Gen. Elec. Co. v. Joiner, 522 U.S. 136, 146 (1997))).

189. For two of the adverse outcomes—ischemic stroke and cardiac injuries—there was no epidemiology and an individual expert was proposed to testify about each of these causal matters. *Id.* at 1246–51. What is notable is that, once again, Judge Rothstein's *Daubert* analysis was based on the scientific evidence—case reports, structural similarity with other known toxic agents, biological plausibility, and treatise assessments—rather than on any expert methodology. *Id.*

190. *Id.* at 1251.

Nor is the *PPA* opinion unique in this regard. In *In re Zoloft (Sertraline Hydrocloride) Products Liability Litigation*,[191] the court excluded the testimony of two experts, an embryologist and a molecular development biologist, that the antidepressant taken during pregnancy causes birth defects.[192] The experts based their opinions on nonepidemiologic evidence, and they did not address the lack of support for their conclusion in the epidemiologic research that existed.[193]

> Therefore, the Court holds that any litigation experts on human causation in this MDL must address the epidemiological research. Where that body of research does not support the conclusions drawn by the experts, the experts must endeavor to reconcile the inconsistent epidemiological data with their opinions. Here, the experts have given scant attention to the epidemiology research in their reports, and have failed to reconcile inconsistent epidemiological evidence with their opinions on human causation.[194]

To us, this passage is similar to the position taken in *O'Connor v. Pennsylvania Railroad* and in a number of the Bendectin cases. The experts' nonepidemiologic research is some evidence of a relationship between Zoloft and birth defects, but it is overcome by the contrary epidemiologic evidence.[195]

Thus far, we have discussed three types of cases: (1) mass tort cases, such as those in the Bendectin and silicone gel breast implant litigation, where nearly every court agreed that the evidence was insufficient because of an overwhelming body of contrary evidence;[196] (2) individual cases where various shortcomings caused

---

191. 26 F. Supp. 3d 466 (E.D. Pa. 2014).

192. *Id.* at 470–73, 481.

193. *Id.* at 475–77. In an earlier opinion, the court excluded the testimony of an expert epidemiologist. *In re* Zoloft (Sertraline Hydrochloride) Prods. Liab. Litig., 26 F. Supp. 3d 449, 455, 465 (E.D. Pa. 2014).

194. *Zoloft*, 26 F. Supp. 3d at 476–77 (footnote omitted).

195. *Id.* at 477. Still, a third example with a slightly different twist is presented in *In re Viagra Products Liability Litigation*, 572 F. Supp. 2d 1071 (D. Minn. 2008). There, the court refused to exclude the testimony of an epidemiologist based on his review of the admittedly mixed results of epidemiologic studies—that Viagra may cause a vision-loss disorder called nonarteritic anterior ischemic optic neuropathy. *Id.* at 1079–82. It left the ultimate admissibility question to transferor courts. *Id.* at 1082. However, the court excluded the testimony of several other experts who based their conclusions on case studies or clinical observations. *Id.* at 1082–86.

196. *E.g.*, Norris v. Baxter Healthcare Corp., 397 F.3d 878, 884–87 (10th Cir. 2005); *In re* Breast Implant Litig., 11 F. Supp. 2d 1217, 1227–44 (D. Colo. 1998); Richardson v. Richardson-Merrell, Inc., 649 F. Supp. 799, 800–03 (D.D.C. 1986), *aff'd*, 857 F.2d 823 (D.C. Cir. 1988). One of the first breast implant cases, *Hopkins v. Dow Corning Corp.*, 33 F.3d 1116 (9th Cir. 1994), resulted in a plaintiff recovery. *Id.* at 1119–20. This occurred at a time prior to the accumulation of the large body of epidemiologic evidence discussed above.

the courts to question the sufficiency of the plaintiffs' evidence; and (3) MDL opinions assessing whether the existing scientific evidence is sufficient to support various expert opinions about causation. As some of the cases we review suggest, inevitably there are cases on the cusp. One group of such cases involves the drug Parlodel, which was prescribed for postpartum women to suppress lactation.[197] Ultimately, it was removed from the market by the FDA.[198]

The frequency of strokes in postpartum women is greater than it is among nonpostpartum women similarly situated.[199] Nevertheless, even among this group, the incidence is

---

197. *Compare* Brasher v. Sandoz Pharm. Corp., 160 F. Supp. 2d 1291 (N.D. Ala. 2001) (admitting expert witness testimony under *Daubert*), Eve v. Sandoz Pharm. Corp., IP 98-1429-C-Y/S, 2001 U.S. Dist. LEXIS 4531 (S.D. Ind. Mar. 7, 2001) (admitting testimony), Globetti v. Sandoz Pharm., Corp., 111 F. Supp. 2d 1174 (N.D. Ala. 2000) (admitting testimony), *and* Hyman & Armstrong, P.S.C. v. Gunderson, 279 S.W.3d 93 (Ky. 2008) (admitting testimony), *with* Dunn v. Sandoz. Pharm. Corp., 275 F. Supp. 2d 672 (M.D.N.C. 2003) (excluding testimony), Soldo v. Sandoz Pharm. Corp., 244 F. Supp. 2d 434 (W.D. Pa. 2003) (excluding testimony), Caraker v. Sandoz Pharm. Corp., 188 F. Supp. 2d 1026 (S.D. Ill. 2001) (excluding testimony), Glastetter v. Novartis Pharm. Corp., 107 F. Supp. 2d 1015 (E.D. Mo. 2000) (excluding testimony), *aff'd*, 252 F.3d 986 (8th Cir. 2001), *and* Hollander v. Sandoz Pharm. Corp., 95 F. Supp. 2d 1230 (W.D. Okla. 2000) (excluding testimony), *aff'd in part and remanded*, 289 F.3d 1193 (10th Cir. 2002).

The trial judge in *Soldo* appointed an expert panel to assist him in his Parlodel admissibility ruling. *Soldo*, 244 F. Supp. 2d at 441–42. Each of the three experts filed a report to the court that assessed the proffered testimony of two plaintiff experts. *Id.* at 442. Two of the court-appointed experts concluded that the testimony of neither of the plaintiff's experts was based on reliable methodology. *Id.* at 503. The third expert concluded that one of the plaintiff's experts used an unreliable methodology, but that the other expert's differential diagnosis was scientifically acceptable. *Id.* at 503–04. The court excluded the plaintiff's experts, basing its decision in part on the fact that none of the available epidemiologic studies demonstrated a statistically significant relationship between the drug and intracerebral hemorrhage. *Id.* at 533–37.

198. *See Glastetter*, 107 F. Supp. 2d at 1035. As some courts have noted, the regulatory threshold is lower than that required in tort claims. With respect to the decision of the FDA to withdraw approval of Parlodel, the court in *Glastetter* commented on the FDA's withdrawal statement:

> [The withdrawal statement] does not establish that the FDA had concluded that bromocriptine can cause an [intracereberal hemorrhage]; instead, it indicates that in light of the limited social utility of bromocriptine in treating lactation and the reports of possible adverse effects, the drug should no longer be used for that purpose. For these reasons, the court does not believe that the FDA statement alone establishes the reliability of plaintiffs' experts' causation testimony.

Id. at 1036.

199. In *Soldo*, the court concluded that the plaintiff's experts did not rule out other possible causes of the plaintiff's stroke. 244 F. Supp. 2d at 518–24. According to the court, one study found that when compared to nonpostpartum women, postpartum women had a relative risk of stroke of 28.3,

exceedingly rare. Unlike strong toxicants like thalidomide or diethylstilbestrol, if Parlodel does cause strokes, it does so at a modestly increased rate, making epidemiologic investigation into this question extremely difficult.[200] The evidence that did exist included a study conducted by Professor Kenneth Rothman, a prominent epidemiologist.[201] Rothman employed a case-control methodology—one that has the best potential to find an effect when the incidence of the investigated disease is very small.[202] However, because strokes in the period after pregnancy are so rare (on the order of 5 per 100,000 pregnancies), Rothman's study could only identify ten cases.[203] Rothman and his coauthors concluded—although they found a modest association—that such a small case series would not permit a statistically meaningful inference of causation to be drawn.[204] Three other epidemiologic studies were inconclusive because of methodological limitations.[205] The plaintiffs'

---

and the expert had not ruled out the postpartum condition itself as a cause of the stroke. *Id.* at 552.

200. *Id.* at 506–16. This problem exists to some extent in the *Zoloft*, *PPA*, and *Viagra* cases noted earlier. However, at least with respect to *PPA* and *Viagra*, the number of exposed individuals and the frequency of the alleged adverse events make epidemiologic investigation less problematic.

201. *See* Siharath v. Sandoz Pharm. Corp., 131 F. Supp. 2d 1347, 1356 (N.D. Ga. 2001).

202. *Id.*

203. *Id.*

204. *Id.* at 1356–57. The *Siharath* case is interesting in the way in which the court characterized the issues, although the procedural context is a *Daubert* motion. The court began, "[T]he Court must address the recurring issue of what is the quantity and quality of scientific evidence that a plaintiff must present on the issue of medical causation in a world of imperfect scientific knowledge." *Id.* at 1349. The court continued, "To survive Defendant's Motions for Summary Judgment, Plaintiffs must produce evidence that would allow a reasonable jury to find to a reasonable degree of medical certainty that Parlodel® is (1) capable of causing stroke and (2) that Parlodel® did in fact cause their strokes." *Id.* at 1352.

For another case conforming to the sufficiency approach, see Caraker v. Sandoz Pharm. Corp., 188 F. Supp. 2d 1026, 1036–37 (S.D. Ill. 2001) (finding that opinions based on animal studies that did not examine whether Parlodel causes intracerebral hematomas do not meet the "fit" requirement).

205. *Siharath*, 131 F. Supp. 2d at 1357–58. Susan Haack explains how one court assessed the evidence about Parlodel as a cause of postpartum stroke:

> Then, to conform to the language of *Daubert*, the court puts all [four of the epidemiologic studies of Parlodel] in terms of the plaintiff's failure to falsify the null hypothesis—here, the hypothesis that any difference, in the sample studied, between the rate of postpartum stroke among women who take Parlodel and those who do not is the result of chance—and reasons that:
>
> > To "falsify" a hypothesis in this context means to prove that the "null hypothesis"—that Parlodel® has no effect on the risk of postpartum stroke—is false, i.e., that Parlodel® in fact significantly increases the risk of postpartum stroke. The failure of plaintiff's experts to show any study proving that the null

experts found adverse drug event reports that were argued to support causation.[206] They also relied on biological mechanism evidence derived from chemical structure similarity—the active ingredient in Parlodel is within a class of drugs that are capable of blood vessel constriction in some patients, and vasoconstriction is at least one of the causes of strokes that result from insufficient blood flow.[207] In vivo animal studies were also proffered, although their primary value was confirming aspects of the biological mechanism evidence—Parlodel can lead to vasoconstriction, which can lead to strokes.[208]

There has been much hand-wringing over the fact that courts came to different conclusions about the admissibility of plaintiffs' experts. Professor Carl Cranor reports, "[T]he courts disagreed on whether to admit essentially the same kind of evidence and same kinds of experts. Moreover, both district and appellate courts are disagreeing about whether to admit expert testimony and supporting scientific evidence that Parlodel has the potential for strokes and heart attacks."[209] Some have called for reconsideration of *Joiner's* decision to adopt the same abuse of discretion standard for other evidentiary rulings as the standard of review of *Daubert* determinations.[210]

While we do not disagree that these inconsistent determinations are unfortunate, we are more sanguine about them than the handwringers. The explanation for our lack of discomfort is that this same phenomenon has long occurred with regard to evaluations of the sufficiency of the evidence. Circumstantial evidence that is neither hopelessly weak nor quite powerful puts courts in the difficult position of policing the line between reasonable inference—a matter for the jury—and impermissible speculation—the basis for a ruling as a matter of law against the party who has the burden of proof on that issue. The *Restatement (Third) of Torts* explains:

> hypothesis has been falsified demonstrates that their causal hypothesis has not been tested or verified by the means of science.
>
> But as the phrase "tested or verified" suggests, *what this really says is that the plaintiff's experts have produced no statistically significant evidence supporting the claim that Parlodel increases the risk of postpartum stroke.*

Haack, *supra* note 75, at 424 (emphasis added) (quoting Soldo v. Sandoz Pharm. Corp., 244 F. Supp. 2d 434, 457 (W.D. Pa. 2003).

206. *Siharath*, 131 F. Supp. 2d at 1359–63.

207. *Id.* at 1363–65.

208. *Id.* at 1366–69.

209. CRANOR, *supra* note 164, at 25 (2006).

210. *See* David L. Faigman, *Appellate Review of Scientific Evidence Under* Daubert *and* Joiner, 48 HASTINGS L.J. 969, 973–79 (1997); Randolph N. Jonakait, *The Standard of Appellate Review for Scientific Evidence: Beyond* Joiner *and* Scheffer, 32 U.C. DAVIS L. REV. 289, 291–92, 335–40 (1999).

> The difficulty that courts confront is that the line between reasonable inference and prohibited speculation is one of the more indistinct lines that exists in law and also is one on which reasonable minds can and do differ. Different courts draw those lines at different points at different times; comparison of cases is very difficult because modest differences in the evidence can substantially affect the power of an inference. Thus, it is not possible to state specific rules that locate the line between permissible inference and prohibited speculation.[211]

A pair of cases in New York illustrates the principle in the *Restatement (Third).* In *Wolf v. Kaufman*,[212] the plaintiff's decedent fell down a flight of unlit stairs, and the defendant was allegedly negligent for not repairing the lights.[213] With regard to the question of whether the lack of lights was the factual cause of the fall, the court held, per Justice Finch, that the evidence was insufficient because "it would be solely a conjecture for a jury to draw the conclusion that the deceased fell down the stairs because of the absence of light."[214] Nine years later, in *Ingersoll v. Liberty Bank*,[215] the issue in the case was whether the plaintiff's decedent fell down a flight of stairs due to a heart attack or dizziness while carrying a thirty-two-pound box on the one hand, or whether, on the other hand, the decedent fell on the second-to-last step and the injuries in the fall caused his death.[216] The step was found broken after the accident and had been negligently maintained by defendant.[217] The same Justice Finch—here, Judge Finch on the Court of Appeals—wrote that the question of causation was a matter for the jury to decide.[218]

Whether the two cases are sufficiently different factually to justify a different outcome is not important.[219] Rather, we think that *Parlodel* is a modern incarnation of rulings that courts have long made about when the evidence introduced in support of an issue is sufficient for a jury to rule on that issue. Appreciating that

---

211. RESTATEMENT (THIRD) OF TORTS: LIABILITY FOR PHYSICAL AND EMOTIONAL HARM § 28 cmt. b (AM. LAW INST. 2010).

212. 237 N.Y.S. 550 (N.Y. App. Div. 1929).

213. *Id.* at 551.

214. *Id.*

215. 14 N.E.2d 828 (N.Y. 1938).

216. *Id.* at 828–29.

217. *Id.*

218. *Id.* at 830; *see also* Reynolds v. Texas & Pac. Ry. Co., 37 La. Ann. 694 (La. 1885) (holding, on facts similar to *Wolf*, that the case was properly one for the fact finder).

219. In fact, there were differences among the Parlodel cases that might have distinguished one case from another—the diseases in individual cases involved two different types of strokes, heart attack, and seizure, and the risk factors for the disease in question varied among the victims. *See supra* notes 197–210 and accompanying text.

connection, then, we should not be surprised that Parlodel courts disagree about where the line is that the *Restatement (Third)* identifies in the quotation above.

CONCLUSION

If we are correct that admissibility decisions are best thought of as sufficiency of the evidence rulings, then several implications and issues follow. Because a full resolution of each of the issues would constitute an article in itself, we choose here only to present them for identification and future discussion.

First, it is our sense that courts are in fact moving in the direction of treating admissibility rulings as at least very close cousins of sufficiency rulings. The "analytical gap" language in *Joiner* and the focus on the case at hand in *Kumho Tire* both push courts away from a *Daubert*-factor analysis and toward a decision as to whether there is sufficient evidence to support a plaintiff verdict. The current language of Federal Rule 702 reflects this trend through its express mention of the sufficiency of the data as an important component of admissibility rulings.[220] Our sense of things, however, is simply that—a sense. It would be quite helpful to examine a set of cases more systematically in order to see if we are correct that sufficiency analyses are becoming more common and to assess when courts most frequently apply sufficiency reasoning in making *Daubert* rulings.

A second aspect of a movement toward sufficiency is its impact on the relative importance of the various *Daubert* factors. Earlier, we quoted Professor Michael Saks' observation that while the *Frye* test asks the judge to focus on the opinions of other experts, *Daubert* moves the inquiry in the direction of asking the judge to assess the evidence herself.[221] Nevertheless, the *Daubert* factors save an important role for the opinions of others, both by retaining the general acceptance test and by telling judges that they should look to the peer review and publication record of an idea or theory. Insofar as courts move toward a sufficiency analysis, one should anticipate a reduced role for these two factors. In order to decide whether the evidence of causation in a tort context is sufficient to send to a jury, the judge must ultimately decide this question on the merits, not the opinion of others. This fact is consistent with our observation that in most admissibility decisions, these two factors serve merely as makeweights. Very, very few admissibility decisions turn on general acceptance or peer review and publication alone. These two factors may play an even lesser role in the future. A sufficiency approach tells judges, "Look for yourself."

---

220. *See supra* text accompanying note 111.

221. *See supra* note 73.

Third, and perhaps controversially, we believe a recognition that admissibility rulings are often sufficiency judgments presents an opportunity to revisit the *Joiner* court's determination that these rulings should be reviewed using an abuse of discretion standard. In a typical *Daubert* case where a trial court excludes the testimony of the plaintiff's only causation witnesses, the following occurs: The court grants a motion to exclude the evidence and then immediately grants a summary judgment to the defendant. When the case reaches an appellate court governed by *Joiner*, that court reviews the admissibility decision under an abuse of discretion standard and then reviews the summary judgment de novo. Of course, if the appellate court affirms the exclusion of the expert testimony, then the de novo review of the summary judgment is pro forma.

From the time the abuse of discretion standard was adopted, some criticized it because of the potential for inconsistent results.[222] They argue that if one of the justifications for a heightened admissibility standard is to avoid the unseemliness of inconsistent jury verdicts, as occurred in the Bendectin litigation, the abuse of discretion standard threatens the same sort of inconsistency, only now at the admissibility stage.[223] At least theoretically, appellate courts could affirm both the admission and the exclusion of testimony in nearly identical cases. Some would argue that this occurred in the Parlodel litigation.

However, most jurisdictions do not use an abuse of discretion standard to review sufficiency decisions, which is why granting or denying summary judgment or judgment as a matter of law is reviewed de novo.[224] Viewing admissibility decisions themselves as sufficiency determinations raises the question of whether appellate courts should review these decisions under a heightened standard.

Of course, that is exactly what the Eleventh Circuit did in *Joiner*, and this position was expressly rejected by the Supreme Court.[225] It is unlikely that the Court will reverse course. Nor is it entirely clear that it should do so. Ultimately, the pair of decisions described above present a complex question of mixed law and fact, and finding the proper standard for judicial review in such

---

222. *See, e.g.*, Margaret A. Berger, *What Has a Decade of* Daubert *Wrought?*, 95 AM. J. PUB. HEALTH S59, S62–S63 (Supp. 1, 2005); Cassandra H. Welch, Note, *Flexible Standards, Differential Review:* Daubert's *Legacy of Confusion*, 29 HARV. J.L. & PUB. POL'Y 1085, 1091–94 (2006).

223. *See* Berger, *supra* note 222, at S62 ("The different approaches courts take on the admissibility of expert proof in toxic tort cases is exemplified by the inconsistent results in the line of cases involving Parlodel."); Welch, *supra* note 222, at 1094 ("Not only does the flexible standard raise junk science concerns, but it could also lead to inconsistent decisions.").

224. *See* Szekeres v. CSX Transp., Inc., 731 F.3d 592, 597 (6th Cir. 2013); Mays v. United States, 763 F.2d 1295, 1296–97 (11th Cir. 1985).

225. Gen. Elec. Co. v. Joiner, 522 U.S. 136, 143 (1997).

situations is never easy. Nevertheless, our analysis suggests that the topic is worthy of further discussion and that something like the "hard look" standard of appellate review proposed by Judge Becker has much to offer as a potential admissibility standard.[226]

Fourth, and finally, our analysis raises the following question: if this is all a matter of sufficiency, should courts simply wait until the case is tried and then make a determination as to whether the evidence is sufficient to support a verdict for the plaintiff? We do not think so. As the introduction to this Article suggested, in an adversarial system employing lay fact finders, there are multiple reasons for imposing a reliability filter on expert evidence. These reasons are not affected by our argument that many admissibility rulings are in fact sufficiency rulings.

Equally importantly for this analysis, the nearly twenty years of experience with *Daubert* hearings demonstrates that courts can, at the outset of a case, both have the plaintiff's expert put up her evidence and afford the defendant's experts the opportunity to rebut—explain why that evidence is insufficient—and proffer their own contradictory evidence. That, after all, is precisely what happens with summary judgment. When such procedures are followed, the plaintiff does get a fair chance to demonstrate sufficiency. Typically, there is not a large body of additional evidence that could come out at trial that would reverse such a decision.[227] Thus, existing procedures have the additional virtue that, in many cases, they have the potential to save substantial administrative costs.

In sum, our goal is not to abolish *Daubert* hearings. It is much more modest. It is simply to point out that Rule 702 admissibility rulings are in many situations indistinguishable from sufficiency rulings. *Daubert* hearings would be more transparent, coherent, and clearly reasoned if courts were to consciously recognize this fact.

---

226. *In re* Paoli R.R. Yard PCB Litig., 35 F.3d 717, 749–50 (3d Cir. 1994).

227. The greater problem is that after a case is resolved, new scientific evidence may appear that makes the plaintiff's case stronger or weaker. However, this potential problem exists no matter when the sufficiency determination occurs.