

University at Buffalo School of Law

Digital Commons @ University at Buffalo School of Law

Journal Articles

Faculty Scholarship

Winter 1999

Guarding the Gate to the Courthouse: How Trial Judges Are Using Their Evidentiary Screening Role to Remake Tort Causation Rules

Lucinda M. Finley

University at Buffalo School of Law

Follow this and additional works at: https://digitalcommons.law.buffalo.edu/journal_articles



Part of the [Evidence Commons](#), and the [Torts Commons](#)

Recommended Citation

Lucinda M. Finley, *Guarding the Gate to the Courthouse: How Trial Judges Are Using Their Evidentiary Screening Role to Remake Tort Causation Rules*, 49 DePaul L. Rev. 335 (1999).

Available at: https://digitalcommons.law.buffalo.edu/journal_articles/196



This Article is brought to you for free and open access by the Faculty Scholarship at Digital Commons @ University at Buffalo School of Law. It has been accepted for inclusion in Journal Articles by an authorized administrator of Digital Commons @ University at Buffalo School of Law. For more information, please contact lawscholar@buffalo.edu.

GUARDING THE GATE TO THE COURTHOUSE: HOW TRIAL JUDGES ARE USING THEIR EVIDENTIARY SCREENING ROLE TO REMAKE TORT CAUSATION RULES

*Lucinda M. Finley**

I. INTRODUCTION

Vigorously exercising their role as evidentiary “gatekeepers”—a task assigned to them by the United States Supreme Court in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*¹—federal trial judges in products liability cases have been doing far more than screening proposed expert testimony to determine admissibility. The *Daubert* gatekeeper power has become a potent tool of tort lawmaking. Under the guise of admissibility determinations, federal judges have been making significant substantive legal rules on causation by substantially raising the threshold of scientific proof plaintiffs need to get their expert causation testimony admitted, and thus survive summary judgment. While the decisions purport to be no more than deferential nods to the criteria of science, judges have actually been making legal rules about what types and strengths of scientific evidence are necessary in order to prove causation. The emerging legal rule is that plaintiffs’ experts must be able to base their opinions about causation on epidemiological studies, and that these studies standing alone must show that the population-wide risk of developing the disease in question, if exposed to defendants’ products, is at least double the risk without exposure.

In the process of developing this legal rule, judges in products liability cases have been making profoundly normative judgments about the social allocation of risk and who should bear the burden of scien-

* Professor of Law, State University of New York at Buffalo. I would like to thank the other presenters and participants in the Clifford Symposium for their encouragement and very helpful comments on this paper, particularly Michael Green, Richard Wright, Marshall Shapo, Gary Schwartz, and Terrence Kiely. I also owe a debt of gratitude to Dr. John DeFazio for taking time out of his busy medical practice to read a draft and give me the insights of a clinician who must often make causal diagnostic judgments based on less than absolute scientific certainty, and to Neil Buchanan for many provocative conversations about the illusory scientific nature of statistical significance.

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

tific uncertainty or controversy—injured people or manufacturers of the products alleged to have caused those injuries? Few of the opinions announcing or applying this emerging causal proof standard acknowledge awareness of the normative nature or implications of their decisions. Thus, an important underlying policy debate remains submerged. Should the rules of the tort system put the onus for uncertainty about the risks of a product on the manufacturer who has marketed it, perhaps without sufficient testing or warning? By doing so, causation rules would enhance the tort policies of deterring marketing of relatively untested products and promoting expanded research on both the effectiveness and hazards of drugs and medical devices. The tort system would also align itself more with the public health protective values that govern the FDA regulatory arena, where a drug is presumed not safe for marketing unless the manufacturer can prove its safety.² Or, should the tort system instead embrace the conservative values of the scientific discipline of epidemiology, whose internal disciplinary standards start with a hypothesis of lack of risk, and demand stringent statistical proof of a doubling or tripling of the risk of a disease before entertaining the possibility of a causal association?³ Judges have been using their evidentiary gatekeeper power to squarely align tort law with the conservative causal normative principles of epidemiology, thus moving the law sharply away from the more consumer protective social policies about risk embodied in the safety regulatory system.

While adopting substantive changes in causation law through the rubric of evidentiary admissibility decisions, judges have also frequently conflated admissibility decisions and sufficiency of evidence decisions. This has effected another profound but concealed change in tort law. Judges have applied *Daubert* to subject each item of expert proof proffered by plaintiffs to substantive causation law scrutiny, to see if it, standing alone, would prove both general and specific cau-

2. See, e.g., David A. Kessler, *The Basis of the FDA's Decision on Breast Implants*, 326 *NEW ENG. J. MED.* 1713 (1992). In explaining the FDA's decision to limit the use of silicone gel breast implants to controlled clinical trials, former FDA Commissioner David Kessler explained that "[c]aveat emptor has never been . . . the philosophy at the FDA." *Id.* at 1715. He noted that the law governing FDA approval of drugs and medical devices "requires a positive demonstration of safety - and the burden of proof rests squarely with the manufacturer." *Id.* at 1713.

3. For an explanation of basic epidemiological precepts, including the degrees of increased risk ratio that epidemiologists prefer to see before they will suggest the possibility of a causal association, see Melissa Moore Thompson, *Causal Inference in Epidemiology: Implications for Toxic Tort Litigation*, 71 *N.C. L. REV.* 247, 249-54 (1992). For a discussion of the values that underlie epidemiology and their differences from the values and goals of the legal system, see Margaret G. Farrell, *Daubert v. Merrell Dow Pharmaceuticals, Inc.: Epistemology and Legal Process*, 15 *CARDOZO L. REV.* 2183 (1994).

sation. If the scientific studies underlying an expert's opinion are not alone sufficient, then the expert's testimony is deemed inadmissible.⁴ This stands in stark contrast to traditional and proper practice, which sees the admissibility of evidence as a question quite distinct from the sufficiency of evidence to meet a plaintiff's burden of proof. The sufficiency inquiry is supposed to view plaintiffs' evidence in its entirety to see if, taken as a whole, it would support a conclusion that causation is more likely than not.⁵ By calling what is really a sufficiency of the evidence determination an admissibility decision, judges are using their evidentiary gatekeeper power to close the gate on plaintiffs' opportunities to have their proof evaluated as a cumulative whole. This subtly but substantially increases plaintiffs' burden of proving individual causation, and it also furthers the trend in toxic tort cases to shift the allocation of power away from juries to judges.⁶ Because a trial judge's decision to exclude evidence is reviewed under the lenient abuse of discretion standard of review,⁷ this new heightened substantive standard of causation and judges' applications of it are largely insulated from meaningful appellate review.

This article explores these substantive causation law implications, and the consequent social policy import, of the way federal trial judges in products liability cases have exercised their *Daubert* evidentiary gatekeeper role. First, the article will briefly describe the recent important Supreme Court trilogy of products liability evidence decisions: *Daubert v. Merrell Dow Pharmaceuticals, Inc.*,⁸ *General Electric Co. v. Joiner*,⁹ and *Kumho Tire Co. v. Carmichael*.¹⁰ It will then examine how trial judges have used these decisions, despite *Daubert's* announced intention to liberalize decisions on admitting scientific expert causation testimony, to erect conservative, and in some instances, virtually insurmountable barriers to plaintiffs' ability to prove causation. These barriers stem from substantive legal decisions about causation law, rather than from scientific principles or case by case assessments

4. See *infra* text accompanying notes 25-33 (discussing cases that take this step of turning admissibility decisions into sufficiency of the evidence rulings).

5. The blurring of the line between the admissibility of evidence and the sufficiency of that evidence to prove causation is explored in the context of cases dealing with whether the morning sickness drug Bendectin caused birth defects. See JOSEPH SANDERS, *BENDECTIN ON TRIAL: A STUDY OF MASS TORT LITIGATION* 155-56 (1998); Susan R. Poulter, *Science and Toxic Torts: Is There a Rational Solution to the Problem of Causation?*, 7 *HIGH TECH. L.J.* 189, 192-94 (1992).

6. See SANDERS, *supra* note 5, at 185-93 (discussing the trend in Bendectin cases to shift decisional power away from juries to judges).

7. See *General Electric Co. v. Joiner*, 522 U.S. 136, 141 (1997).

8. 509 U.S. 579 (1993).

9. 522 U.S. 136 (1997).

10. 526 U.S. 137 (1999).

of proffered testimony. Finally, the article analyzes some of the societal implications of these legal developments, including differential impact on social groups whose health problems have tended to be ignored or underexplored by the scientific research community; a diminished ability of the tort system to stimulate scientific research and knowledge; and a premature sense of scientific closure from the legal declaration of the supposedly definitive state of science, which can shut down further research on scientific issues that are still uncertain or contested.

II. THE PRODUCTS LIABILITY EVIDENCE TRILOGY

Daubert emerged from Bendectin litigation, which was a significant source of the controversy over whether courts were admitting "junk science," the epithet that critics of the tort system placed on expert opinions they regarded as not well-founded in scientific methodology or departures from the mainstream of scientific opinion.¹¹ When Bendectin was initially marketed to prevent or reduce morning sickness in pregnant women, the pharmaceutical company had done little research into potential adverse effects, including the potential to produce birth defects.¹² Several women who took Bendectin and then bore children with deformed limbs filed suit on behalf of their damaged children, claiming that the drug was the cause of the birth defects. When some of these initial Bendectin cases went to trial, juries were disturbed by the paucity of safety research conducted by the manufacturer, and by the company's efforts to ignore or misrepresent warning signs in animal studies. The ensuing plaintiffs' verdicts stimulated an exponential growth in Bendectin litigation.¹³

The burgeoning litigation and accompanying media attention helped to push the issue of Bendectin's teratogenicity onto the scientific community's research agenda. Numerous epidemiological studies, some with the financial support of the drug company, were conducted, and the published results indicated that taking Bendectin

11. The term was coined by Peter Huber in PETER HUBER, *GALILEO'S REVENGE: JUNK SCIENCE IN THE COURTROOM* (1991). For a thorough dissection of Huber's work as nonscholarly and based on inaccurate, or largely unsupported or overblown anecdotes, see Kenneth J. Chesebro, *Galileo's Retort: Peter Huber's Junk Scholarship*, 42 AM. U. L. REV. 1637 (1993). See also DAVID L. FAIGMAN, *LEGAL ALCHEMY: THE USE AND MISUSE OF SCIENCE IN THE LAW* 210, n.29 (1999) (describing Huber's book as a myopic attack on plaintiffs' experts).

12. For a history of Bendectin, its testing, marketing, and the subsequent litigation, see MICHAEL GREEN, *BENDECTIN AND BIRTH DEFECTS: THE CHALLENGES OF MASS TOXIC SUBSTANCES LITIGATION* (1996); SANDERS, *supra* note 5.

13. In his book on the Bendectin litigation, Joseph Sanders provides important insights into why the strong evidence of manufacturer misconduct swayed juries to downplay or disregard uncertain, weak or negative evidence on causation. See SANDERS, *supra* note 5, at 12-13, 130-34.

during pregnancy did not appreciably increase the risk of bearing a child with birth defects.¹⁴ Despite this mounting scientific evidence tending to exonerate the drug, plaintiffs produced experts who conducted, for purposes of the litigation, reanalyses of the existing studies. These reanalyses critiqued some of the defense oriented studies for including in the group of women deemed exposed to Bendectin those who did not take the drug until after the crucial time of fetal limb formation. These reanalyses, not surprisingly, did show that Bendectin more than doubled the risk of birth defect development.¹⁵

When presented with plaintiffs' reanalyses of the epidemiology, and with animal study data and experts comparing the chemical structure of Bendectin to that of known teratogens, some juries still returned verdicts in favor of plaintiffs, despite the growing body of contrary epidemiology. These verdicts were regarded as aberrational by many members of the scientific community, the defense bar, some trial judges, and virtually all appellate courts reviewing Bendectin cases. A judicial consensus emerged to parallel the scientific consensus that Bendectin had not been proven to be a teratogen. In response to this scientific consensus, judges employed the technique of judgment notwithstanding the verdict,¹⁶ or ruled that plaintiffs lacked sufficient evidence to satisfy the burden of proof,¹⁷ to gradually adopt a legal rule that unless plaintiffs could produce a consistent body of statistically significant epidemiological studies that showed that Bendectin at least doubled the risk of birth defects, plaintiffs did not have sufficient evidence of causation to support a verdict.

Other courts faced with Bendectin cases adopted similar reasoning about the essential threshold of scientific proof, but instead of supplanting a jury verdict, they deemed inadmissible the opinion of any plaintiff's expert who attempted to draw a causal inference based on anything other than statistically significant, peer reviewed, published epidemiological studies that showed a relative risk above the background risk level of two or greater.¹⁸ In *Daubert*, the Ninth Circuit held that plaintiffs' expert testimony on causation was inadmissible

14. For a thorough discussion of the epidemiological studies on Bendectin, see *id.* at 61-86; GREEN, *supra* note 12, at 26-34; see also Joseph Sanders, *From Science to Evidence: The Testimony on Causation in the Bendectin Cases*, 46 STAN. L. REV. 1 (1993).

15. For a discussion of the methodological, selection bias, and other confounding errors of these reanalyses by plaintiffs' experts, see SANDERS, *supra* note 5, at 181-83.

16. See *Richardson v. Richardson-Merrell, Inc.*, 649 F. Supp. 799 (D.D.C. 1986).

17. See *Brock v. Merrell Dow Pharms., Inc.*, 874 F.2d 307 (5th Cir. 1989).

18. See *Lynch v. Merrell Nat'l Labs.*, 830 F.2d 1190 (1st Cir. 1987); *Ealy v. Richardson-Merrell, Inc.*, 897 F.2d 1159 (D.C. Cir. 1990); *Richardson v. Richardson-Merrell*, 857 F.2d 823 (D.C. Cir. 1988).

because it did not meet this legal and scientific threshold. The court invoked the venerable ruling in *Frye v. United States*,¹⁹ a 1923 case about the admissibility of polygraph evidence. *Frye* required that in order to be admissible, a scientific opinion must be “generally accepted” within the scientific community, and that peer review and publication were inviolate requirements of general acceptance. Opinions that were contrary to the weight of a growing body of scientific studies also could not, by definition, satisfy the “general acceptance” standard.²⁰

Complaining that the *Frye* “general acceptance” rule was outdated, had been superseded by the more liberal Federal Rules of Evidence, and unfairly operated to exclude the scientifically valid, but novel or controversial opinion, the plaintiffs petitioned the Supreme Court to review the Ninth Circuit decision. Limited to this narrow issue of whether *Frye* was consistent with the Federal Rules of Evidence, the Court granted certiorari, and reversed, holding that the ostensibly more liberal standards of Federal Rule 702 governed the admissibility of scientific expert testimony.²¹

The Court first emphasized that under Rule 104(a), the federal trial judge has an obligation to serve as a “gatekeeper” to screen evidence and admit only scientific opinion testimony that comports with Rule 702. This aspect of the ruling rejected the “just check credentials and let it all in” attitude that some judges and commentators had previously adopted towards expert testimony.²² Rule 702 speaks of “scientific, technical, or other specialized knowledge” that will assist the trier of fact. Thus, the Court held, “the subject of an expert’s testimony must be ‘scientific knowledge’”—it must be grounded in the methods and procedures of science.²³ The evidentiary reliability of expert testimony turns on whether it meets the criteria of “scientific knowledge,” which is judged according to its “scientific validity.”²⁴

Justice Blackmun’s opinion then attempted to provide some factors to help trial judges assess scientific validity: 1) scientific knowledge is “falsifiable”—it can be and has been tested to see if the results can be replicated or disproven; 2) scrutiny of the scientific community through peer review or publication is a factor bearing on validity, but it is not the *sine qua non* as it is under *Frye*; 3) the court should con-

19. 293 F. 1013 (D.C. Cir. 1923).

20. *Daubert v. Merrell Dow Pharms., Inc.*, 951 F.2d 1128 (9th Cir. 1991).

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

22. *See, e.g., Ferebee v. Chevron Chem. Co.*, 736 F.2d 1529 (D.C. Cir. 1984) (example of “let it all in” attitude).

23. *Daubert*, 509 U.S. at 589-90.

24. *Id.* at 590 n.9.

sider the known or potential error rates of the scientific technique or method; and 4) general acceptance within the relevant scientific community is a permissible, but not determinative, factor.²⁵ The Court stressed that this assessment of validity under Rule 702 was meant to be flexible, and cautioned that judges should focus on the principles and methodology, not on the conclusions drawn from scientifically valid methodology.²⁶ Finally, using the general Rule 402 admissibility criterion of relevance, the Court admonished that scientific expert opinion testimony must “bear a valid scientific connection to the pertinent inquiry.”²⁷

Because *Daubert* rejected the *Frye* test’s required deference to scientific conventional wisdom, and stressed the liberality and flexibility of the Federal Rules of Evidence, some hailed it as a pro-plaintiff decision that would make it easier for toxic tort and products liability claimants to prove causation.²⁸ Others predicted that because the decision required trial judges closely to scrutinize the bases for scientific opinion, and to defer to the standards of the scientific community, it would lead to more widespread exclusion of plaintiffs’ proposed experts, particularly where plaintiffs were pushing new or still controversial scientific theories of causation.²⁹

Those who predicted that trial judges would flex their gatekeeper muscles to exclude vast quantities of plaintiffs’ proposed expert causation opinion testimony in products liability cases have turned out to be right. The post-*Daubert* era can fairly be described as the period of “strict scrutiny” of science by non-scientifically trained judges. Judges now routinely assess the reliability of scientific studies by scrutinizing the criticisms leveled at the methodology of particular studies. What few predicted, however, is the way *Daubert* has been expanded well

25. *Id.* at 593-94.

26. *Id.* at 594-95.

27. *Id.* at 592.

28. For example, the *Cardozo Law Review* sponsored a symposium on *Daubert*, with academics and practitioners as participants. The opinions at this symposium were closely divided between those who thought that *Daubert* would prove to be a boon to plaintiffs, and those who opined that in application, it would favor defendants. For examples of those who hailed it as a pro-plaintiff ruling, see, for example, Anthony Z. Roisman, *Conflict Resolution in the Courts: The Role of Science*, 15 *CARDOZO L. REV.* 1945 (1994); Kenneth Chesebro, *Taking Daubert’s “Focus” Seriously: The Methodology/Conclusion Distinction*, 15 *CARDOZO L. REV.* 1745 (1994).

29. See, e.g., David E. Bernstein, *The Admissibility of Scientific Evidence After Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 15 *CARDOZO L. REV.* 2139 (1994); Bert Black, *The Supreme Court’s View of Science: Has Daubert Exercised the Certainty Demon?*, 15 *CARDOZO L. REV.* 2129 (1994); Clifton T. Hutchinson & Danny S. Ashby, *Daubert v. Merrell Dow Pharmaceuticals, Inc.: Redefining the Bases for Admissibility of Expert Scientific Testimony*, 15 *CARDOZO L. REV.* 1875 (1994); Marc S. Klein, *After Daubert: Going Forward With Lessons from the Past*, 15 *CARDOZO L. REV.* 2219 (1994).

beyond its Bendectin context to usher in heightened substantive rules of causation. The remaining two cases in the expert admissibility trilogy have facilitated this trend towards making substantive causation law in the guise of evidentiary admissibility determinations.

Shortly after the *Daubert* decision, some judges began to question, or chip away at the Court's caution that admissibility determinations were to be made on the basis of the scientific validity of methodology, and not the conclusions. The United States Court of Appeals for the Third Circuit became the most forceful proponent for the view that methodology and conclusions are not so distinct as the Court seemed to indicate in *Daubert*. In the ongoing *Paoli Railroad Yard Litigation*, the court reasoned that to exercise their role as gatekeepers, judges must not only determine if the underlying methodology on which an expert relies is scientifically valid, but must also evaluate whether the conclusion drawn from the methodology is scientifically supportable.³⁰ This growing tendency of trial judges to throw out expert testimony because the conclusions were deemed questionable—often because controversial or not “generally accepted”—prompted some federal appellate courts to insist on maintaining the distinction between the admissibility of evidence and the weight to be given to it by the trier of fact.³¹

The Eleventh Circuit was one of the courts that took district judges to task for scrutinizing an expert's conclusions while passing on the admissibility of the opinion.³² This prompted the next case to go to the Supreme Court, *General Electric v. Joiner*.³³ In a case alleging that plaintiff developed cancer from occupational exposure to PCBs, the Eleventh Circuit reversed the trial court's decision to exclude plaintiff's expert testimony on causation. The court ruled that because the Federal Rules displayed a liberal preference for admissibility, and the judge had treaded too far into evaluating conclusions and thus usurped the jury's function to weigh the competing expert conclusions, appellate courts should apply a “stringent standard of review” to decisions to exclude evidence.

30. See *In re Paoli R.R. Yard PCB Litig.*, 35 F.3d 717, 745-46 (3d Cir. 1994).

31. The two circuits that have been most concerned with limiting admissibility decisions used by scientific experts, and not allowing the decision whether to admit evidence to blur into an assessment of its weight or the soundness of the conclusions drawn from a body of scientific evidence, are the District of Columbia and Second Circuits. See, e.g., *In re Joint E. & S. Dist. Asbestos Litig.*, 52 F.3d 1124 (2d Cir. 1995); *Ambrosini v. Labarraque*, 101 F.3d 129, 141 (D.C. Cir. 1996).

32. See *Joiner v. General Elec.*, 78 F.3d 524, 529-30 (11th Cir. 1996).

33. 522 U.S. 136 (1997).

Joiner was an electrician who had long experienced occupational exposure to a PCB contaminated coolant fluid. When he contracted small cell lung cancer, he sued the manufacturer of PCBs, and the manufacturer of the transformers and coolant he had worked with, contending that his exposure to PCBs promoted his cancer. To prove causation, Joiner's experts relied on animal studies showing that PCBs injected into mice caused cancer, and on epidemiological studies of groups of workers exposed to PCBs who developed higher than expected rates of cancer. The district court picked apart these studies, despite the fact that all were concededly based on scientifically valid methodologies. The animals were infant mice, not adults; the dose they were given far exceeded the maximum PCB dose that Joiner was ever exposed to; the type of cancer the mice developed was different than the type afflicting Joiner. One epidemiologic study showing a heightened incidence of lung cancer in PCB production workers was not statistically significant; another did not rule out toxins other than PCBs to which the workers had also been exposed. For all these reasons, the district court ruled, any conclusions drawn by Joiner's experts that these studies could support a finding that PCBs could cause lung cancer in humans were not supportable.³⁴

The Supreme Court reversed the Eleventh Circuit's holding that decisions to exclude expert evidence must be reviewed under a stringent standard. Instead, the Court held, like any other evidentiary decision, trial judges' decisions to exclude scientific expert testimony under *Daubert* should be assessed under the lenient "abuse of discretion" standard of review. But the Court did not stop there; it went on to substantially undermine the line between methodology and conclusions it had drawn in *Daubert*. Endorsing the view propounded by the Third Circuit, and rejecting the Eleventh Circuit's determination that it is legal error to reject expert testimony because the trial judge disagrees with the conclusions an expert draws from the data, the Court stated:

[C]onclusions and methodology are not entirely distinct from one another. Trained experts commonly extrapolate from existing data. But nothing in *Daubert* or the Federal Rules of Evidence requires a district court to admit opinion evidence which 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. [citations omitted] That is what the

34. For a discussion of the district court's reasons for rejecting the studies relied on by plaintiff's experts, see *id.* at 144-46.

District Court did here, and we hold that it did not abuse its discretion in so doing.³⁵

This aspect of the *Joiner* majority opinion prompted a partial dissent from Justice Stevens. He argued for continued adherence to *Daubert's* bright line between methodology and conclusions, and cautioned that scrutiny of an expert's conclusions at the admissibility phase usurps the traditional function of the jury to assess the validity or strength of an expert's conclusions.³⁶

Justice Breyer, in a concurring opinion, endorsed the majority view that in exercising their gatekeeper role, judges must "make subtle and sophisticated determinations about scientific methodology and its relation to the conclusions an expert witness seeks to offer. . . ."³⁷ The need for judges to make their own judgments about an expert's conclusions, Justice Breyer opined, is especially important "where the science itself is tentative or uncertain."³⁸ This is an open invitation to trial judges to make normative decisions about how much scientific controversy or uncertainty the tort system should tolerate when determining whether to let plaintiffs in products liability suits have an opportunity to present their case to a jury. Indeed, Justice Breyer made his value judgment explicit:

[M]odern life, including good health as well as economic well-being, depends upon the use of artificial or manufactured substances, such as chemicals. And it may, therefore, prove particularly important to see that judges fulfill their *Daubert* gatekeeping function, so that they help assure that the powerful engine of tort liability, which can generate strong financial incentives to reduce, or to eliminate, production, points towards the right substances and does not destroy the wrong ones.³⁹

The lesson of *Joiner* is that judges may scrutinize a scientists' conclusions, even when doing so requires judges to make "subtle and sophisticated" scientific evaluations and to wade into areas of scientific controversy and uncertainty. Moreover, judges should be particularly careful about admitting conclusions that are still considered tenuous or debatable, even when based on scientifically valid methodology, lest the "engine of tort liability" prematurely condemn a product that may, upon further scientific study, prove not to be harmful. In this regard, *Joiner* risks resuscitating the *Frye* "general acceptance" test

35. *Id.*

36. *Id.* at 154-55 (Stevens, J., dissenting).

37. *Id.* at 147 (Breyer, J., concurring).

38. *Id.* at 147-48.

39. *Joiner*, 522 U.S. at 148-49.

supposedly repudiated in *Daubert*.⁴⁰ And, no matter how the judge rules on a plaintiff's proffered evidence, that decision will be largely insulated from searching appellate scrutiny, because it is difficult to conclude that a trial judge has abused her discretion.

For these reasons, *Joiner* is far more than a mere lesson in the standard of review to be applied to evidentiary rulings. It expresses a normative judgment that judges are to be trusted more than juries (and sometimes more than scientists) in areas where law intersects with science. It also urges caution on the tort system, expressing a preference to wait for the slow attainment of scientific certainty rather than to make a decision that may be ahead of scientific consensus, which often then prompts further scientific inquiry. In Justice Breyer's view, a legal decision that a product is harmful when science is not yet certain presents greater policy problems than the alternative of allowing continued marketing and barring the courthouse door to ill people whose claims of causation may in fact later be widely embraced by the scientific community.

The final step in the Court's expert testimony trilogy occurred in March, 1999, when the Court decided *Kumho Tire Co. v. Carmichael*.⁴¹ Given the emphasis in *Daubert* and *Joiner* on scientific expert testimony, lower courts had disputed whether trial judges should also act as gatekeepers to strictly scrutinize the testimony of non-scientific or technical experts, such as engineers, forensic experts, psychologists, economists, sociologists, or clinicians. *Kumho* was a products liability case stemming from a tire blowout which caused a fatal traffic accident. The plaintiffs sued the tire's manufacturer and distributor, claiming the blowout was caused by a defect in the tire that made its tread separate from the steel-belted tire carcass. Their defect and causation case relied extensively on the opinion testimony of an expert in tire failure analysis. When faced with the defendants' challenge to the admissibility of this expert's opinion, the district court engaged in a rather rote application of the *Daubert* factors—peer review and publication, potential error rate of the methodology, and degree of acceptance within the relevant scientific community—to conclude that the expert's methods and conclusions he drew from his analysis were not "reliable." The Eleventh Circuit reversed the decision to exclude the expert's testimony, holding that the *Daubert* factors were relevant

40. See, e.g., *Diaz v. Johnson Matthey Inc.*, 893 F. Supp. 358, 375 (D.N.J. 1995) (offering the insight of a federal district court in New Jersey in a pre-*Joiner* decision: "If conclusive [scientific] evidence were necessary to admit . . . a theory on general causation, we might as well be proceeding under the Frye general acceptance test . . .").

41. 119 S. Ct. 1167 (1999).

only to “science,” and thus should not be applied to non-scientific, experience-based observation.⁴²

The Supreme Court reversed the Eleventh Circuit’s decision. First, it rejected the distinction lower courts had tried to draw between “scientific” and “nonscientific” experts, because, by its terms, Rule 702 applies to all “technical or specialized knowledge.” Thus, trial judges must perform the gatekeeping task of closely scrutinizing all expert testimony. The Court further concluded that while the specific *Daubert* factors may not all apply outside the scientific realm, *Daubert*’s “general principles” apply. Thus, when an expert’s “factual basis, data, principles, methods, or their application are called sufficiently into question . . . the trial judge must determine whether the testimony has a reliable basis in the knowledge and experience of the relevant discipline.”⁴³ By including an expert’s application of methods and principles as among the matters to be assessed, the Court again signaled its apparent approval of scrutinizing the conclusions an expert draws from data, studies, or visual inspection. Indeed, the Court itself painstakingly evaluated the expert’s conclusions. It did not question the methodology of visually inspecting tires, but stressed that the expert’s conclusion that the particular tire in question had a defect was not reliably supported by his visual inspection, given the presence of known wear factors that could counter the defectiveness conclusion.

The combined implication of this trilogy of Court decisions is that in all areas of products liability and tort law that rest on any sort of specialized, scientific, or technical expert testimony, federal trial courts must exercise a vigorous gatekeeping function to carefully evaluate the basis and conclusions of plaintiffs’ proposed expert testimony, whether going to the issue of defect, clinical diagnosis or standard of care in medical malpractice cases,⁴⁴ or causation in toxic tort cases. In

42. *Carmichael v. Samyang Tire, Inc.*, 131 F.3d 1433 (11th Cir. 1997).

43. *Id.*

44. The Court’s decision in *Kumho Tire* seems to have effectively resolved a brewing controversy in the lower courts over whether clinical medical testimony must be scrutinized under the *Daubert* factors. See, e.g., *Moore v. Ashland Chem. Inc.*, 151 F.3d 269 (5th Cir. 1998) (setting aside a contrary panel decision, and holding that a medical expert’s diagnostic opinion, while highly subjective and not “scientific” in the classic sense of being subject to experimental falsifiability, must nonetheless be scrutinized under the *Daubert* gatekeeping factors). While *Moore* involved a clinical diagnostic and causation opinion, the gatekeeping role could also be applied to standard of care expert testimony in medical malpractice cases, including whether the defendant health care providers conform to or deviate from standard medical judgment or practice. In recognition of the fact that clinical medical testimony often is based on experience and observation of the plaintiff and similar cases, and on what is known as “differential diagnosis”—ruling out other possible causes—many courts have not required clinical causation experts to

exercising this gatekeeping role, courts may evaluate the expert's conclusions as well as the methodology, and, no matter how outcome determinative the evidentiary rulings, or no matter how much they really rest on substantive legal decisions, district judge's decisions may only be reviewed for abuse of discretion,⁴⁵ rather than reviewed *de novo*, as the substantive legal rulings they really often are.

The next section explores how this evidentiary trilogy—particularly *Daubert* and *Joiner*, because it is still too early to assess the impact of *Kumho Tire*—has enabled courts to effect significant substantive legal changes in the tort law regarding the type and strength of causal proof that plaintiffs are required to produce.

III. USING THE GATEKEEPING ROLE TO CHANGE THE SUBSTANCE OF TORT LAW: ELEVATING THE SUBSTANTIVE REQUIREMENTS FOR PROVING CAUSATION

The *Daubert* decision stimulated an extensive amount of commentary, including ruminations on the epistemological dilemma of how judges without scientific training could evaluate the scientific validity and reliability of evidence, and whether in trying to do so, judges would be making scientific determinations or legal rulings.⁴⁶ A review of post-*Daubert* cases reveals that many judges have resolved this apparent epistemological conundrum by making what are plainly legal rulings about the type and nature of scientific evidence plaintiffs must produce to even attempt to prove causation. Some judges have used their gatekeeper power to wade into disputed scientific territory to try to resolve or choose sides in scientific controversies. Ignoring the

support their opinions with epidemiological studies. *See, e.g.*, *Kennedy v. Collagen Corp.*, 161 F.3d 1226 (9th Cir. 1998); *Zuchowitz v. United States*, 140 F.3d 381 (2d Cir. 1998); *McCulloch v. H.B. Fuller Co.*, 61 F.3d 1038 (2d Cir. 1995).

45. Some appellate courts are still trying to perform meaningful appellate review, and to hew to the traditional line that in making admissibility decisions, district courts are not to evaluate the weight to be given to competing expert claims. Meaningful appellate review is procedurally more possible when a trial judge decides that a plaintiff's expert causation testimony is insufficient to raise a jury issue, and thus grants summary judgment to the defendant, or judgment as a matter of law setting aside a jury verdict. The standard of review for summary judgment and judgment as a matter of law is *de novo*, not the relatively cursory abuse of discretion standard. *See, e.g.*, *Kennedy v. Collagen Corp.*, 161 F.3d 1226 (9th Cir. 1998). In *Kennedy*, the court reversed the trial judge's grant of summary judgment in rejecting plaintiffs' expert causation testimony. The Ninth Circuit held that trial judges "should not exclude expert testimony simply because they disagree with the conclusions of the expert." *Id.* at 1230. *See also In re Joint E. & S. Dist. Asbestos Litig.*, 52 F.3d 1124 (2d Cir. 1995) (reversing trial judge's grant of judgment notwithstanding the verdict to defendant, and finding that trial judge had rejected plaintiffs' expert causation testimony because of disagreement with expert's conclusions).

46. *See, e.g.*, Alexander Capron, *Daubert and the Quest for Value-Free "Scientific Knowledge" in the Courtroom*, 30 U. RICH. L. REV. 85 (1996); Farrell, *supra* note 3, at 2183.

usual legal rule that decisions about the admissibility of evidence are meant to be distinct from decisions about the weight or sufficiency of evidence, these courts have adopted the legal position that there is no difference under *Daubert-Joiner* between the sufficiency and admissibility determinations.⁴⁷ Other judges, perhaps wishing to steer clear of the difficult task of parsing scientific evidence, have instead seized on familiar legal criteria such as “relevance” to make substantive causation law rulings with precedential effects well beyond the particular evidence at issue in the case. Several courts have also avoided the difficult implications of having to assess the reliability of scientific methodology by turning the supposedly flexible and advisory *Daubert* factors of peer review, publication, and degree of widespread acceptance in the scientific community into simplistic legal litmus tests.⁴⁸

The most noteworthy trend in using admissibility determinations to make substantive causation rules—noteworthy because it is seriously scientifically and legally misguided—is a growing number of court rulings stating that in order for an expert opinion about causation to be relevant and thus admissible, the expert must base her testimony on epidemiological studies that demonstrate the product in question at least doubles the risk of the disease from which plaintiff suffers.⁴⁹

47. See, e.g., *Merrell Dow Pharms., Inc. v. Havner*, 953 S.W.2d 706 (Tex. 1997); *Minnesota Mining & Mfg. Co. v. Atterbury*, 978 S.W.2d 183, 189, 192 (Tex. App. 1998). See also *In re Joint E. & S. Dist. Asbestos Litig.*, 827 F. Supp. 1014, 1025 (S.D.N.Y. 1993) (finding plaintiffs' evidence insufficient because the judge found their studies subject to criticism, or found the defense studies more conclusive or less plagued by criticism). The Second Circuit reversed, holding that by doing so the judge had impermissibly crossed the boundary between evaluating the sufficiency of evidence and substituting his judgment about its weight or credibility for the jury's. *In re Joint E. & S. Dist. Asbestos Litig.*, 52 F.3d 1124, 1137 (2d Cir. 1995). The appellate court also rejected defendants' argument that *Daubert* eroded the line between admissibility and sufficiency of the evidence. See *id.* at 1132-33. This tendency of judges to blur questions of the sufficiency of evidence and the admissibility of evidence started in the Bendectin cases. Despite growing epidemiological support for the defendants, some juries continued to return verdicts in favor of plaintiffs, and judges grew openly skeptical of juries' abilities to correctly evaluate complex science. See Joseph Sanders, *Scientific Validity, Admissibility, and Mass Torts after Daubert*, 78 MINN. L. REV. 1387, 1431-33 (1994).

48. Examples of courts simplistically using *Daubert* factors such as peer review, publication, and general acceptance include *Forsyth v. Eli Lilly Co.*, 1998 U.S. Dist. Lexis 541 (D. Haw. 1998), *Minnesota Mining & Manufacturing Co. v. Atterbury*, 978 S.W.2d 183 (Tex. App. 1998). See also *Allison v. McGhan Medical Corp.*, 184 F.3d 1300 (11th Cir. 1999), where the court seemed to do little more than check off as if from a list whether an expert's theories had been published, peer reviewed, and were generally accepted. *Id.* at 1313. The court of appeals affirmed the decision to exclude the testimony, but for more searching reasons. *Id.* at 1322. Some commentators specifically warned that *Daubert* presented a risk that some judges would grasp at simplistic formulations as screening devices. See Carl F. Cranor et al., *Judicial Boundary Drawing and the Need for Context-Sensitive Science in Toxic Torts After Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 16 VA. ENVTL. L.J. 1 (1996).

49. Epidemiological studies ascertain a “relative risk” for a certain disease if one is exposed to a product or toxin. Epidemiologists compare the number of cases of a disease in a sample group

Courts have used two primary routes to arrive at the legal proof requirement of epidemiology with a relative risk of 2.0 or greater. One is to use the "reliability" factor from *Daubert*, which is supposed to focus on the validity and scientific soundness of the methodology, to declare any epidemiological study with a lower relative risk "unreliable" as a matter of law.⁵⁰ The second common move is to conflate the admissibility criteria of relevance with burden of proof, to reason that since a plaintiff's burden of proof is to show that the product to which she was exposed "more likely than not" caused her disease, and an epidemiological study that shows a doubling of risk (a relative risk of 2.0 in statistical terms) means that it is 50% likely that any particular case of the disease is attributable to the exposure rather than unex-

of the unexposed population, known as the "background risk," with the number of cases of the disease in a sample group of people exposed to the factor being studied. See Linda A. Bailey et al., *Reference Guide on Epidemiology*, in FEDERAL JUDICIAL CENTER, REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 168 (1994). The resulting ratio is known as the relative risk. *Id.* A relative risk of 2.0 means there are twice as many cases of the disease in the exposed group than in the unexposed group. *Id.* To put it another way, it means that the increased risk above background levels faced by the group is 100%. When this group risk is translated to the individual level, it means that in the case of any person with the disease, it is 50% likely that their illness is attributable to their exposure. *Id.* For other excellent explanations of the principles and methods of epidemiology, including the relative risk concept, see SANDERS, *supra* note 5, at 47-60; Michael D. Green, *Expert Witnesses and Sufficiency of Evidence in Toxic Substance Litigation: The Legacy of Agent Orange and Bendectin Litigation*, 86 Nw. U. L. REV. 643, 647 (1992); Thompson, *supra* note 3, at 250.

Courts also require that the risk ratio in a study be "statistically significant," which is a statistical measurement of the likelihood that any detected association has occurred by chance, or is due to the exposure. Tests of statistical significance are intended to guard against what are called "Type I" errors, or falsely ascribing a relationship when there in fact is not one (a false positive). See SANDERS, *supra* note 5, at 51. The discipline of epidemiology is inherently conservative in making causal ascriptions, and regards Type I errors as more serious than Type II errors, or falsely assuming no association when in fact there is one (false negative). Thus, epidemiology conventionally requires a 95% level of statistical significance, i.e. that in statistical terms it is 95% likely that the association is due to exposure, rather than to chance. See *id.* at 50-52; Thompson, *supra* note 3, at 256-58. Despite courts' use of statistical significance as an evidentiary screening device, this measurement has nothing to do with causation. It is most reflective of a study's sample size, the relative rarity of the disease being studied, and the variance in study populations. Thompson, *supra* note 3, at 256.

Some judges have improperly blurred the two very distinct concepts of relative risk and statistical significance by labeling the results of epidemiological studies which derived a relative risk of less than 2.0 as "statistically insignificant." See, e.g., *Maiorana v. U.S. Mineral Products Co.*, 827 F. Supp. 1014, 1041-42 (S.D.N.Y. 1993) describing epidemiological studies that yielded relative risks between 1.0 and 1.5 as "statistically insignificant".

50. See, e.g., *Allison v. McGhan Med. Corp.*, 184 F.3d 1300, 1315 (11th Cir. 1999) (affirming the district court's rejection of epidemiological studies as unreliable because with a relative risk below 2.0 there cannot be a consensus in the general scientific community that an association exists). The *Allison* court also noted that this verged on resurrecting the *Frye* general acceptance test, but read *Joiner* as essentially permitting that resurrection. *Id.* See also *Minnesota Mining & Mfg. Co. v. Atterbury*, 978 S.W.2d 183, 199-202 (Tex. App. 1998); *Kelley v. American Heyer-Schulte Corp.*, 957 F. Supp. 873, 878 (W.D. Tex. 1997).

plained causes, or “background risk,” then only such epidemiological studies satisfy the burden of proof.⁵¹ Causation evidence that does not satisfy the burden of proof, these courts reason, is not “relevant” to the issue of causation, and thus can be excluded.⁵² These courts also reason that since epidemiology is the only type of science that studies causal associations in human populations, other types of scientific evidence, such as animal studies, toxicology reports, chemical structure analysis, and clinical differential diagnosis, even if “scientifically valid” evidence of the type relied on by scientists to make risk assessments or diagnostic causal attributions, are not legally relevant unless also supported by epidemiology with the requisite increase in relative risk.⁵³

This development in causation law initially appeared in Judge Weinstein’s opinion in the Agent Orange litigation,⁵⁴ and judges presiding over Bendectin cases picked up on it and increasingly adopted it as a

51. In the eyes of some judges, The Federal Judicial Center’s *Reference Guide on Epidemiology* seems to endorse the approach of making admissibility determinations turn on the ultimate burden of proof, when it informs judges:

The civil burden of proof is described most often as requiring the fact finder to “believe that what is sought to be proved . . . is more likely true than not true.” The relative risk from an epidemiological study can be adapted to this 50% plus standard to yield a probability or likelihood that an agent caused an individual’s disease. The threshold for concluding that an agent was more likely the cause of a disease than not is a relative risk greater than 2.0.

Reference Guide on Epidemiology, *supra* note 49, at 168-69. See, e.g., Merrell Dow Pharm.s, Inc. v. Havner, 953 S.W.2d 706, 721-22 (Tex. 1997); *In re Hanford Nuclear Reservation Litig.*, 1998 U.S. Dist. Lexis 15028, at *15 (E.D. Wash. 1998). The Reference Manual, however, was not making any recommendation about admissibility of evidence or the sufficiency of all plaintiffs’ evidence taken as a whole. Nor was it saying that only epidemiology that reached this threshold could be relevant to the issue of causation, and neither was it instructing judges that no plaintiff with other types of scientific proof or with epidemiology with a lower relative risk but other strong evidence of individual causation could get before a jury. Rather, the Manual was attempting to educate judges about the principles of epidemiology, and illustrate those principles by equating them to legal concepts with which the judges were familiar. It was not offering prescriptive legal recommendations about how to rule on the evidence in any particular case.

52. See, e.g., Allison v. McGhan Med. Corp., 184 F.3d 1300, 1315 (11th Cir. 1999); Schudel v. General Elec. Co., 120 F.3d 991 (9th Cir. 1997); Daubert v. Merrell Dow Pharms., Inc., 43 F.3d 1311 (9th Cir. 1995); *In re Hanford Nuclear Reservation Litig.*, 1998 U.S. Dist. Lexis 15028 (E.D. Wash. 1998); Sanderson v. International Flavors & Fragrances, Inc., 950 F. Supp. 981 (C.D. Cal. 1996); *In re Breast Implant Litig.*, 11 F. Supp. 2d 1217 (D. Colo. 1998); Forsyth v. Eli Lilly Co., 1998 U.S. Dist. Lexis 541 (D. Haw. 1998); Hall v. Baxter Healthcare Corp., 947 F. Supp. 1387 (D. Or. 1996); *Havner*, 953 S.W.2d at 706.

53. See, e.g., *In re Breast Implant Litig.*, 11 F. Supp. 2d at 1217; *Hall*, 947 F. Supp. at 1387.

54. *In re “Agent Orange” Product Liability Litigation*, 611 F. Supp. 1223 (E.D.N.Y. 1985). Professor Michael Green provides an extensive discussion of how the Agent Orange litigation has influenced subsequent legal causation developments. See Green, *supra* note 49, at 671-74. For an in-depth story of this seminal toxic tort litigation and the factors that influenced Judge Weinstein to rule as he did, see PETER SCHUCK, *AGENT ORANGE ON TRIAL* (1986). Both Professors Green and Schuck are highly critical, both from a scientific and legal perspective, of the way

substantive rule.⁵⁵ For example, on remand from the Supreme Court, the Ninth Circuit in *Daubert* reasoned that because the burden of proof in tort is “more likely than not,” only epidemiology that showed the increase in the relative risk to be 2.0 or greater could meet the “fit,” or relevance prong elucidated by the Supreme Court.⁵⁶ This principle of causation law—that to sustain a toxic exposure claim, plaintiffs must be able to offer statistically significant epidemiology that demonstrates that the risk of disease is at least doubled upon exposure to the product in question—has now been firmly entrenched in silicone gel breast implant litigation as well.⁵⁷

The rejection of plaintiffs’ expert testimony on causation may be justified in the context of Bendectin and breast implant litigation, because during the course of the litigation, the body of epidemiology grew and matured, and failed to confirm any substantial increase in risk. This cast plaintiffs’ experts in the position of arguing against a developing scientific consensus and having to admit that their opinion as to a causal relationship was fraught with uncertainty. The well developed epidemiological context of these two medical products is a significant explanatory reason why courts increasingly came to reject admitting plaintiffs’ expert testimony.⁵⁸

in which Judge Weinstein elevated epidemiology to make it foundational to a plaintiffs’ causation case. See SCHUCK, *supra*; Green, *supra*.

55. See, e.g., *Daubert v. Merrell Dow Pharms., Inc.*, 43 F.3d 1311 (9th Cir. 1995); *Brock v. Merrell Dow Pharms., Inc.*, 874 F.2d 307 (5th Cir. 1989).

56. 43 F.3d at 1320-22.

57. See, e.g., *Allison*, 184 F.3d at 1300; *In re Breast Implant Litig.*, 11 F. Supp. 2d at 1217; *Kelley v. American Heyer-Schulte Corp.*, 957 F. Supp. 873 (W.D. Tex. 1997); *Hall*, 947 F. Supp. at 1387; *Minnesota Mining & Mfg. Co. v. Atterbury*, 978 S.W.2d 183 (Tex. App. 1998). Contrary opinions in breast implant litigation, rejecting epidemiology with a relative risk of greater than 2.0 as a required element for a plaintiff to prove causation, include *Vassallo v. Baxter Healthcare Corp.*, 696 N.E.2d 909, 917-18 (Mass. Sup. Ct. 1998); and *Jennings v. Baxter Healthcare Corp.*, 954 P.2d 829 (Or. App. 1998). See also *Hopkins v. Dow Corning Corp.*, 33 F.3d 1116, 1124 (9th Cir. 1994) (decided prior to publication of major epidemiological studies of breast implants and auto-immune diseases). The court in *Hopkins* held that where there is no solid body of epidemiology to review, it is appropriate to admit expert testimony that draws on other types of valid and reliable scientific data). *Id.* at 1124-1125. Cf. *Pick v. American Med. Sys., Inc.*, 958 F. Supp. 1151 (E.D. La. 1997) (silicone penile implant case where plaintiff’s experts relied largely on epidemiology concerning silicone breast implants). The court rejected the proposition that epidemiology with a relative risk greater than 2.0 is required to admit general causation testimony. *Id.* at 1160.

58. See, e.g., *Ambrosini v. Labarraque*, 101 F.3d 129, 138-39 (D.C. Cir. 1996) (reasoning that Bendectin rulings were premised on the fact that there was an “overwhelming body of contradictory epidemiological evidence” aligned against plaintiffs’ experts, and should be limited to that context) (quoting *Richardson v. Richardson-Merrell*, 857 F.2d 823, 830 (D.C. 1988)). See also GREEN, *supra* note 12, at 307; Green, *supra* note 49, at 671-72 (arguing that the courts’ insistence on epidemiology for both admissibility and sufficiency of evidence rulings in Bendectin cases should be limited to situations where strong, well-developed, consistent body of epidemiology exists). Cf. SANDERS, *supra* note 5, at 197 (warning that Bendectin litigation sets a dangerous

The precedential effect of these rulings have been felt well outside their scientific contexts, however. Courts facing disputes about the admissibility or sufficiency of plaintiffs' expert causation testimony in cases dealing with other products have followed the "law" of Bendectin and breast implants to rule that without some epidemiological studies that show an increase in relative risk of 2.0 or more, plaintiffs have no relevant, reliable, or admissible evidence. Underscoring the precedential, substantive law treatment accorded the Bendectin and breast implant decisions, district courts in the Ninth Circuit, which produced the *Daubert II* ruling that the only relevant causation evidence is testimony based on epidemiology that shows a doubling of the risk, have been most likely to apply this rule to other products. And, they have adopted this substantive threshold for proof of causation even where there is not yet any epidemiology studying a product, or when the epidemiology is still very tentative and preliminary.⁵⁹

Because the post-*Daubert* breast implant cases have the potential for a growing influence on cases dealing with other products with quite different scientific records, it is worth parsing the reasoning of two of the most extensive breast implant discussions of the admissibility of plaintiffs' evidence. Although purporting to rest their admissibility decisions on the *Daubert* factors of scientific validity and reliability, the judges in these two cases failed to heed the advice of neutral scientific experts, and instead made substantive legal rulings based on relevancy and burdens of proof. The case of *Hall v. Baxter Healthcare Corp.*,⁶⁰ is illustrative of a judge's use of the legal criterion

example for other contexts because in Bendectin the well-developed epidemiological record that exonerated Bendectin "made it relatively easy to determine that substantial parts of the plaintiffs' expert testimony was 'bad' when judged by the ordinary criteria of science").

59. Non-Bendectin and non-breast implant cases in which courts insist on causation testimony must be predicated on epidemiology that demonstrates at least a doubling of the risk. See, e.g., *Schudel v. General Elec.*, 120 F.3d 991 (9th Cir. 1997) (alleging that workplace exposure to industrial solvents trichloroethane (TCA) and perchloroethylene (perc) caused neurological and respiratory impairments); *Sanderson v. International Flavors & Fragrances, Inc.*, 950 F. Supp. 981, 1000 (C.D. Cal. 1996) (connecting between aldehydes in perfume products and multiple chemical sensitivity); *Forsyth v. Eli Lilly Co.*, No. 95-00185, 1998 U.S. Dist. Lexis 541 (D. Haw. 1998) (connecting the anti-depressant drug Prozac and suicide); *In re Hanford Nuclear Reservation Litig.*, 1998 U.S. Dist. Lexis 15028 (E.D. Wash. 1998) (connecting exposure to radioactive emissions from nuclear weapons plant and various cancers and thyroid diseases). See also *Bartley v. Euclid*, 158 F.3d 261 (5th Cir. 1998) (in case claiming back injuries from vibrations in coal hauling equipment, court declined to reach issue of whether epidemiology with a relative risk of greater than 2.0 is a foundational legal requirement, but then ruling plaintiffs' expert causation testimony admissible because it meets this epidemiological threshold). This list does not purport to be exhaustive; it is a sampling of cases obtained through a post-*Daubert* date restricted LEXIS search for cases containing the terms "causation and epidemiology! and relative w/5 risk."

60. 947 F. Supp. 1387 (D. Or. 1996).

of “relevancy” to reject all scientific evidence that does not rest on epidemiology with a relative risk greater than 2.0, despite its scientific validity, and despite the fact that scientists reasonably rely on the types of rejected evidence to form judgments about causation. This case involved the consolidated claims of several plaintiffs who alleged that their leaking or ruptured silicone gel breast implants had caused various auto-immune diseases. The defendants brought a motion in limine to exclude all testimony by plaintiffs’ experts about a causal link between implants and the alleged diseases. To exercise its *Daubert* gatekeeping role, the trial judge appointed a panel of neutral scientific experts, pursuant to Federal Rule of Evidence 706, to review the parties’ experts’ submissions and scientific studies. The judge asked these neutral “technical advisors” from the fields of epidemiology, immunology/toxicology, rheumatology, and chemistry, to issue a report focusing on five questions drawn from *Daubert*: (1) Was a particular plaintiff’s expert’s opinion supported by scientific reasoning and methodology that is generally accepted in the field? (2) Is the opinion based upon scientifically reliable data? (3) If epidemiological studies are inconclusive, what other types of scientific evidence could justify a conclusion about the cause of the disease at issue? (4) Do the methodology and data support the expert’s conclusions? (5) Does the data relied on by the expert apply to the disease in issue in the case, or does it speak to some other disease or syndrome?⁶¹

The neutral epidemiology expert appointed by the court, Dr. Merwyn Greenlick, reviewed the opinions and underlying studies of the two epidemiology experts, Dr. Goldsmith for the plaintiffs, and the defense expert, Dr. Ory. If the trial judge had followed the advice of his own neutral scientific expert, and that expert’s interpretation of the *Daubert*-derived questions, the judge would have admitted the opinions of both experts, and left to the jury the question of which differing interpretation should be given greater weight. Dr. Greenlick concluded that both Dr. Goldsmith and Dr. Ory based their epidemiological opinions on scientifically valid data, and that their “somewhat different positions [were] a result of different, but legitimate, interpretations” of the sixteen underlying studies.⁶² Dr. Greenlick castigated a defense lawyer for inappropriately criticizing or diminishing Dr. Goldsmith’s testimony, and advised the judge that “Dr. Goldsmith’s statement is at the heart of science. Interpreted through the eyes of an epidemiologist, Dr. Goldsmith is saying that work in the area has

61. *Id.* at 1393-94.

62. *Id.* at 1448.

progressed to the point . . . where he and others have begun . . . to take the possibility of an association seriously.”⁶³

In response to the court’s question about what types of evidence could be used to determine causation when faced with inconclusive epidemiology, Dr. Greenlick advised the court that in light of the fact that the epidemiology was not definitively negative or positive, but was sufficient to raise a serious question about the possibility of an association between silicone implants and auto-immune diseases, it was scientifically appropriate also to consider animal studies, biophysical data, medical records, differential diagnosis and other types of scientific data. He analogized the causation issue facing the court to that of a clinician treating a patient. A clinical treating physician cannot wait for the slow evolution of epidemiological certainty about population-wide issues, but must make a judgment on all available scientifically legitimate sources about what is causing the disease in a particular patient.⁶⁴ Dr. Greenlick then specifically concluded that “there is sufficient scientific data upon which to base the opinions both of Dr. Goldsmith and Dr. Ory, the principal epidemiology witnesses, and I believe it would be appropriate to accept both their testimonies.”⁶⁵

Judge Jones did not follow this scientific expert advice. He decided to exclude the testimony of Dr. Goldsmith⁶⁶ not because he disagreed with the neutral expert’s scientific analysis, but because he ruled that as a matter of law, to prove causation “plaintiffs must be able to show a relative risk of greater than 2.0.”⁶⁷ Consequently, any epidemiological expert’s opinion testimony, no matter how scientifically valid, is simply not legally relevant to the issue of causation unless it rests on studies showing the requisite doubling of risk. In arriving at this ruling, Judge Jones relied not on his scientific experts’ reports or on scientific treatises, but on prior court decisions and judicial evidence manuals—classic sources of law. Thus, this ruling, like those precedents upon which he relied, is not an example of a judge following the standards of good science to make a case specific and fact specific

63. *Id.* at 1448-49.

64. *Id.*

65. *Id.* at 1450.

66. The judge did not have to rule on the admissibility of Dr. Ory’s testimony because once he excluded all of the plaintiffs’ expert causation testimony, the plaintiffs had no case, and there was no need for the defense to put on any evidence. This demonstrates the inherent one-sidedness of the *Daubert* rule—it is most likely applied against plaintiffs, but defense experts rarely have to be subjected to the same scrutiny, although neutral experts and judges in fact take defense experts’ criticisms of plaintiffs’ experts’ conclusions into account when making admissibility determinations about plaintiffs’ experts.

67. *Hall*, 947 F. Supp. at 1403.

evidentiary admissibility determination. It is a legal policy decision, a conscious choice to reject scientific principles as legally irrelevant, which could set a precedent with potentially binding relevance far beyond the realm of breast implant litigation.

Another important feature of the new toxic causation law fashioned by the judge in *Hall* is that it conflates the plaintiffs' overall burden of proof on causation with the determination of whether to admit any particular piece of evidence. Once the court declared the legally privileged and determinative role of statistically significant epidemiology with a relative risk of 2.0 or greater, and then excluded the plaintiffs' epidemiology experts for not meeting this legal standard, the court proceeded to exclude the remainder of the plaintiffs' scientific experts because of the lack of epidemiology experts.⁶⁸ What the court did, essentially, was to evaluate each proposed expert's testimony and assess whether it, standing alone, would be sufficient to satisfy the plaintiff's overall burden of proof to show that causation is more likely than not. The judge concluded that none of it was legally relevant or sufficient because it did not rest on admissible epidemiology.

In the guise of evidentiary admissibility determinations, the court made sufficiency determinations about each individual piece of evidence. Although admissibility decisions are made on individual items of evidence, this obvious fact does not logically lead to the conclusion that in order to be admissible, each individual item of evidence must be itself sufficient, standing alone, to present a jury question.⁶⁹ To so hold creates an admissibility and sufficiency criterion that is virtually impossible for any plaintiff to meet, because it deprives plaintiffs of

68. Again, Judge Jones had to disregard the advice of one of his own appointed neutral experts. The neutral chemistry expert, Dr. Robert McClard, concluded that the testimony of plaintiffs' expert Dr. Garrido, who formed an opinion about the capability of silicone to degrade to silica, a well-known harmful agent, in the body, was supported by scientifically valid reasoning and methodology, and was clearly scientifically relevant. *Id.* at 1473 app. e. See also *Vassallo v. Baxter Healthcare Corp.*, 696 N.E.2d 909 (Mass. 1998), where the trial judge relied on the same report from Dr. McClard to conclude that the testimony of the plaintiff's expert was admissible under the *Daubert* criteria. *Id.* at 916.

69. As the Second Circuit noted in cautioning trial judges that they must keep the two inquiries distinct, "admissibility entails a threshold inquiry over whether a certain piece of evidence ought to be admitted at trial," while the sufficiency of the evidence inquiry "asks whether the collective weight of a litigant's evidence is adequate to present a jury question." *In re Joint E. & S. Dist. Asbestos Litig.*, 52 F.3d 1124, 1132 (2d Cir. 1995). The Second Circuit held that *Daubert* was not meant to erode the line between admissibility determinations and sufficiency determinations, and that in making admissibility determinations trial judges should not engage in a parsing of the weight to be given an expert's conclusions. *Id.* Accord *Berry v. CSX Transp., Inc.*, 709 So. 2d 552 (Fla. App. 1998). Florida still follows the *Frye* "general acceptance" standard, but the court has since ruled that it is inappropriate for trial judges to assess the weight to be given to a study, or whether an expert's conclusions are unassailable or subject to some criticism, in making admissibility determinations.

the entitlement to have their evidence evaluated as a cumulative whole unless they can produce the magic bullet of the requisite strength of epidemiological studies.⁷⁰ Justice Stevens, writing separately in *Joiner*, warned against district judges using their gatekeeping power to erect such a legal barrier, because it is contrary both to *Daubert* and to the methodologies of sound science which *Daubert* commanded courts to respect. Castigating the district judge for separately examining one by one each study on which plaintiffs' experts relied, rather than assessing them in their cumulative whole, Justice Stevens noted that many scientists, including the defendants' experts, used the cumulative weight of the evidence approach in making causal judgments. "It is not intrinsically 'unscientific' for experienced professionals to arrive at a conclusion by weighing all available scientific evidence—this is not the sort of 'junk science' with which *Daubert* was concerned."⁷¹ But the district judge in *Hall* in effect declared it legally invalid for scientists to look at the cumulative whole of the scientific record by, for example, adding to epidemiology with relative risk above 1.0 but below 2.0, evidence gleaned from animal studies, toxicology, chemical analyses, pathological examinations of plaintiff's tissues, and the differential diagnosis technique of ruling out other known causal agents, to arrive at an opinion about causation.

In a subsequent breast implant case that relied heavily on *Hall*, as well as Bendectin precedents, *In re Breast Implant Litigation*,⁷² the United States District Court in Colorado refined *Hall's* principle of causation law by linking the requisite type of epidemiology to the two elements of causation lurking in every toxic tort case: general and specific causation. The general causation inquiry asks whether the product at issue is capable, in general, of causing the disease. Clearly, epidemiology, which is the study in large populations of associations between exposure and illness, is highly relevant evidence to suggest general causation. The Colorado court, however, declared it to be the *only* relevant evidence on general causation, and only if an epidemio-

70. In contrast to the cases discussed *supra*, courts in Texas have interpreted *Daubert* as eroding the distinction between the admissibility of evidence and the sufficiency of the evidence to prove causation. See, e.g., *Minnesota Mining & Mfg. Co. v. Atterbury*, 978 S.W.2d 183, 189 (Tex. App. 1989) (interpreting the Texas Supreme Court decision in *Merrell Dow Pharmaceuticals Inc. v. Havner*, 953 S.W.2d 706 (Tex. 1997), as adopting the same criteria for the two inquiries). The *Atterbury* court explained that defendants have "two bites at the *Daubert* apple"—a motion to exclude the evidence, and then, if the trial judge does not grant that, a motion challenging the sufficiency of the evidence. *Atterbury, supra* at 192.

71. *Id.* at 153.

72. 11 F. Supp. 2d 1217 (D. Colo. 1998).

logical study demonstrates a doubling of risk in the general population.

This ruling is statistically, scientifically, and legally suspect. It is legally dubious because it uses a specific causation standard—whether it is more likely than not that the product at issue caused the plaintiff's illness—to adopt a substantive rule about the type and strength of evidence required for general causation. In other words, the court made the admissibility of general causation evidence hinge on whether it, standing alone, would also meet the specific causation burden of proof. By thus collapsing the specific and general causation inquiries, the Colorado court ruled that as a matter of law, unless plaintiffs can show that for every single member of the exposed population, it is 50% or more likely that her illness is attributable to her exposure to the product at issue, then no individual plaintiff is entitled to try to show that she is one of the ones for whom a population-wide risk of something less than 50% risk became reality.

The plaintiffs' case proceeded on the basis that because epidemiological studies did show some increased risk of auto-immune diseases among women exposed to silicone gel implants, which did suggest that a certain percentage (although not 50%) of women with such diseases may in fact have become ill because of their implants, they should have been able to offer additional kinds of scientific evidence which tended to show that it was in fact more likely than not that the individual plaintiffs were among those whose illnesses could be linked to their implants. The plaintiffs were seeking to move to the specific causation case because the epidemiology did show some increased risk in general, and, when coupled with other scientific evidence, certainly did not rule out general causation. The court, however, refused to admit all of the scientific evidence which spoke to specific causation in the individual plaintiffs, such as clinical testimony, differential diagnosis, and examination of silicone in tissue samples of the plaintiffs, not because it lacked scientific validity or relevance to the specific causation issues, but because it did not "fit" the general causation issue because it did not rest on epidemiology with a relative risk greater than 2.0.⁷³

In addition to collapsing general causation into the standard for proving specific causation, this ruling also collapsed the plaintiff's burden of proof on the fact of individual causation into the population-

73. See *In re Breast Implant Litig.*, 11 F. Supp. 2d at 1232 (discussing specific causation testimony of Dr. Kassan); 1236-37 (discussing specific causation testimony of Dr. Klapper).

wide probability of causation.⁷⁴ As a matter of statistics, it is improper to equate a population-wide risk in an epidemiological study with the individual probability of causation. Population-wide risk estimates simply do not address, and thus cannot be translated to, the probability of causation in any one individual.⁷⁵ Epidemiology, which is inherently population based, represents a range of values across the group, not a single risk probability value that can be applied to any individual.⁷⁶ The population-wide risk probability cannot take into account particular susceptibilities of individuals based on factors such as genetics,⁷⁷ overall state of health, age, gender, race, amount and duration of exposure. Furthermore, the probability of individual causation cannot be computed solely from the population-wide relative risk, because in some individuals the exposure may accelerate the onset of disease that they might have gotten later in life as a result of age or exposure to other risk factors, and thus in fact, be a substantial contributing factor to the individual's disease. But statistically, these cases of accelerated onset tend to get accounted for as background risk cases, rather than as exposure cases, which can lead to mathemati-

74. See Steve Gold, *Causation in Toxic Torts: Burdens of Proof, Standards of Persuasion, and Statistical Evidence*, 96 YALE L.J. 376 (1986). As Gold explains, while traditionally only the preponderance of the evidence standard of persuasion was defined probabilistically,

in toxic torts, the statistical causation evidence is also expressed probabilistically - as a factual estimate of the defendant's contribution to the plaintiff's risk. The failure to distinguish between the two kinds of probability has led to the collapse of the factual burden and preponderance standard into a single test: does the factual probability of causation exceed 50%?

Id. at 379. As Gold then demonstrates, this subtly yet substantively alters causation law, tightening the standard plaintiffs must meet, and narrowing the evidentiary focus to probabilistic statistical evidence such as epidemiology—precisely what the courts effectuated in *Hall* and *In re Breast Implant Litigation*. *Id.* at 392.

75. *Id.* at 390.

76. See James Robins & Sander Greenland, *The Probability of Causation Under a Stochastic Model for Individual Risk*, 45 BIOMETRICS 1125, 1126 (1989); Thompson, *supra* note 3, at 254-56.

77. See, e.g., Cranor et al., *supra* note 48, at 39-40. For example, population-wide breast implant studies included in the exposed groups both women with implants which were still intact, and presumably, given the high rupture rates, women whose implants had ruptured, whether they knew of the rupture or not. Thus, the studies were not designed to take into account whether an individual's implant had ruptured, how long ago it had ruptured, how much silicone had leaked into her body, or where it had migrated to. Yet these individualistic factors of the amount and duration of exposure to silicone in the body could affect an individual's susceptibility in immune responses, i.e. they might make it more likely in an individual case that the individual was one of the people upon whom the slightly elevated population-wide relative risk in fact came to rest.

cal underestimation of both the population wide relative risk, and the individual risk probability.⁷⁸

Statisticians, epidemiologists, and many other types of scientists understand the numerous reasons why it is invalid to equate the increase in relative risk in the population being studied with the probability of causation in any individual case. But these reasons seem to have eluded the comprehension of those judges who have made the misleadingly simple equation of legal relevance and the burden of proof on either general or specific causation with epidemiology that demonstrates a relative risk of 2.0 or greater. Dr. Greenlick, the court-appointed neutral expert in *Hall*, argued that from the perspective of science, and epidemiology in particular, plaintiffs should be able to try to prove a specific causation case when the epidemiology shows a relative risk of less than 2.0:

From a scientific point of view it is not appropriate to disregard relative risks of less than 2.0. First of all, relative risk is a term that applies to a population, not to an individual. While it is possible to estimate the average increased risk for members of a population, it is really not appropriate to assume that each individual in a population actually had a similar risk. It is much more appropriate to believe that the average increased risk is made up of a wide range in individual risks in the population⁷⁹

Even Dr. Marcia Angell, the executive editor of the *New England Journal of Medicine*, who has been so stridently critical of the breast implant litigation, and argues forcefully for elevating epidemiology above all other forms of scientific evidence in the courtroom,⁸⁰ has had to acknowledge that one of the breast implant epidemiological studies showing a relative risk of 1.24,⁸¹ could mean that breast implants have in fact caused the auto-immune disease in some individual women. This relative risk could mean, she writes, that in every two out of twelve ill women with breast implants,

the implants were the sole cause of their disease and in the other 10 they played no role. Or it could mean that implants played a major role in 3 or 4 women and a very small one in the others. Or it could

78. See Sander Greenland, *Relation of Probability of Causation to Relative Risk and Doubling Dose: A Methodologic Error That Has Become a Social Problem*, 89 AM. J. PUB. HEALTH 1166, 1167 (1999).

79. *Hall v. Baxter Healthcare Corp.*, 947 F. Supp. 1387, 1450-51 app. B (D. Or. 1996).

80. See MARCIA ANGELL, *SCIENCE ON TRIAL* (1996).

81. Charles H. Hennekens et al., *Self-reported Breast Implants and Connective-Tissue Diseases in Female Health Professionals*, 275 JAMA 616, 618 (1996).

mean that implants contributed a varying amount to the disease in all 12.⁸²

Some courts who have adopted the evidentiary requirement of epidemiology with a relative risk greater than 2.0 apparently, but erroneously think they are just adhering to the teaching of *Daubert* by following the accepted tenets of the scientific discipline of epidemiology. Other courts, displaying a far better understanding of science and the limits of epidemiology, have explicitly rejected the substantive causation rule that plaintiffs must have epidemiology with a relative risk of greater than 2.0 in order to prove causation. Adhering to the precepts of epidemiology, these courts have recognized that because studies that define risk factors do not address the cause of a disease in any particular person, it is irrational and unfair to require an individual plaintiff to produce a magic bullet level of increased population-wide risk. As a New Jersey court explained:

an epidemiologist cannot state that it is more likely than not that a particular case of colon cancer was caused by the asbestos. . . . A medical doctor, however, or even one otherwise acquainted with the physiology of a particular patient and the progress of the disease, may make a medical judgment concerning the origin of the disease,

factoring together epidemiological studies, other types of scientific evidence, and known individual risk factors or individual susceptibility that might enhance the risk of exposure, "even though the risk in the study fell short of the 2.0 correlation."⁸³ The court went on to reject the defendants' argument that experts should not be able to testify to causation unless epidemiology established a threshold relative risk of 2.0 as "proving too much." Under the defendants' proposed rule, no plaintiff could recover if a study showed the increased risk was 1.99, but all could potentially recover if it were 2.1, a result that "makes little sense, scientifically or legally."⁸⁴

Similarly, other courts rejecting the epidemiological threshold have recognized that scientists themselves factor in evidence other than epidemiology in making causal judgments, especially individual causal judgments. It is scientifically legitimate to use epidemiological evi-

82. ANGELL, *supra* note 80, at 197 (quoted in *In re Breast Implant Litigation*, 11 F. Supp. 2d 1217, 1226 (D. Colo. 1998)). Judge Sparr failed to appreciate the import of this quote, however, when he rejected all of the plaintiffs' specific causation testimony because the elevated risk in the population in general was not greater than 50%.

83. *Grassis v. Johns-Manville Corp.*, 591 A.2d 671, 675 (N.J. Super. Ct. App. Div. 1991) (decision that epidemiology of a relative risk greater than 2.0 was not required to prove causation later adopted by the New Jersey Supreme Court in *Landrigan v. Celotex Corp.*, 605 A.2d 1079 (N.J. 1992)).

84. *Id.*

dence that falls short of the 2.0 relative risk level in combination with clinical or other experimental evidence which strengthens the inference of causation in an individual case.⁸⁵ Other courts have allowed expert medical causation testimony when no epidemiology had even been conducted on whether there is a statistical link between exposure to the toxin in question and the plaintiff's illness, recognizing that it can be medically and scientifically acceptable to base diagnostic judgments on other types of evidence.⁸⁶ As the Third Circuit explained:

[i]n the actual practice of medicine, physicians do not wait for conclusive, or even published and peer-reviewed, studies to make diagnoses to a reasonable degree of medical certainty. However, experience with hundreds of patients, discussions with peers, attendance at conferences and seminars, detailed review of a patient's family, personal, and medical histories, and thorough physical examinations are the tools of the trade, and should suffice for the making of a differential diagnosis even in those cases in which peer-reviewed studies do not exist to confirm the diagnosis of the physician.⁸⁷

The consistency of this reasoning with accepted precepts of science and medicine is reflected in its similarity to the recommendation of Dr. Greenlick, the neutral expert epidemiologist in *Hall*. Displaying notable perceptiveness that the causation issue in a toxic tort case is more like the individual clinical medical judgment than the population-wide statistical risk estimates of the epidemiologist, he noted that courts, like clinicians, have to make concrete decisions about what is causing a particular person's illness, and often have to do so in the face of less than definitive science. As Dr. Greenlick explained, clinicians making diagnoses do not have the luxury of waiting for epidemiology to arrive at more conclusive, relatively certain decisions about risk levels in general populations.

Physicians [must] do the best they can in an uncertain situation. They use all of the sources of information at their disposal, including animal studies, case reports, small or inconclusive epidemiologi-

85. See, e.g., *In re Joint E. & S. Dist. Asbestos Litig.*, 52 F.3d 1124 (2d Cir. 1995); *Ambrosini v. Labarraque*, 101 F.3d 129 (D.C. Cir. 1996).

86. See, e.g., *Heller v. Shaw Indus., Inc.*, 167 F.3d 146 (3d Cir. 1999); *Kennedy v. Collagen Corp.*, 161 F.3d 1226 (9th Cir. 1998); *Zuchowitz v. United States*, 140 F.3d 381 (2d Cir. 1998); *Benedi v. McNeil-PPC Inc.*, 66 F.3d 1378 (4th Cir. 1995); *McCulloch v. H.B. Fuller Co.*, 61 F.3d 1038 (2d Cir. 1995).

87. *Heller*, 167 F.3d at 155. This case demonstrates that this does not mean abandonment of all judicial scrutiny of an expert's opinion testimony. However, the court went on to approve the exclusion of the central part of plaintiff's medical expert testimony because the experts' conclusions were not supported by the temporal relationship between exposure and onset of symptoms, and cessation of exposure and the end of symptoms. *Id.* at 164-65.

cal studies, expert opinions, conversations with colleagues, and, most importantly, their own clinical experience and judgment. . . . [P]hysicians are not dealing with the general scientific question "What causes disease in women?"; they are dealing with the particularistic clinical question "What is causing disease in this particular woman?" Those are very different questions. And the physician must draw some working causal model in the case of a single patient, even in the face of a great deal of uncertainty. The scientist has the luxury of reporting that there isn't yet sufficient data to draw a conclusion. That luxury isn't available to the clinician, because the decision to do nothing in a clinical situation is selecting a specific course of action.⁸⁸

The courts that insist on epidemiology of a certain risk probability have not explicitly rejected the view that the causation judgment is more akin to an individual clinical judgment than a probabilistic risk assessment. Instead, they acknowledge that they are not following the principles of science, but are making a legal policy determination to equate epidemiology, relative risk, general causation, and the burden of proof on individual causation. Conceding that epidemiology and other methods of science cannot directly prove causation, the Texas Supreme Court in *Merrell Dow Pharmaceuticals, Inc. v. Havner* explained its decision to nevertheless require plaintiffs to produce epidemiological studies with a relative risk of greater than 2.0 as a policy determination that "strikes a balance between the needs of our legal system and the limits of science."⁸⁹ The California Federal District Court in *Sanderson v. International Flavors and Fragrances, Inc.*,⁹⁰ was even more blunt about the policy reasons underlying its decision. When faced with the plaintiff's argument that the fact that her alleged problem had never made its way onto the research agenda of any epidemiologist created an unfair dilemma, the court said her only option was to "Wait." A question that is capable of being answered by science should await such an answer, and the courts, given their needs for certainty, efficiency, finality, and social legitimacy, should await the "laggardly pace" of the scientific research process.

Given the difficulties in legal resolution of scientific controversies . . . and the great benefit to society in scientific resolution of scientific controversies - meaning the progressive and continuous rejection, refinement, and confirmation of hypotheses - the Court cannot avoid thinking that the framers of Rule 702 did well to make law a prisoner to science, not the other way around.⁹¹

88. *Hall v. Baxter Healthcare Corp.*, 947 F. Supp. 1387, 1449 (D. Or. 1996) (quoting Greenlick report).

89. *Id.* at 718.

90. 950 F. Supp. 981 (C.D. Cal. 1996).

91. *Id.* at 1004.

Thus, in the evolving post-*Daubert* case law, we have a sharp divergence of judicial opinion. One view is that judges should allow testimony if it is based on the types of evidence reasonably relied on by scientists, including clinical diagnosticians, and should not insist on epidemiology, or a certain required level of relative risk in epidemiological studies, because science itself does not always insist on such evidence. Under this view, it is fundamentally erroneous, both as a matter of law and science, to adopt the illusory equation of population based relative risk with the burden of proof on causation. Under this view, the law does not have to await scientific or epidemiological certainty before considering a casual judgment, because scientists themselves do not so insist. The contrasting legal view is that relative risk and individual burdens of proof can, as a matter of legal policy, be equated, because that is a simple device for judges to screen out scientific evidence, especially when a scientific issue is novel, or hotly debated, or highly uncertain. According to this view, the law should wait for a relative degree of scientific certainty, at least in the form of a strong epidemiological risk association. Law thus exalts epidemiology above other scientific disciplines, and makes the same highly conservative value choice on causation that epidemiology itself makes: it is worse to conclude that there is a causal relationship when in fact there turns out not to be one (a Type I false positive error), than it is to conclude that there is no causal relationship when in fact there is such a relationship (a Type II false negative error).

The growing number of courts that are adopting this position, influenced by the legal rulings adopted in *Bendectin* and breast implant cases, are only considering one side of the normative and policy picture. They are insisting on a strict, statistically inaccurate adherence to burdens of proof on causation, because of reluctance to move the law ahead of scientific uncertainty, out of concern for the potentially adverse consequences for defendants whose products might erroneously be judged to have caused a particular type of harm. But these courts have failed to consider many important value and policy considerations: the consequences for the tort goal of deterring risky corporate behavior, the potential for discouraging a more adequate level of safety research, the synergistic effect that the legal system has on stimulating scientific research, the potential disparate impact on certain social groups who have traditionally been neglected by the scientific research community, and the burdens on injured people who may well have causally legitimate claims that they have been injured by an inadequately tested drug or carelessly made device or improperly disposed or dispersed toxic chemical. The final section of this article ex-

plores these important policy implications of the legal trend to require an epidemiological threshold for proof of causation.

IV. THE NORMATIVE AND POLICY IMPLICATIONS OF THE HEIGHTENED CAUSATION STANDARDS

When a judge adopts a substantive rule of tort causation law that bars plaintiffs from proceeding to try to prove individual causation unless they have a requisite degree of epidemiology, the court is making a legal policy decision to align the values of the law with the inherently conservative posture of the discipline of epidemiology towards risk. Dr. Greenlick, the court-appointed epidemiological expert in *Hall*, displayed notable perception about the different value systems of epidemiology and the law when he posited that the task facing judges and juries ought to be conceived of as more like that of the clinician than the epidemiologist. As Dr. Greenlick explained, epidemiology proceeds cautiously and conservatively, loathe to give up the assumption that there is no relationship between a particular agent and disease.

The process of epidemiological inquiry can be characterized by the formation of a suspicion of causality, followed by slow and careful, sometimes agonizing, movement along a continuum from a zero level of certainty about causality toward, but rarely to, a level of absolute certainty. The language that epidemiologists and other scientist use to describe their level of certainty is a very conservative language, one that generally understates their degree of certainty about an association between a factor and a disease and about a causal link between these factors.⁹²

Indeed, epidemiology is so inherently conservative in its reluctance to abandon the null hypothesis that it is far more willing to tolerate false negatives—the rejection of a causal association when one may actually exist—than false positives—the attribution of an association when one does not exist.⁹³ This inherent conservatism about accepting the possibility of a causal association is reflected in the high relative risk threshold preferred by epidemiologists, and in the discipline's preference that an association should be replicated in several studies before embracing a causal ascription.⁹⁴ In terms relevant to the products liability arena, the discipline of epidemiology posits that it is better to be

92. *Hall*, 947 F. Supp. at 1449 app. B.

93. See Farrell, *supra* note 3, at 2210.

94. See Thompson, *supra* note 3, at 251-56 (noting that while any relative risk greater than 1.0 indicates a positive association between exposure to the product in question and the disease, epidemiologists often do not embrace a causal relationship until studies have demonstrated a very strong level of association, such as a relative risk of greater than 3.0).

wrong by saying a product is completely safe when it is actually producing some harm, than it is to erroneously say it is risky. These rigorous statistical significance standards of certainty about an effect and willingness to accept false positives are the normative values of epidemiology "because nothing less than the establishment or confirmation of a predictive rule of nature rests on these test results."⁹⁵

By being reluctant to abandon the null hypothesis, and thus insisting on very strong associations before doing so, epidemiologists are quite willing to tolerate uncertainty and ambiguity. Courts, however, have to be careful not to be misled by the qualifying tenuousness and unwillingness to speak in legal cause terms that characterize epidemiologists' written or testimonial reports. As one appellate court explained, in chastising a trial judge for rejecting plaintiffs' expert causation testimony because the epidemiological studies were equivocal about causation and called for further study, epidemiology usually speaks in terms of uncertainty because it is not designed to make causal determinations, but only to identify probabilities of a risk association.⁹⁶ Moreover, medical science research articles qualify their conclusions and call for further study not as "an expression of ignorance, but rather [as] an expression that all scientific fields are open-ended and can progress from their present state."⁹⁷ Epidemiologists do not have to make decisions about who should financially bear a risk, or about how responsibility for ascertaining and reducing a risk should be allocated.

A court, on the other hand, has to make concrete decisions about what is causing a particular person's illness, a question on which statistical associational probabilities are only one form of relevant evidence, and only one relevant policy consideration. In making these individualistic causation judgments, tort law has often recognized that it is doing more than just ascertaining scientific "truth." It is making policy judgments about which party should bear the responsibility for causal uncertainty, and which party is in the best position to learn more about and absorb or spread the costs of the risks.⁹⁸ But, by in-

95. Farrell, *supra* note 3, at 2210.

96. See *Berry v. CSX Transp., Inc.*, 709 So. 2d 552, 568 (Fla. Dist. Ct. App. 1998).

97. *Id.* at 568 n.12.

98. That the law of causation has often been driven by consideration of the policy goals of tort law has long been recognized by scholars. See, e.g., Guido Calabresi, *Concerning Cause and the Law of Torts: An Essay for Harry Kalven*, 43 U. CHI. L. REV. 69 (1975); Wex S. Malone, *Ruminations in Cause in Fact*, 9 STAN. L. REV. 60 (1956). Courts have frequently crafted principles of causation, including relaxing the traditional requirements for plaintiffs, in order to achieve or balance tort goals such as compensation of innocent plaintiffs, deterrence, risk spreading, or fairness. See *Sindell v. Abbott Labs.*, 607 P.2d 924 (Cal. 1980); *Summers v. Tice*, 199 P.2d 1 (Cal. 1948) (shifting the burden of proof on causation to defendants where plaintiff was unable to

sisting on an attainment of a scientific certainty which is rarely likely to exist, and which cannot be supplied by epidemiology, a judge is not, like a scientist, just deferring decision until more research becomes available. Rather, a judge is selecting a specific course of action that definitively resolves important social and legal rights: this plaintiff, or this group of injured people, will receive no compensation from the manufacturer, and this product is therefore likely to continue to be marketed and labeled as "safe," even if the epidemiological studies have too quickly embraced a false negative,⁹⁹ or when they say that the increased risk factor is only 1.9, rather than 2.1, or when epidemiology shows some increase in risk and other scientific and individual diagnostic evidence bolsters the conclusion that the risk did in fact fall on the plaintiff.¹⁰⁰

The tort system should not slavishly follow the values of epidemiology because its purposes and social functions have always included a "justice" role that is broader in scope than whether scientists have arrived at a conclusion, or whatever happens to be the scientific "truth" consensus of the moment.¹⁰¹ Policy or "justice" considerations which are appropriately central concerns of tort law include fac-

prove which of two negligent wrongdoers caused his single injury); *Hymowitz v. Eli Lilly & Co.*, 539 N.E.2d 1069 (N.Y. 1989) (adopting market share proportionate liability approach when plaintiffs, through no fault of their own, could not prove which of several negligent manufacturers made the drug that caused their individual injuries); *McKellips v. Saint Francis Hosp. Inc.*, 741 P.2d 467 (Okla. 1987) (allowing plaintiffs to recover proportionate damages when misdiagnosis deprived them of chance of surviving cancer, even though plaintiffs could not prove the malpractice caused their death); *Herskovits v. Group Health Coop.*, 664 P.2d 474 (Wash. 1983) (allowing plaintiff to recover percentage of damages when doctor's delayed diagnosis reduced plaintiff's chance of recovery from terminal illness).

99. Some courts and commentators have noted that because the law makes causal decisions for a different purpose than science does, and uses different criteria for where to place the burden of persuasion, the law should not always insist on the 95% level of statistical significance (designed to screen out Type I errors) before admitting a scientific study. See *DeLuca v. Merrell Dow Pharm.s, Inc.*, 911 F.2d 941, 947 (3d Cir. 1993); Farrell, *supra* note 3, at 2210-13. For normative reasons, where the purpose of a legal decision is to award compensation for personal loss, the law should adopt a lower degree of certainty, perhaps an 80% standard of statistical significance, and thus display greater tolerance for false positives. Farrell, *supra* note 3, at 2211.

100. Many people, including scientists, clinicians, public health regulators, and especially potential consumers, would consider a relative risk of 1.9 or even 1.5 to be a serious matter. If the relative risk is 1.9, this means that for every 100 cases of the disease in the unexposed population, there are 190 cases in the exposed population—a 90% increase, albeit not the 100% increase in numbers of cases required by the relative risk standard of 2.0. Translated into what is known as the Attributable Risk Fraction, the percentage likelihood that instances of the disease among the exposed group are related to the exposure, a relative risk of 1.9 means it is 47.37% likely that any exposed persons' disease is related to the exposure.

101. For discussions of the different value systems and purposes of scientific research and legal decisionmaking, see DAVID L. FAIGMAN, *LEGAL ALCHEMY: THE USE AND MISUSE OF SCIENCE IN THE LAW* (1999); SHEILA JASANOFF, *SCIENCE AT THE BAR: LAW, SCIENCE, AND TECHNOLOGY IN AMERICA* (1995); Rochelle Cooper Dreyfuss, *Galileo's Tribute: Using Medical*

tors such as which party is better able to bear the risk of injury and scientific uncertainty—the injured plaintiff or the manufacturer, who may be able to spread the risk among all consumers. They also include the value judgment that if the uncertainty is due largely to the manufacturer’s failure to conduct sufficiently rigorous pre-marketing testing, or post-marketing testing when reports of potential problems arose, then it may be both fair to the injured person and consistent with the goal of deterrence to make the manufacturer absorb some of the cost of injuries.¹⁰² It may also include the value judgment that it is better to withhold or withdraw a product from the market—especially a cosmetic, or convenience, or non-life saving product, or a product with several equally effective and possibly safer alternatives—until it is more conclusively proven safe, than the alternative of continuing to market a product until it is conclusively proven unsafe.¹⁰³ This is the social value judgment that most consumers endorse. Indeed, consumers often believe that the regulatory system is far more rigorous and effective than it is, and thus assume that if a drug or medical device or food supplement is on the market, it must have been thoroughly tested by the manufacturer or the government and demonstrated to be safe. For this reason, when juries in products liability cases are presented with evidence of minimal testing by manufacturers, or conscious decisions not to follow up on trouble signs in animal tests or adverse post-marketing safety reports, they often find manufacturers

Evidence in Court, 95 Mich. L. Rev. 2055 (1977); Farrell, *supra* note 3; Peter Schuck, *Multiculturalism Redux: Science, Law and Politics*, 11 YALE L. & POL’Y. REV. 1 (1993).

102. Some commentators discussing the problems scientific uncertainty poses for tort causation law have proposed shifting the burden of proof on causation to manufacturers when the uncertainty is largely due to their failure adequately to test, follow-up, and sponsor research about the risks of their product. See, e.g., Margaret A. Berger, *Eliminating General Causation: Notes Towards a New Theory of Justice and Toxic Torts*, 97 COLUM. L. REV. 2117 (1997) (proposing dispensing with traditional proof of causation requirement and instead resting liability on proof that the defendant inadequately tested or disclosed information necessary to assess latent risks); Heidi Li Feldman, *Science and Uncertainty in Mass Exposure Litigation*, 74 TEX. L. REV. 1, 45 (1995) (proposing shifting the burden of proof on cause-in-fact, or allowing proportionate 50% damages recovery, when plaintiffs can demonstrate strong scientific uncertainty about causation); Wendy E. Wagner, *Choosing Ignorance in the Manufacture of Toxic Products*, 82 CORNELL L. REV. 773 (1997) (proposing the adoption of a presumption that insufficiently tested product caused plaintiffs’ harm, which the manufacturer could overcome by proving it did adequate research and that research proves the product is safe).

103. This value judgment parallels the public health protective values of the regulatory system, such as the FDA. See Kessler, *supra* note 2, at 1713. The FDA was prompted by this value calculus when it acted to withdraw the wildly popular anti-obesity drug combination “fen-phen” in advance of large scale epidemiological studies that later substantiated the risk first alerted in clusters of case reports. See Milo Garibaldi, *An American Fiasco: Diet Pills 1996-97*, 5 PHARMACEUTICAL NEWS 8-12 (1998); Press Release, Food and Drug Administration, *FDA Announces Withdrawal of Fenfluramine and Dexfenfluramine*, Sept. 15, 1997, available at <www.fda.gov/cder/news/fenphenpr81597.html>.

liable, including for punitive damages, even in the face of inconclusive or weak individual causation evidence.¹⁰⁴ While manufacturers may certainly view this tendency to determine liability despite uncertain science as alarming, it is reflective of strongly held community values that one should not market a product, especially one to be placed in the human body or ingested by people, without extensive safety testing and heightened concern about danger signs. When the tort system does not insist on well settled conclusive epidemiology as a threshold requirement for proof of causation, it allows expression of these important community values, it avoids rewarding manufacturers who deliberately choose not to do adequate safety research, and it thus encourages more socially and scientifically beneficial safety behavior by manufacturers.¹⁰⁵

Those courts that have made epidemiology the *sine qua non* of proof of causation are making a normative judgment to tolerate or even encourage a high degree of uncertainty about the dangers of some products, to reward manufacturers for ignoring risk, and to require consumers to bear more of the burden of scientific uncertainty than manufacturers. The troubling policy implications of this normative choice become even more apparent when one considers the often unappreciated synergistic effects between the tort system and scientific research. A surprisingly large number of products, drugs, and medical devices are brought to market with disturbingly little safety testing, especially for long-term effects, or for effects on population groups, such as women or children, who were not included among the clinical trial subjects.¹⁰⁶ Moreover, many drugs wind up being pre-

104. This aspect of jury decision making has been labeled "commingling" evidence of liability with evidence of causation, thus blending strong evidence on one element with weak evidence on the other, to arrive at a balancing judgment. See SANDERS, *supra* note 5, at 130-39 (analyzing this commingling by juries in Bendectin cases). For a discussion of the possibility of jurors distrusting manufacturer's uncertain evidence of safety in light of the woefully egregious state of their testing in breast implant litigation, see Rebecca S. Dresser et al., *Breast Implants Revisited: Beyond Science on Trial*, 1997 WIS. L. REV. 705, 740-43 (1997). A plaintiffs' lawyer who was very successful in obtaining plaintiff verdicts in breast implant litigation openly admitted his trial strategy was to focus on the powerful evidence of the manufacturer misconduct, and to downplay the causation issue. See Amy Singer, *Look Over Here: How John O'Quinn Directed A Jury's Attention Toward the Strongest Elements of His Case - And Won The Largest Verdict Ever In A Breast Implant Suit*, AM. LAW., March 1993, at 86.

105. See Wagner, *supra* note 102, at 828-32; Dresser et al., *supra* note 104, at 740-43.

106. Wagner extensively discusses the paucity of safety information available for many drugs or chemicals to which workers or citizens are frequently exposed. See Wagner, *supra* note 102. Even when a drug has been relatively thoroughly tested, we may have virtually no information about its safety or efficacy in women, because until recently, women were excluded from clinical trials. See R. Alta Charo, *Protecting Us To Death: Women, Pregnancy, and Clinical Research Trials*, 38 ST. LOUIS U. L.J. 135, 137-40 (1993); Vanessa Merton, *The Exclusion of Pregnant, Pregnable, and Once-Pregnable People (a.k.a. Women) from Biomedical Research*, 19 AM. J. L.

scribed for conditions or under usage circumstances other than the ones for which they were tested and approved—known as “off-label” use.¹⁰⁷ For these and other reasons, safety problems often do not become apparent until after regulatory approval, once a product or drug is on the market.

The products liability system has played a notable role in bringing some of these serious safety problems to light, and in prompting long-overdue regulatory action and scientific research. Scientific research and publication agendas are not set in a political vacuum.¹⁰⁸ A sudden onslaught of tort cases claiming a product has caused a disease can elevate an issue to the forefront of scientific interest, of the interest of research funders, including the government and manufacturers, and of the interest of the editors of scientific journals. Tort cases, particularly when they initially result in large plaintiffs’ verdicts, also

& MED. 369, 370-73 (1993). Even when a product like a contraceptive must be tested on women because it is designed for use only in women’s bodies, risks may be ignored or underestimated because women’s health complaints are trivialized or dismissed as attributable to emotional problems. See Lucinda M. Finley, *The Pharmaceutical Industry and Women’s Reproductive Health*, in CORPORATE VICTIMIZATION OF WOMEN 59-97 (Elizabeth Szockyj & James G. Fox eds., 1996); Nancy Krieger & Elizabeth Fee, *Man-Made Medicine and Women’s Health: The Biopolitics of Sex/Gender and Race/Ethnicity*, in MAN-MADE MEDICINE: WOMEN’S HEALTH, PUBLIC POLICY, AND REFORM 15-27 (Kary L. Moss ed., 1996).

107. See Steven R. Salbu, *Off-Label Use, Prescription, and Marketing of FDA-Approved Drugs: An Assessment of Legislative and Regulatory Policy*, 51 FLA. L. REV. 181 (1999); Jaime A. Wilsker, *One-Half Phen in the Morning/One Fen Before Dinner: A Proposal For FDA Regulation of Off-Label Uses of Drugs*, 6 J.L. & POL’Y 795 (1998). The FDA recently promulgated new regulations governing the off-label uses of drugs, and the information manufacturers may circulate to doctors about such uses. See Dissemination of Information on Unapproved New Uses for Marketed New Drugs, Biologics, and Devices, 63 Fed. Reg. 64,556 (Nov. 20, 1998).

108. Some courts have recognized this reality, and have taken it into consideration to reach conflicting decisions about whether a plaintiff can have sufficient proof of causation without any, or extensive epidemiology. Compare *Zuchowitz v. United States*, 140 F.3d 381, 385 (2d Cir. 1998) (finding that no studies of toxic effects of drug Danocrine at doses higher than recommended dose had been done, that no studies of intentional overdoses could ethically be performed, and number of cases of women exposed to overdoses was too small to conduct meaningful epidemiology, and taking these factors into account as one of bases for decision that plaintiff had sufficient proof linking overdose to her fatal lung disease despite lack of epidemiological proof), and *Ambrosini v. LaBarraque*, 101 F.3d 129 (D.C. Cir. 1996) (noting that it would be impossible and unethical to conduct epidemiological study to further quantify an association between the contraceptive drug Depo-Provera and birth defects because FDA had banned use of drug during pregnancy out of concern about possible birth defects, and holding that plaintiff does not need well-settled epidemiological studies linking her precise defect with Depo-Provera), with *Sanderson v. International Flavors & Fragrances*, 950 F. Supp. 981 (C.D. Cal. 1996). In *Sanderson*, experts testified that the dearth of research on the connection between fragrance products and various illnesses was attributable to the fact that the research community had no political or monetary reason to be interested in the issue. 950 F. Supp. at 994. The court held that a plaintiff has no case until the day when science decides to study the problem and reaches epidemiological threshold of extensive studies that show doubling of relative risk. *Id.* at 1003-04. The decision in *Sanderson* virtually guarantees that no such research ever will be done.

stimulate media and public awareness of potential health and safety problems, with consequent demands for firmer answers, which also pushes the research and regulatory agenda. The tort system played a highly significant, perhaps crucial role in bringing to public attention, and prompting further scientific investigation and understanding of the risks of asbestos, agent orange, thalidomide, mer/29, accutane, DES, the Dalkon shield, Bjork-Shiley heart valves, and now the fenfluramine diet drugs.¹⁰⁹ Even in the instances of Bendectin and silicone breast implants, where the science stimulated by the tort suits¹¹⁰ has turned out to show that the risk is either much less than initially feared or so small as to be unprovable by the limited detection power of epidemiology, both the public health and individuals are still better off for the research into the effects of these products that had long been used without sufficient testing. Science gained new understanding into the processes and risk factors for birth defects; cavalier attitudes about giving any drug to pregnant women were changed; less expensive and equally or more effective alternatives for treating morning sickness were promoted; doctors and potential breast implant recipients learned far more about the risks of product failure, rupture, leakage, and local complications such as infections, granulomas, and hardening around the implants; implant manufacturers were stimulated to produce better quality and safer saline implants; women were reassured about the possibility of auto-immune disorders; and concerns over the safety of the silicone implants also prompted increased biomaterials research and quests for devices that could be better biologically integrated with natural tissue.

Under the new heightened, epidemiology-deferential causation law, these often beneficial effects of the tort system in stimulating scientific research may be shut down. If tort law insists that plaintiffs cannot proceed to trial without already well-developed science, manufactur-

109. For discussions of how litigation over some of these products has stimulated renewed research interest and greater scientific, regulatory, and public understanding of the dangers, see, for example, Finley, *supra* note 106, at 59-97; Dreyfuss, *supra* note 101, at 2066-67; Wagner, *supra* note 102, at 822-25; Dresser et al., *supra* note 104, at 743-45; Paul D. Rheingold, *The MER/29 Story B An Instance of Successful Disaster Litigation*, 56 CAL. L. REV. 116 (1968).

110. SANDERS, *supra* note 5, at 61-83 (charting how the scientific research into Bendectin grew in response to the litigation, and then tailed off dramatically as the litigation came to an end). Dresser shows a timeline relating virtually all of the research into the connection between breast implants and auto-immune diseases to the onset and growth of the litigation. Dresser et al., *supra* note 104, at 743-45. Dresser and colleagues posit that but for the litigation, this product that had been used untested for over thirty years in women's bodies would have remained unstudied. *Id.* The FDA, which failed to require manufacturers to submit safety data for eight years after it first requested that they do so, also might never have acted to take a second look at implants and require the first ever controlled clinical trials, as well as more accurate reporting of rupture rates and other localized complications. *See id.* at 735-39.

ers will have a perverse incentive to avoid or suppress adequate safety research.¹¹¹ There will be fewer reasons for epidemiologists to become interested in studying products when the courthouse door is prematurely slammed. Similarly, there will be little reason for funders, especially pharmaceutical companies and product manufacturers, to support epidemiological research into the health effects of products already on the market. If the consequence of no epidemiological research being performed is to largely insulate a manufacturer from liability, then why would any reasonable manufacturer want to fund post-marketing studies when they may lead to sustaining lawsuits? Manufacturers would also have an incentive to lobby regulatory agencies and government research funding entities to keep a product off their research radar screen, or to retain the leading experts and even pay them to refrain from investigating the possibility of a causal association between a product and a disease.

The way judges have applied the new heightened causation rules and interpreted tentative or inconclusive epidemiological studies may also stifle further scientific research, because the tort system has rushed to premature "closure" on some products or scientific issues. The adversarial nature of the legal system demands "either/or" or "yes/no" answers to questions such as causation. This feature, coupled with the dramatic financial stakes hinging on the outcome of mass torts, has produced a tendency on the part of judges, prompted by their own misunderstanding of epidemiology,¹¹² or by misleading attributions of conclusiveness by defense counsel and defense expert witnesses, to make illusory ascriptions of certainty to epidemiological

111. As Wendy Wagner has pointed out, tort causation law already creates too many disincentives for manufacturers against thoroughly studying the risks of their products. See Wagner, *supra* note 102, at 774-75. See also Berger, *supra* note 102, at 2138-40; Feldman, *supra* note 102, at 44-45. Since large scale epidemiology on the effects of drug use in the general population can only ethically be done retrospectively (once a drug is already marketed), this new heightened standard of causal proof also reduces incentives against following up on post-marketing adverse incident reports or other signals of risk that come to light once large numbers of people are taking a drug outside the careful confines of the clinical trial. See Berger, *supra* note 102, at 2128.

112. Judge Weinstein's opinion in the *Agent Orange Litigation* is an example of a judge misunderstanding the language and reasoning of epidemiology to declare studies that arrived at no certain conclusion either way as conclusively negative, i.e. as positively ruling out an association between the dioxin in Agent Orange and cancers and other health problems of Vietnam veterans. See *In re Agent Orange Product Liability Litig.*, 611 F. Supp. 1223 (E.D.N.Y. 1985). Judge Weinstein's use of uncertain epidemiology has been criticized by commentators studying the *Agent Orange Litigation*. See Schuck, *supra* note 101, at 11-14; Green, *supra* note 49.

studies that by their terms say the matter is hardly closed, and call for further inquiry.¹¹³

Breast implant litigation provides an illuminating example of this tendency of the adversarial tort system to translate inconclusive epidemiology as positively ruling out any association. The two largest scale epidemiological studies of breast implants and a variety of auto-immune diseases or symptoms did show some slight increased risk of a few diseases or clusters of reported symptoms that could indicate immune system dysfunction. While the authors concluded that their research failed to demonstrate a strong or clear link between silicone implants and auto-immune disorders, they carefully noted that neither did the studies disprove or otherwise rule out an association. Like the good researchers that they were, they called for further studies.¹¹⁴ Defense experts, when asked to testify in the “yes or no—causation or not” terms of legal discourse, erroneously described these studies as conclusively establishing lack of causation.¹¹⁵ Dr. Greenlick, the court-appointed neutral expert in *Hall*, took a defense expert to task for this, and noted that his assessment of the epidemiology as showing that silicone was *not* associated with auto-immune diseases was “at odds with the way some, or perhaps most, epidemiologists would assess the literature.”¹¹⁶ “When expert witnesses testify within the culture of the adversarial judicial process, they construct their opinions on a base of science, but they easily depart from the norms of science in the language they use to offer their opinions,” Dr. Greenlick explained.¹¹⁷ “While they are cautious and conservative as they state opinions in formal scientific discourse, in the courtroom they describe their views in far stronger and more certain language.”¹¹⁸ Thus were studies that said the issue of an association between silicone and various auto-immune syndromes was still open, with sufficient signs of risk worth taking seriously and requiring further study, translated into studies that “proved” there was no connection.

The court rulings saying plaintiffs did not have proof of causation were often couched in similarly conclusive terms about the state of the

113. See SANDERS, *supra* note 5, at 110-15, 195-97 (discussing how the adversarial process of litigation leads to an unhealthy deconstruction of science, overemphasizing the flaws in studies, and ascribing greater negative certitude or positive strength than is scientifically warranted for inevitably inconclusive studies).

114. See Sherine E. Gabriel et al., *Risk of Connective-Tissue Diseases and Other Disorders After Breast Implantation*, 330 NEW ENG. J. MED. 1697, 1701-02 (1994); Hennekens et al., *supra* note 81, at 681.

115. See, e.g., *Hall v. Baxter Healthcare Corp.*, 947 F. Supp. 1387 (D. Or. 1996).

116. *Id.* at 1449.

117. *Id.* at 1448.

118. *Id.*

science: the legal system has now largely declared silicone gel breast implants conclusively "safe," even though no scientific study does so. Through this process, the new causation law has led to illusory certainty and premature closure on scientific and public health issues that warrant further investigation.¹¹⁹ But once a court has declared a scientific case closed, what incentive is there to keep performing time consuming and expensive long-term, large-scale epidemiological studies? Manufacturers will want to embrace the legal declaration of closure; journal editors and peer reviewers are likely either to consider continued research not sufficiently important or original to deserve publication, or to cast aspersions on the researcher for defying conventional scientific wisdom. Unless there is an organized and politically powerful victim's advocacy group that continues to demand research, as there was with veteran's groups demanding further study of the health effects of Agent Orange,¹²⁰ or the women's health advocacy group DES Action, which worked with researchers and legislators for increased funding for research and education about the risks,¹²¹ the legal system's premature declarations of scientific closure are likely to have a powerful impact on the scientific and regulatory communities.

There are also likely to be subtle, but significant gender, race, and class implications of the heightened requirements for proof of causation. Certain social groups have traditionally drawn greater research interest and research dollars. For example, medical problems of middle-aged white men have received a disproportionate amount of research attention, while the problems of women, the poor, and members of minority racial and ethnic groups have received less attention.¹²² Political groups or other organizations can also stimulate research attention to potential health problems—for example, the unions have played an important role in pushing for research into asbes-

119. See Julie M. Spanbauer, *Breast Implants as Beauty Ritual: Woman's Sceptre and Prison*, 9 YALE J.L. & FEMINISM 157 (1997) (exploring how the legal rush to judgment has led to scientific and media conclusions that implants are perfectly safe, falsely reassuring women and submerging awareness and concern about well documented localized adverse effects—as distinct from systemic diseases—of implants).

120. See Schuck, *supra* note 101, at 11-12 (discussing the continued pressure by veterans groups on the Veterans Administration to sponsor further research and address health concerns about Agent Orange).

121. See Finley, *supra* note 106, at 69.

122. See *id.* See also Lisa C. Ikemoto, *The Fuzzy Logic of Race and Gender in the Mismeasure of Asian American Women's Health Needs*, 65 U. CIN. L. REV. 799 (1997); Paul Cotton, *Examples Abound of Gaps in Medical Knowledge Because of Groups Excluded From Scientific Study*, 263 JAMA 1051 (1990); Paul Cotton, *Is There Still Too Much Extrapolation From Data on Middle-aged White Men?*, 263 JAMA 1049 (1990).

tos and other occupational exposures; veterans groups have continued to demand better Agent Orange and Gulf War Syndrome research.¹²³ Until the relatively recent attention devoted by civil rights groups to toxic exposure issues in poor minority communities,¹²⁴ and the activism of women's groups around breast cancer research or DES research, the advocacy groups that prodded regulatory agencies or the research community were more likely to represent largely male constituencies, such as industrial workers. For these reasons, the products, exposures, and diseases for which there is likely to be well-developed epidemiologic research will not be gender, race, or class-neutral. The societal groups most likely to be under-studied by the research community are often going to be the same groups whose health concerns have received less initial scrutiny from product or drug manufacturers.¹²⁵ If the epidemiologic community has not produced enough research into some types of the exposures and risk factors facing women, minorities, and the poor, then these groups will be inherently disadvantaged when they try to use the tort system to redress their health problems and to stimulate more serious manufacturer and researcher attention to their concerns. The long-standing inequities of medical research can lead to differential race, gender, and class-based access to the tort system.

It is not coincidental that the two products that have led courts to devise the new heightened proof of causation rules were products used largely by women—Bendectin, and breast implants. In both cases, these products were over-hyped and over-used, based on very limited or nonexistent initial manufacturer sponsored research about safety, which is consistent with the unfortunate history of insufficient research concern devoted to women's health issues. Their use by so many women stemmed from culturally induced fears or pressures surrounding the quintessentially female concerns of pregnancy and breast endowment.¹²⁶ In light of societal health priorities, and long-standing attitudes by manufacturers and in the medical research community towards women's problems, it is unlikely that either of these products ever would have received much scientific attention without the intervention of the tort system. While the increased research attention may have alleviated some of the initial health concerns about these

123. See Dreyfuss, *supra* note 101, at 2066-67.

124. See ROBERT BULLARD, *UNEQUAL PROTECTION ENVIRONMENTAL JUSTICE AND COMMUNITIES OF COLOR* (1994).

125. See, e.g., Finley, *supra* note 106 (discussing how manufacturers of contraceptive devices routinely ignored or dismissed women's reports of problems).

126. See, e.g., SUSAN M. ZIMMERMANN, *SILICONE SURVIVORS: WOMEN'S EXPERIENCES WITH BREAST IMPLANTS* (1998); Dresser et al., *supra* note 104; Spanbauer, *supra* note 119.

products, it has shed important new health information on the localized dangers and rupture or degradation rates of implants. Unless courts resist the scientifically and legally problematic emerging trend to require well-developed epidemiology in order for injured plaintiffs to enter the courthouse gate, we may rarely have future opportunities to learn more about new understudied but potentially ineffective or risky products that are now being marketed, or will in the future be marketed and used largely by, or have impacts primarily on the poor, minorities, or women.

In order to alleviate these myriad policy pitfalls and health research disincentives of the emerging causation law, courts faced with making a *Daubert* admissibility determination about a plaintiff's proffered causation evidence must not place all the burden of scientific uncertainty on plaintiffs, especially when that uncertainty is due largely to manufacturers' derelictions in adequately testing their products. While courts do not yet need go so far as overtly shifting the burden of proof on causation to manufacturers, as some have proposed,¹²⁷ they can avoid the scientifically and statistically unsound step of elevating epidemiology into the sole way to meet the burden of proof by permitting plaintiffs to go forward based on the cumulative weight of all available forms of scientific evidence and medical judgment, so long as they are all types of methodologies reasonably relied on by experts in the medical scientific disciplines. Courts must remain open to the scientific reality that epidemiology with a relative risk of less than 2.0 does not rule out individual causation, and is not the only relevant scientific or medical evidence for proving it. Courts must also carefully distinguish between the decision to admit evidence and the evaluation of that evidence for its sufficiency, or the strength and conclusiveness of the science. If lay judges routinely start substituting their judgments about the conclusions that may properly be based on scientific research for the judgments of scientists, that is an overreaction to concerns about litigation-driven pseudo-science corrupting the trial process. Judges are too likely to substitute simplistic, illusory formulations, such as the equation of epidemiology, with legal relevance and the burden of proof for the careful, scientifically sensitive judgment process that *Daubert*, sound deterrence policy, and considerations of equitable access to the courts and the science stimulating effects of tort law demand. And while much of this new judicial causation law may reflect concerns about jurors' abilities to assess complex and disputed scientific evidence, jurors' judgments are often

127. See sources cited *supra* note 102.

reflective of legal and community values of equal or greater importance to the value of the “right” or “certain” scientific conclusion, values such as the need to send deterrence signals to manufacturers who have wilfully ignored their obligation to thoroughly assess the human health risks of their drugs and products.