Toxic Torts
Science, Law, and the Possibility of Justice

Carl F. Cranor
University of California, Riverside

CAMBRIDGE UNIVERSITY PRESS
decisions by the scientific community. However, these developments have had some adverse consequences as well. To the extent that courts do not review the science well, this will undermine the aims of the decisions. It also appears that there will remain too little institutional concern with the uncertainties concerning the safety of products before they enter commerce and following them once they are in commerce. Even worse, under the procedures of Daubert there are temptations for litigants defending their actions to utilize uncertainty to slow regulation or frustrate tort suits, and to construct a view of science for courts that is at odds with most working scientists, creating an “insidious science” with a patina of respectability to it. Similar incentives can motivate litigants to distort scientific studies and the literature for legal purposes. These consequences can corrupt the science and our (and judges’) view of it.

Decisions about the science needed to assess risks and harms in the tort law can adversely affect institutional structures that already appear not to work well and to fall well short of reasonable goals. Chapter 8 argues that the Daubert trilogy and its implementation create structural incentives to decrease testing for risks from products. It creates motivations to reduce scientific and institutional understanding of the potential toxicity of products in our midst. Apart from a major overhaul of our legal structure on the regulatory side (for which there likely is little political will), courts could take some steps toward addressing these issues by modifying their reviews of expert testimony within a law that has a causation requirement. Or, they could modify tort law liability rules themselves in order to provide greater incentives to test products for adverse consequences and to monitor products once they are in commerce. Will we as a society go beyond the current half measures that have several counterproductive consequences? Will we stay with the status quo, in which there is too little legal concern with uncertainties concerning the safety of products before they enter commerce, creating incentives that increase risks to the safety of the public, the workforce, and the environment? This remains to be seen.

2
Legal Background

INTRODUCTION

The law is one of the complex institutions that must be understood in order to identify the science-law issues, to see why they are so critical to the functioning of the legal system and to understand why mistaken decisions about the admission of expert testimony can be of wider social concern. Moreover, for both historical and ongoing disputes, it is important to understand why some of the legal changes have occurred.

This chapter first provides some institutional background about the tort law, including some specific steps in civil procedure, in order to identify the stage at which courts consider the admissibility of evidence. This reveals why admissibility decisions at this point in the timeline leading to a trial can be so crucial to the litigants (mainly the plaintiffs), to the law, and to society more generally. Second, it sketches the context in which the U.S. Supreme Court decided to take the legal admissibility of scientific evidence. Finally, it considers three recent U.S. Supreme Court cases and how these have modified the admissibility of expert testimony and its scientific basis, and some recent amendments to the Federal Rules of Evidence subsequent to the Court cases.

THE TORT LAW

The legal actions that are of concern arise in the tort or personal injury law. Tort law is that

body of law which is directed toward the compensation of individuals, rather than the public, for losses which they have suffered within the scope of their legally recognized interests generally, rather than one interest only [such as contracts], where the law considers that compensation is required.1

Tort law is often contrasted with criminal law, which, as a common account describes it, is typically "concerned with the protection of interests common to the public at large, as they are represented by the entity which we call the state; often it accomplishes its ends by exacting a penalty from the wrongdoer." It is also contrasted with contract law, which imposes liability "for the protection of a single, limited interest, that of having the promises of others performed," and with quasi-contractual liability that has been "created for the prevention of unjust enrichment of one person at the expense of another, and the restitution of benefits which in good conscience belong to the plaintiff." The tort law is concerned with compensation for injuries a person has suffered that were intentionally or negligently inflicted by others, or inflicted "without fault" for which a person can recover under strict liability laws.

The conception of justice on which the tort law rests is historically traceable to Aristotle's principle of rectificatory justice: it seeks to rectify wrongs – make matters right – that have been done to persons and to restore them to the condition they were in before the injury occurred. The tort law is one part of a legal system concerned with citizens' failure to comply with the law. In a perfect world there might be little or no use for the tort law – if citizens conformed to the law, were extremely careful about how their activities affected others, and voluntarily and immediately compensated those harmed by any inevitable accidents. Alas, we do not live in such a world. Not all carefully conform to the law; not all are as careful as they should be; injuries inevitably occur (and even in a perfect world, it might be difficult to avoid them). In addition, of course, when people should be compensated is a contentious issue usually requiring an authoritative body to adjudicate the issues, to determine when compensation is owed and to ensure that it is paid. Consequently, there is a need for an institution that permits citizens injured by others to bring a legal action to rectify those harms and to compel compensation.

Moreover, as I noted in the first chapter, if other legal institutions functioned impeccably to identify and remove risks from chemical products before they materialized into harm, there would be a lesser need for torts. However, neither premarket laws, such as drug or pesticide laws, nor postmarket laws that seek to reduce risks of products in commerce before they materialize into harm appear to work well (as I consider later).

The tort law is privately enforced by those who believe they have been injured by others (contrasted with the criminal law that is publicly enforced by an agency of the state). Plaintiffs who have been exposed to substances that they believe harmed them may file suits seeking compensation for the injuries suffered in order to restore them to the status quo ante.

I have introduced this generic conception of the tort law to place it within the legal landscape. However, the specific issues of concern in this book and their importance are revealed by locating them within legal procedures.

A LEGAL CASE IN OUTLINE

How does a person bring a legal action in torts? Suppose that as in the case of Walter Allen, his wife believes that he was exposed to a toxic substance (which he was) and that as a result of that exposure he contracted brain cancer and died. What are some of the legal steps she or her legal representative must take in order to have a jury trial on the issue and, ultimately, to recover damages for injuries she might have suffered?

Complaint and Answer

Mrs. Allen would typically secure a lawyer, preferably quite a good one. The lawyer would determine whether she has a plausible theory of liability for Walter Allen's injuries, and in what legal jurisdiction her case should be heard. Her lawyer would then draft and file a "complaint" with the court in question. This complaint would "allege facts showing the defendant's primary duty toward him, defendant's breach of that duty, and (in many cases but not all) actual injury and a causal relation between defendant's breach of duty and that injury" for which plaintiff is entitled to legal relief. Because the Allen case involved the product ethylene oxide and was subject to product liability law, she has only to show that the product injured her husband, that she suffered legally compensable injuries and that she should be compensated for the injuries.

The party(ies) named as defendant(s), in this case, American Sterilizer and Pennsylvania Engineering, would have an opportunity to respond by means of an "answer." Defendant's answer would normally deny the factual allegations on which the case rested and might even deny that there was a substantive theory of liability under which plaintiff could receive legal relief, even if the facts were as she alleged. That is, the defendants might deny that plaintiff was exposed to the substances in question, and might go on to claim that the substances

1 Prosser and Keeton, 5.
2 Prosser and Keeton, 5.
3 Prosser and Keeton, 32.
5 Ravelo, A Theory of Justice, 246, 351 (the tort and criminal laws are part of imperfect compliance with principles of justice (351)).
7 This particular style of a complaint is chosen as fairly representative, although in an earlier time complaints had to be both more stylized and in some ways more specific. The Federal Rules of Civil Procedure only require "a short and plain statement of the claim showing that the pleader is entitled to relief" (Federal Rules of Civil Procedure 8(a) (2)).
alleged were not toxic, that even if exposure had occurred, the concentrations of the substance were not harmful, and that, even if plaintiff were exposed to the substances in concentrations which were toxic, plaintiff was not entitled to legal relief under the substantive law of the jurisdiction. Such responses are one of the rituals of the law.10

Discovery

Next, the two sides would engage in a discovery process. These are official procedures that are made available so that litigants can discover the “facts and possible evidence in the case, and at least ascertain in part what detailed fact issues may arise for trial, as well as the opponent’s positions concerning factual matters.”11 Each party may address interrogatories to the opponent, seek documents from the other side, orally examine witnesses under oath who have filed depositions, as well as examining anyone who may have some knowledge of the subject matter of the suit.12

Pretrial Conferences

Once discovery is completed, there is a pretrial conference (or conferences) to address issues between parties. Following that, the trial begins.

However, it is at this point, during the pretrial hearings, that plaintiffs face some of the issues that arise concerning the admissibility of scientific testimony and evidence. In particular, the plaintiff has the “burden of producing evidence” or the “burden of production.”

Our system [of law] leaves it to the parties to [investigate the case or furnish the evidence upon which they are to be decided]. If, now, neither party offers any evidence at the trial, what will happen? The answer is that one party loses. He may, therefore, be said to bear the risk of this consequence of nonproduction of evidence. Or as we more often say, he bears the burden of producing at least some evidence.14

10 The last part of defendants’ answer is only the first of many instances in which plaintiffs may face motions by defendants aiming to end a case. Roughly, a defendant claims that even assuming the correctness of the evidence that plaintiffs have alleged (in the complaint in this case), or discovered prior to trial or established at trial, such evidence is insufficient as a matter of law to establish a basis for legal recovery for alleged harm suffered. These are all legal devices by which one party or the other might seek to have a judge rule as a matter of law that the legal theory or the evidence (as the case may be) on the other side of the issue was so lacking that the legal process should be terminated. Typically, plaintiff faces these issues more than defendant because the plaintiff has the initial burden of showing that there is some reason to change the legal status quo and the burden of producing enough evidence to justify the legal process continuing.


12 McCormick on Evidence, 3.

13 James and Hazard, Civil Procedure, 245.

14 James and Hazard, Civil Procedure, 245.

In a toxic tort suit, the plaintiff must introduce sufficient evidence that she was exposed to a potentially toxic substance and that it more likely than not caused her injuries (along with other evidence about defendant’s responsibility, and so on). Typically, because of the scientific nature of the required evidence, plaintiffs must rely upon both scientific studies and the testimony of expert witnesses. It is these that have become the focus of debates in light of the Supreme Court’s decision in Daubert.15

Before or during pretrial conferences, plaintiffs may face a variety of motions petitioning the judge legally to review experts to evaluate whether they should be “admitted” to testify in the case and these are likely to be accompanied by a motion to dismiss the case. In particular defendants may file a variety of motions concerning expert testimony plus a motion for summary judgment. Rule 26 of the Federal Rules of Civil Procedure “requires the disclosure of not only the full opinions to be offered by certain experts, but also the bases for the opinions.”16 At this point, an opponent may challenge the admissibility of one or more expert witnesses. Such challenges may take the form of a motion to strike the evidence or a motion for an in limine hearing (a motion to have evidence or expert testimony reviewed by the judge in a specific hearing before trial).17 That is, a court can issue a decision based on papers the litigants have submitted or hold a hearing where “issues in the case can be more fully explored.”18 Some courts have required hearings in certain instances, but in general there appears to be considerable flexibility on how expert testimony is reviewed.19 Although the judge has authority to “require parties to present objections to expert testimony at any point during the case,” typically such objections would be made “shortly after the close of expert discovery” or “shortly before trial.”20

At the time experts are challenged, opponents may file a motion for a summary judgment. This would allege that based on plaintiffs’ “pleadings, depositions, answers to interrogatories, and admissions on file, together with the affidavits [by experts and others], if any . . . ” plaintiffs’ case lacked an adequate evidentiary basis to establish all the elements needed for their theory of liability.21 If critical experts have not been admitted to testify, opponents may succeed in ending the case at this point.

Once plaintiff and defendants have passed any pretrial motions or hurdles they might face, including the adequacy of the scientific basis of their claims, the case proceeds to trial.


20 Federal Rules of Evidence 506(c).
Plaintiff's Case-in-Chief

Plaintiff presents her case first, because she has a general burden to show that there is some reason to change the legal status quo and because she has a burden to produce evidence, the burden of production. That is, plaintiff must present legal theories and sufficient evidence to justify "a finding in [her] favor..."21 Plaintiff presents her evidence and experts, subject, of course, to cross-examination. At the end of this presentation, her case-in-chief, she might once again face a motion from defendants seeking to end the legal proceedings on the grounds there is not a legally adequate evidentiary basis for a reasonable jury to find for her or that there was no legal theory under which she could be entitled to recover compensation. That is, defendants might well move for a "directed verdict."22 If this motion is denied, then the presentation shifts to the defense.

Defendant’s Case-in-Chief

Next, the defense presents its case-in-chief - legal theories and evidence to counter plaintiff’s claims on the issues in question. Following this presentation, defendants might move for a directed verdict in their favor on the grounds that plaintiff’s evidence, taking it in the sense most favorable to her, when compared with defendant’s evidence does not present a substantial issue of fact to go to a jury.23 If this fails, the sides proceed to closing arguments.

Closing Arguments and Proposed Jury Instructions

After both sides have made their concluding arguments, each submits proposed jury instructions. These are proposals about how the judge should instruct the jury on factual and legal issues. The judge reviews them, decides what instructions should be used, and then instructs the jury.

Plaintiff’s Burden of Persuasion

When the case goes to the jury, the plaintiff faces a second burden of proof, the "burden of persuasion."

If... the trier [of fact – the jury] is operating under a system which requires him to decide the question one way or the other, then to avoid caprice that system must furnish him with a rule for deciding the question when he finds his mind in this kind of doubt or equipoise. Where the parties to a civil action are in dispute over a material issue of fact, then that party who will lose if the trier’s mind is in equipoise may be said to bear the risk that the trier will not be affirmatively persuaded or the risk of nonpersuasion upon that issue.24

In less convoluted prose, if the plaintiff does not establish her factual and legal claims to the required standard of proof, plaintiff loses. Moreover, it is the members of the jury who must be persuaded: "the judge is not directly concerned in solving the problem."25

Plaintiff’s Standard of Proof

The burden of persuasion is one thing (concerning who must establish a factual or legal issue), how difficult it is to lift or carry the burden is another – that is the "standard of proof." The standard of proof is the degree of certainty to which a litigant must make her claims to the satisfaction of a fact finder (typically the jury) in order to prevail on an issue. In the criminal law, the state has the burden of proof to establish the necessary elements of a crime in order to obtain a conviction of an accused. It must establish its factual and legal claims "beyond a reasonable doubt" to the satisfaction of a jury. The standard of proof needed to establish legal claims can and do differ from one area of the law to another.

In most areas of the tort law the plaintiff must establish her case by a preponderance of evidence to the satisfaction of a jury.26 Equivalent formulations are that she must establish her claims by a "greater weight of the evidence" or that she must establish the required claim by showing that it is "more likely than not true." Carrying this burden to the required standard refers "not to the number of witnesses or quantity of evidence but to the convincing force of the evidence"27 (emphasis added).

If plaintiff has persuaded the jury that her claims are more likely than not true, the jury will find in her favor. If not, the jury will find for the defendant. Once a jury decision has been reached, the trial judge may review it and if he agrees, enters a judgment for the winning party. However, jury verdicts are subject to post-jury reviews by the judge and, of course, subject to appeal to higher courts.

SUBSTANTIVE ISSUES IN THE TORT LAW

In addition to these generic issues of civil procedure, a plaintiff must establish the substantive elements required of tort law in order to obtain legal relief for injuries suffered. In particular a plaintiff must show (1) that defendant had a

21 James and Hazard, Civil Procedure, 268.
22 Federal Rules of Evidence 50(a)(1), and James and Hazard, 280–288.
26 James and Hazard, Civil Procedure, 243.
Causation in Toxic Tort Suits

The element of concern for our discussion is that plaintiff must show by appropriate evidence that defendant caused plaintiff’s harm. In toxic tort cases in federal jurisdictions, this showing typically breaks into two distinct causation claims that must be established (although as we will see in Chapter 8, sometimes this rigid distinction can be misleading and has been rejected by some states). First, plaintiff must show general causation: that defendant’s substance can cause or is capable of causing harms of the kind from which plaintiff suffered. Second, she must show causation specific to that plaintiff: or specific causation: that defendant’s action or product caused or contributed to plaintiff’s disease. That is, if a plaintiff claims that ETO caused her husband’s injuries, she must show that it is more likely than not that ETO can cause the kinds of injuries from which he suffered and that it is more likely than not that defendant’s ETO (and not something else) caused or contributed to the injuries. Such bifurcation of causation is not needed in more ordinary tort cases, such as car accidents, because it will be obvious that a five-thousand-pound car traveling at moderate to high rates of speed can cause considerable physical damage to many objects.

Because most diseases have multiple causes, to meet the specific causation requirement litigants must rule out other significant possible contributors to plaintiff’s condition. Litigants might use epidemiological studies to establish that in a general population of people exposed to the substance suspected of causing the harm there was a sufficiently elevated rate of disease to satisfy the burdens and standards of proof on general causation. They can and do also utilize animal and other toxicological studies to provide compelling evidence that a substance can cause a certain adverse effect in humans, for example, that the anticancer drug CCNU (1-(2-chloroethyl)-3-cyclohexyl-1-nitrosourea) can cause cancer in humans. In order to establish that plaintiff’s exposure in question likely caused or contributed to her disease, plaintiff’s expert must infer this conclusion from the general studies and from more particular information about plaintiff’s exposure and circumstances.

Plaintiff must establish both general and specific causation by a preponderance of the evidence; she must show that it is more likely than not that defendant’s substance is capable of causing injuries of the kind from which plaintiff suffered and that it is more likely than not that defendant’s substance caused her specific injuries. Thus, to return to our example, Mrs. Allen must show that it is more likely than not that exposure to ETO is capable of causing brain cancer and that it is more likely than not that ETO (rather than something else) contributed to her husband’s brain cancer.

The Role of Scientific Evidence and Expert Witnesses in Establishing Causation

In order for the plaintiff to establish both general and specific causation required in toxic tort cases, she must present evidence admissible in court that is sufficient to persuade a jury that her injuries more likely than not were caused by exposure to defendant’s substance. In toxic tort cases, this evidence is often established in large part by introducing expert testimony based on appropriate scientific studies. Scientific evidence is typically provided by any studies that have been done concerning the toxicity of substances to which plaintiff was exposed (although it is usually difficult to provide this data) or by the exposure of others to the substance or by other kinds of toxicologic evidence, in order to show that the substance more likely than not can cause such injuries like those the plaintiff suffered. For specific causation, plaintiffs must show that there was sufficient exposure and other evidence to support the claim that defendant’s substance more likely than not did cause plaintiff’s injuries.

To understand this idea and some of the evidentiary requirements for establishing scientific claims in toxic torts, we need to place the use of scientific evidence in the tort law within a broader framework. First, as some of the treatises put it, the common law insists “upon the most reliable sources of information.” For example, witnesses to an event such as an automobile accident that causes injury may testify to that fact, provided that they have firsthand knowledge of the event or transaction at issue. However, even for witnesses to testify to events that can be perceived by the senses, they “must have had an opportunity to observe, and must have actually observed the fact.” By contrast, the expert has something different to contribute. This is the power to draw inferences from the facts, which a jury would not be competent to draw. To warrant the use of expert testimony two general elements are required. First, some courts state that the subject of inference must be so distinctively related to some science, profession, business or occupation as to be beyond the ken of laymen. . . . Second, the witnesses must have sufficient skill, knowledge, or experience in or related to the pertinent field or calling as to make it appear that his opinion or inference will probably aid the trier in the search for the truth.
That is, a person who seeks to testify as an expert must first be admitted as appropriately qualified and then may testify to the science or technical matter at issue.  Thus, scientific expert opinion may be used in support of general or specific causation claims provided the expert is appropriately qualified, the expert's testimony has an appropriate scientific basis, the testimony is appropriately relevant to the factual issues in dispute, and it will assist the jury in coming to conclusions. In a toxic tort case the plaintiff's scientific witnesses must present opinions regarding a causal relationship between exposure to the toxic product and the injury that are appropriately grounded in science.

The Admissibility of Evidence
Courts consider several issues in assessing expert testimony. The first is whether any proffered evidence, including expert testimony, is admissible. Trials proceed according to rules and procedures "that make it clear when proof has been presented so that evidence is officially introduced and the jury may be considered" by the jury. Thus, such evidence must be officially "admitted" to be considered by a jury. Moreover, an expert cannot testify regarding any matter he or she chooses. An expert's credentials would typically be reviewed to ensure that he or she has the proper qualifications, experience and education to testify to the issues in the case. In addition, the scientific foundation of his or her testimony would be reviewed to ensure that the testimony is appropriately grounded in scientific research and methodology. Before 1993, introducing scientific evidence and having experts admitted tended not to be difficult: if a litigant had well-qualified experts whose testimony was relevant to the scientific and technological issues, would assist a jury in understanding them, and their methodology was not based on "novel" techniques or studies, judges tended to admit them and let cross-examination during trial guide the jury in deciding the weight to accord the testimony and whose experts to believe. Most courts utilized the Frye test. This was a principle for judging the underlying basis of expert testimony that had been created when a defendant in a murder case had sought to introduce a precursor of lie detector tests, a "systolic blood pressure detection test." The pertinent language was

Just when a scientific principle or discovery crosses the line between the experimental and demonstrable stages is difficult to define. Somewhere in this twilight zone the evidential force of the principle must be recognized, and while the courts will go a long way in admitting expert testimony deduced from a well-recognized scientific principle or discovery, the thing from which the deduction is made must be sufficiently established to have gained general acceptance in the particular field in which it belongs.

The Frye test required that the generic kinds of studies, tests, or techniques on which an expert might rest expert testimony must be "generally accepted in the pertinent field." The systolic blood pressure test had not been "generally accepted" at the time of Frye so the defense expert in that case was not permitted to testify. It was "easy to apply and required little scientific sophistication on the part of judges," and "if one were a card carrying member of a recognized occupation or profession, one's proffered expert testimony was admitted and the validity of the underlying knowledge was assumed [especially as along as one was not relying on a "novel" study or technique]."

The U.S. Supreme Court did not require such detailed scrutiny of scientific testimony as recently as 1983 in Barefoot v. Estelle, a capital murder case. In this case, an expert for a criminal defendant was permitted to testify, even though his own profession had taken a strong public stand against the position he argued and the Court indicated it was "unreliable." The Court held that contrary testimony and cross-examination could correct any mistaken views he might articulate. Since the 1993 Daubert and related decisions, judges have conducted much more searching reviews of expert testimony and its foundation before trials commence, which we discuss later in this chapter.

Summary Judgment
A second issue concerning a litigant's evidence is that even if a particular expert's testimony is admissible, it may not be adequate to support each element of plaintiffs' legal theory. If not, there is no legal issue for a jury to decide and the case can end right there. For example, if plaintiffs' evidence only supports a claim that defendants' substance might possibly cause the disease in question, there would be "no genuine issue as to [that fact] and the moving party would be entitled to a judgment [in its favor] as a matter of law." If there is not

---

33 At the time the Federal Rules of Evidence stated that
If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training or education, may testify thereto in the form of an opinion or otherwise. (Federal Rules of Evidence, section 702 (1988.).

37 293 F. at 1014.
38 Feigman et. al., Modern Scientific Evidence, 7–8.
41 Federal Rules of Evidence 56(c).
scientifically valid and reliable evidence to support each element of a litigant's case, a judge may find that there is no issue for a jury to consider and issue a decision for the opposing side, a so-called summary judgment.42

**Judgment as a Matter of Law**

A third major issue concerns whether, during a trial, a party to a legal dispute has presented a "sufficient evidentiary basis for a reasonable jury to find for that party on that issue..."43 This finding and accompanying judgment as a matter of law would be made after plaintiffs have presented their case-in-chief or after both sides have presented evidence in support of their respective views, but before the case goes to the jury. Such a review would typically involve a comparative assessment of a party's evidence versus the opposing party's evidence.44 If a court finds one side's evidence so overwhelms the evidence by the opposing party such that no reasonable jury could find for the opposing party, he or she may issue a decision for the first without having a jury decide the case. Or, for a judgment notwithstanding the verdict, if a jury decided, say for the plaintiff, but a judge found that no reasonable jury could come to such a conclusion, he or she can overturn the verdict.

**Some Procedural Puzzles**

Some of the distinctions indicated earlier between the admissibility of evidence, a summary judgment and a judgment as a matter of law appear to have been conflated in litigation concerning Bendectin. I need not settle these issues, but it is important to note them. Some courts ruling on the evidence in the Bendectin cases had a full trial record on which to base their assessment of evidence. Thus, they could properly engage in a comparative assessment of plaintiffs' versus defendants' evidence and issue a judgment as a matter of law that plaintiffs'

Evidence was insufficient for a jury verdict. Other courts not having a trial record, but facing experts prepared to testify that Bendectin caused birth defects, ruled the experts' testimony inadmissible and, then, because there was too little evidence on plaintiffs' side to support their cause of action issued a summary judgment. Whether they should have ruled such evidence inadmissible is controversial. Faced with pressures not to go through other trials when at the end they would in all likelihood have to rule for defendants, they sought to save time and money by declaring plaintiffs' evidence inadmissible before trial rather than later comparing plaintiffs' and defendants' evidence during or after trial. Their actions, however, tended to blur the "line between admissibility and sufficiency" (or judgment as a matter of law in above terminology) in these cases.45 In the Supreme Court Daubert decision (considered below) the Court seemed to endorse this procedure, inviting both summary judgments and judgments as a matter of law as a way of addressing inadequate scientific evidence.46

The qualification of experts, whether they can testify, and what constitutes an appropriate basis or foundation for their testimony are some of the issues that have arisen in several recent Supreme Court cases and that provide the background for the treatment of litigants in toxic tort cases. Focusing on these arcane legal and scientific issues is needed to locate the use of scientific evidence in the tort law and to address some of the tensions between science and the law.

Recent developments require judges to review proposed expert testimony very early in the legal process. The Federal Rules of Civil Procedure were amended in 1993 to

require a party, independently of any discovery request, to disclose the identity of all expert witnesses expected to testify at trial; to provide, among other things, the experts' written signed reports stating all opinions to be offered and support for opinions; and to make the expert available for deposition after the report is submitted.48

42 A motion for summary judgment is "typically supported by affidavits of witnesses who would be competent to testify at trial - affidavits containing statements of fact which would be admissible at trial if made by these witnesses. The movant's opponent then has the charge to submit counter-affidavits of similarly competent witnesses. Both sides may also use the products of discovery - depositions and interrogatories that establish uncontested facts, admissions made upon requests to admit, etc. If the opponent does not controvert the proofs offered in support of the motion, and the movant's affidavits show without contradiction facts which would entitle him to judgment as a matter of law, then summary judgment may be granted. If, on the other hand, the proofs fail to exclude all bases on which judgment might be rendered in favor of the person against whom the motion is made, summary judgment must be denied." James and Hazard, Civil Procedure, 220-221. See also, Federal Rules of Evidence 56(c).

43 Federal Rules of Evidence 50(a).

44 After only one side has presented evidence it might be so insufficient that a reasonable jury could not find for the plaintiff. After both parties have presented evidence, this finding would be explicitly comparative with the judgment that one party's evidence so overwhelms the other party's evidence that there is no factual issue for a jury to decide.


46 Margaret A. Berger, "Procedural Paradigms for Applying the Daubert Test," Minnesota Law Review, 78 (1994): 1325-1386; Sanders, "Scientific Validity, Admissibility and Mass Torts," 1434. See also Joseph Sanders, Bendectin on Trial: A Study of Mass Tort Litigation (Ann Arbor: University of Michigan Press, 2001), 155-156 (arguing that some Bendectin courts mistakenly ruled plaintiffs' evidence inadmissible and then issued a summary judgment that because plaintiffs could "present no competent evidence on the causal question" their evidence was insufficient).

47 Daubert v. Merrill Dow Pharmaceuticals, 509 U.S. 579, at 596 (citing both Federal Rules of Civil Procedure 56 and 50(a)).

These requirements are triggered by the date of trial, but the controlling Daubert decision "suggests that in civil litigation, issues concerning the admissibility or sufficiency of expert testimony should be raised before trial." Thus, in the outline of a legal case discussed earlier, judicial and litigant review of expert witnesses typically occurs after discovery but before a jury is empaneled and before a trial proper begins. This review has become very important and might be "outcome determinative" as some courts have noted. (Whether or not there is an actual hearing on expert witnesses is a separate issue, but there need not be.)

If a litigant, either plaintiff or defendant, fails to have all the necessary experts admitted to testify in a trial, his or her case may be at an end. This was the fate of Walter Allen's widow and son in Allen v. Pennsylvania Engineering and Lisa Soldo. In each instance, the plaintiffs were not permitted to present their case to a jury and in effect had no legal case because the judge decided a critical expert's testimony was not admissible and the scientific basis for causation needed for a tort cause of action could not be presented at trial. Consequently, because they could not offer evidence to establish their legal claims, there was no factual issue for the jury to decide. The judges issued summary judgments legally ending plaintiffs' cases without a public jury trial of the issues involved.

Less frequently, similar exclusions could occur on the defense side. Recently, the City of Chicago was required to compensate a man for brain-stem injuries following an encounter with the police. The city was unable to mount a defense based on an alternative theory of injury because its expert's theory was judged "too speculative" and the expert was not admitted for trial. The City was limited in the defense it could present and lost a jury verdict. Exclusion of defendants' experts might arise once plaintiffs have met their burden of production or perhaps when defendants are put in the position of testifying against well-established research, for example, concerning the effects of asbestos or tobacco smoke.

To summarize: Decisions about the admissibility of experts occurs early in the sequence of legal events leading to a trial before the trial proper in front of a jury ever begins. If a litigant's experts are not admitted to testify in a trial, that party's case may end at that point, unless there is some other way to establish the claims (there typically is not in toxic tort cases).

RECENT DEVELOPMENTS IN THE ADMISSIBILITY OF EXPERT TESTIMONY

The Bendectin Litigation and Related Cases

The legal issues that have modified the law and exacerbated problems for plaintiffs arose with an antinausea drug for pregnant women with the trade name, "Bendectin." Specifically, the social-legal events leading to this litigation began with David Mekdec, who was born in 1975 with a shortened right forearm, missing two fingers from his malformed right hand with limited use of the other two fingers. He was missing pectoral muscles in his chest, limiting his ability to move his right arm. He was anticipated later in life to have a "congenital heart defect." Over time his distressed mother, reflecting on drugs she had taken during her pregnancy with David, inquired about the etiological role those drugs might have had in his birth defects. The family brought suit against Merrell National Laboratories, the manufacturer of Bendectin, that she believed (and there was some evidence suggesting that it) could have contributed to David's malformed limbs. "The day before David's fifth birthday, ... a jury in a federal court in Orlando [Florida] awarded Michael and Betty Mekdec [David's parents] $20,000 in their lawsuit," an amount hopelessly inadequate to compensate plaintiffs or even to pay their lawyers. This verdict was unsatisfactory in a number of ways, was overturned by the trial judge, and the case was unsuccessful on retrial. The Mekdecis not only received nothing, but Merrell, playing hardball, requested over $200,000 in legal expenses from them. The trial judge reduced this to $6,000, which became a lien on their middle-class house (which they were not required to pay).

The personal and legal saga that began with the Mekdecis directly affected not only the thousands of people affected by Bendectin (approximately 2000 cases were filed), but the consequences of that litigation are still reverberating through the legal system today as a result of several U.S. Supreme Court decisions (as well as appellate and district court decisions implementing the Supreme Court's decisions) that arose directly or indirectly from the Bendectin litigation.

Concerns about the Companies

Bendectin had been created, manufactured, and distributed in the United States beginning in the late 1950s by Richardson-Merrell Pharmaceutical. This was

53 Federal Rules of Evidence 56(c).
55 Much of the following discussion concerning the context of this litigation is based upon Green, Bendectin and Birth Defects, 1996).
56 Green, Bendectin, 1.
57 Green, Bendectin, 3, 121–158.
58 Green, Bendectin, 3, 121–158.
was the same company that had been licensed in 1956 to distribute thalidomide, the drug that ultimately caused eight to twelve thousand birth defects in Europe and about forty in the United States. In addition, about the same time Merrell had also developed Mer/29, one of the early anticholesterol drugs. Mer/29 caused adverse side effects, such as cataracts, lost or thinned hair, and mild to severe skin reactions. Data presumably showing the safety of Mer/29 were "faked." Litigation in the U.S. concerning thalidomide and Mer/29 cost Merrell about $100,000,000 in tort damages and settlements. Moreover, according to the judge in the Mer/29 litigation Merrell acted in "reckless disregard of the possibility it would visit serious injury upon persons using it," falsified test data, withheld data from the FDA, misrepresented the safety of the drug to the medical profession, all of which could lead a jury to conclude that Richardson-Merrell acted "with wanton disregard for the safety of all who might use the drug." This background created an atmosphere of suspicion within the FDA and was well known to the plaintiffs' bar, both of which added to the firm's problems when there were concerns that Bendectin could cause birth defects.

Moreover, the discovery process in the early Bendectin litigation revealed that animal studies conducted by Merrell suggesting that Bendectin was a teratogen were not followed up, and Merrell delayed sending the results to the FDA for three years. There had also been extensive behind-the-scenes efforts to manage, even manipulate, the reporting of birth defects in patients whose mothers had taken Bendectin and any potentially adverse news stories that threatened to break about the drug and its effects.

Other firms probably added to this skepticism. For example, Michael Green notes that so much evidence appeared in the Dalkon Shield cases of "corporate wrongdoing and fraud" that it led to routine "multi-million dollar punitive damages."

Perception of a Tort Law Crisis
If the FDA and others were suspicious of Richardson-Merrell Pharmaceutical, some were also concerned about the tort law — the institution that in principle provides redress for wrongly inflicted injuries. The Bendectin litigation occurred during a period in which there was a widespread perception and discussion of a crisis in the tort law. Critics of the tort law pointed to an earlier expansion of liability for defendants, alleged over litigiousness by Americans, increasing damage awards, an increase in punitive damages, and an uncertainty about where it all would end. Also, there was concern that expert witnesses could be found to opine on nearly any issue with the result that judges and juries could be misled and that the system was "biased in favor of plaintiffs, whom sympathetic juries favor." According to this view, the result was overdeterrence — useful technologies would be driven from the market and U.S. industry rendered less competitive in international markets, as it has been alleged that physicians have been driven from medical practice by increased malpractice insurance premiums caused by malpractice suits. Indeed, it was claimed that the tort law could be a substantial drag on the economy. For some the Bendectin litigation became almost an exemplar of the ills of the tort law. Whether such claims were true is another issue to which we briefly return at the end of the chapter.

There also may be a story yet to be fully investigated about how the perception of a tort law crisis was created or arose. There has been the suggestion that a concerted public relations campaign on the part of the National Chamber of Commerce, major firms, their industry groups, supporters, and politically associated think tanks created or significantly contributed to the perception.

The Supreme Court Daubert Litigation
Daubert v. Merrell Dow Pharmaceuticals, Inc.: The specific events leading to the Supreme Court decision began when Jason Daubert and Eric Schuller were born with serious birth defects. Their mothers and others suspected that Bendectin was a contributor to their injuries.

Their cases came fourteen years into this litigation and after a number of epidemiological studies had finally been done on the relationship between

---

60 Green, Bendectin, 19.
61 Green, Bendectin, 89.
62 Green, Bendectin, 88 (quoting the federal judge in Toole v. Richardson-Merrell, Inc.).
63 Green, Bendectin, 129.
64 Green, Bendectin, 15.
65 Green, Bendectin, 19.
66 Green, Bendectin, 20.
67 See, for example, the National Chamber Litigation Center Web page that credits a mid-1970s memorandum from F. Lewis Powell, later a Supreme Court Justice, with triggering a concerted Chamber of Commerce effort to litigate cases favorable to business interests and in general to create an "advocacy program [that] has grown to include all aspects of employment relations, environmental regulation and enforcement, government contracts, as well as other cutting-edge legal issues in the areas of class action reform, product liability, torts, and punitive damages." (Located at: http://www.uschamber.com/nclc/about/anniversary.html. [visited September 2004].) There is considerable anecdotal evidence about many additional activities, but no extensive scholarly work on the subject to my knowledge.

More recently, the Republican Contract with America advocates "commonsense legal reforms" that seek to restrain plaintiffs' attorneys and recovery for certain kinds of damages. (Located at: http://www.house.gov/house/Contract/CONTRACT.html [visited September 2004].)

Finally, "tort reform" is the object of a number of organizations that seek to protect their economic interests by reducing lawsuits. See, for example, John Micklethwait and Adrian Wooldridge, The Right Nation: Conservative Power In America (New York: The Penguin Press, 2004), esp. 130, 158, 176.
69 509 U.S. at 582.
exposure to Bendectin and birth defects. Merrell Dow's expert submitted an affidavit that stated that no published study had found Bendectin to be a human teratogen and that therefore claimed that use of Bendectin during the first trimester of pregnancy had not been shown to increase the risk of birth defects. Plaintiffs' experts concluded that Bendectin could cause birth defects, basing their conclusion on: (1) test tube and animal studies linking Bendectin and malformations; (2) studies showing similarities between the molecular structure of Bendectin and other teratogens; and (3) a reaplanalysisis of published epidemiological studies.

The trial court, likely confusing the admissibility of evidence with its legal sufficiency, agreed with defendants and excluded plaintiffs' experts. It then granted a summary judgment for defendants before trial. The court held that, because there was a plethora of epidemiological evidence regarding Bendectin, plaintiffs' substantial nonepidemiological evidence was not sufficient to create a material issue of fact and defeat the summary judgment motion. The trial court relied on the Frye general acceptance test from a 1923 criminal case.

The Frye rule for the admissibility of scientific evidence was formulated in a criminal case in which a defendant tried to introduce a precursor of lie detector tests. That court held that novel scientific evidence or methodology on which an expert relied to testify had to have "general acceptance" in the relevant scientific community to be admitted for consideration at criminal trial. The original test applied to generic "tests," studies or technological devices, not to opinion testimony (although this appeared to change over time in some jurisdictions, but not others, such as California).

The Ninth Circuit Court of Appeals affirmed the trial court's exclusion of evidence, also following the Frye general acceptance test. The court apparently gave great weight to the fact that other appellate courts had not admitted reanalyses of epidemiological studies regarding the teratogenicity of Bendectin that had never been published or peer-reviewed. Furthermore, it noted that the large number of published studies opposing plaintiffs' position that Bendectin could cause birth defects undermined the efficacy of re-analyses that reached the opposite conclusion. The plaintiffs petitioned the U.S. Supreme Court to hear the case and it did so.

The Supreme Court, seemingly granting a victory to plaintiffs, vacated the Ninth Circuit's decision because it had been based on the Fryer rule, not the Federal Rules of Evidence that had been legislated in full knowledge of the existence of the Frye rule. It then remanded the case for reconsideration under its newly articulated standard for admissibility of scientific opinion evidence.

On remand, the Ninth Circuit Court of Appeals in a controversial opinion decided against the Dauberts without returning the case to the district court of origin. It reasoned that there was no reason to return it to the trial court, "If as a matter of law, the proffered evidence would have to be excluded at trial. The expert testimony would have to be excluded, the court concluded, because plaintiffs could not show that exposure to Bendectin more likely than not doubled their risk of birth defects according to the court. This last point is scientifically and legally controversial. We return to both points in Chapter 6.

The Frye "general acceptance" test had posed several concerns. Some argued that Frye was too conservative, "for it imposes a protracted waiting period that valid scientific evidence and techniques must endure before gaining legal acceptance. On this view it would keep perfectly good, but not yet broadly accepted cutting-edge science (e.g., DNA analysis) or testimony from the courtroom. It could also be seen as vague about what constituted the "general acceptance" of a particular kind of study or test. And, it may be difficult to determine in which particular field a kind of scientific study should be generally accepted. At the same time, it also has been criticized for being quite liberal: "the more narrowly a court defines the pertinent field, the more agreement it is likely to find. Perhaps in the extreme, any expert might conceivably be permitted to testify as long as there was some appropriate self-vouching community of experts who would support the principles and methods of expert testimony, for example, in astrology or necromancy.

Ultimately, on statutory grounds the Supreme Court held that the Congressional adoption of the Federal Rules of Evidence had superseded the Frye

---

70 See 509 U.S. at 582.
76 See 509 U.S. at 583.
77 See 951 F.2d at 1129–30.
78 See 509 U.S. at 582.
79 See 509 U.S. at 583.
80 See 951 F.2d at 1130.
81 See Daubert, 509 U.S. at 597–98.
85 Faigman et al., Modern Scientific Evidence, 8.
87 Faigman et al., Modern Scientific Evidence, 9.
88 Faigman et al., Modern Scientific Evidence, 9.
89 Faigman et al., Modern Scientific Evidence, 9.
rule. In developing a view that supported greater court review of expert testimony, the Court used a policy argument. It noted that the law typically requires firsthand knowledge of the facts as evidence on legal issues. Expert testimony is an exception to this general requirement, because an expert does not necessarily have firsthand knowledge of material to which he or she might testify. However, the relaxation of the firsthand knowledge requirement with its "insistence upon the most reliable sources of information..." is premised on an assumption that the expert's opinion will have a reliable basis in the knowledge and experience of his discipline. Consequently, expert opinion must satisfy some indicators of reliability.

Another way to think about these issues is that the Daubert decision sought to ensure that expert testimony is based on appropriate science pertinent to a legal decision. Thus its aim might be seen as winnowing expert testimony so that a jury decision, which is ultimately based in part on either plaintiffs' or defendants' experts' accounts of the science, will be within the bounds of respectable scientific views about the issue involved or at least not beyond the boundaries where reasonable scientific experts might disagree. Consequently, it might have sought to ensure that whatever a jury decides will not be beyond respectable scientific reasoning on that issue and will be (broadly) scientifically acceptable (within the boundaries of science that the scientific community itself would not find unacceptable). This does not ensure that the overall verdict will be acceptable, but an important aspect of it will be. Because judges were now more involved arbiters of the reliability of expert testimony and its foundation (contrary to the Frye doctrine), the Court held that a judge must review the testimony of a scientific expert to ensure that it is grounded in the methods and procedures of science, which "connotes more than subjective belief or unsupported speculation," although the subject of scientific testimony need not to be known to a certainty. Consequently, the trial judge must... [conduct]... a preliminary assessment of whether the reasoning or methodology underlying the testimony is scientifically valid and whether that reasoning or methodology properly can be applied to the facts in issue. (emphasis added)

90 See Daubert, 509 U.S. at 588-589.
91 Daubert, 509 U.S. at 588-589.
93 Daubert, 509 U.S. at 592. Whether firsthand knowledge is a reliable analogy for scientific "reliability" is a separate and more difficult issue.
94 Grantor, "Daubert and the Acceptability of Legal Decisions," 127-131. There are deeper issues here to which I will return in later chapters (especially Chapter 8).
95 509 U.S. at 590.
96 509 U.S. at 590.
97 Daubert, 509 U.S. at 592-593. In dicta the Court outlined several nonexclusive, nonnecessary factors for courts to consider in evaluating experts' reasoning and methodology: (1) the falsifiability, or testability, of the theory guiding the technique used to reach the offered conclusion; (2) publication and peer review of the theory; (3) any known or potential rate of error of the technique; and (4) general acceptance within the relevant scientific community (593-594).
100 Daubert, 509 U.S. at 594-595 (emphasis added).
101 509 U.S. at 588.
102 509 U.S. at 596.
103 509 U.S. at 596 (citing both Federal Rules of Civil Procedure 56 and 50(a)).
104 509 U.S. at 596-97.
105 509 U.S. at 596-97.
particularly when their incorrectness is shown. However, incorrect hypotheses are of little use in the much quicker and more final context of a particular legal case. The Court strangely seemed to suggest that the scientific basis of expert testimony and a legal decision should be on even firmer ground than a result in the scientific field itself. The Court acknowledged that a judge will occasionally incorrectly exclude valid scientific methodologies, but that such exclusion is part of the balance to be struck in the legal context where the admission of an erroneous technique can have grave and irreparable consequences to the parties involved in an adversarial case.

Thus, the court seemed to reverse the roles of science and the law because it seemed to suggest that it was of greater importance that the science admitted into legal cases should be more certain than the scientific studies that provide the foundation for future scientific developments. This is even more odd when we recall that often legal cases are brought at a time when the relevant science is "on the frontiers of scientific knowledge" and unlikely to be established with great certainty.

The suggestions contained in this part of the court's opinion are particularly troublesome as we will see in Chapter 7 (especially given the Court's views in *Kumho Tire* and the fact of reasonable scientific disagreements).

**Joiner v. General Electric and Kumho Tire v. Carmichael:** The *Daubert* case was quickly followed by *General Electric v. Joiner* (1997) and by *Kumho Tire v. Carmichael* (1999), two Supreme Court decisions focusing primarily on procedural issues. The *Daubert* Court had not addressed the standard of review that should be applied by appellate courts in reviewing trial court decisions on the admission of expert testimony. Subsequent to *Daubert* most circuits held that an "abuse of discretion standard" applies, which means that a trial judge's ruling on the admissibility of scientific evidence must be "manifestly erroneous" or "clearly erroneous" before it can be overturned.

---

107 509 U.S. at 597.
108 509 U.S. at 597.
109 See 509 U.S. at 597. One implicit message indicated in this passage is that it is a permissible social cost for courts to mistakenly exclude "valid scientific methodologies." This contributes to a false negative mistake on scientific grounds. It also has irreparable consequences to the parties involved, if they have been wrongly harmed by others and have no possibility for corrective justice.
110 There is a further issue concerning whether judges in their admissibility decisions ought to be even-handed in protecting against factual false positives or factual false negatives. I argue in later chapters that they should.
112 See, e.g., *American & Foreign Ins. Co. v. General Elec. Co.*, 45 F.3d 135, 139 (6th Cir. 1995) (lower court exclusion of expert testimony on circuit breaker design must be clearly erroneous to show abuse of discretion); *United States v. Dorsey*, 45 F.3d 809, 815-16 (4th Cir. 1995) (applying abuse of discretion standard to lower court ruling on admissibility of forensic anthropologist's testimony); *Bradley v. Brown*, 42 F.3d 434, 436-37 (7th Cir. 1994) (holding that lower court's findings regarding doctors' testimony will not be overturned "unless they are manifestly erroneous").

Third and Eleventh Circuits took somewhat different views because admissibility decisions could be so decisive in determining the outcome of legal cases. The Supreme Court considered these issues in the second of the trilogy of cases to address expert testimony: *Joiner v. General Electric*.

Recall (from Chapter 1) that Robert Joiner worked as an electrician for the city of Thomasville, Georgia, *inter alia*, repairing and cleaning the city's electrical transformers, which used a mineral-based fluid as a coolant. Mr. Joiner alleged that his exposure to polychlorinated biphenyls (PCBs), furans, dioxins and other organic substances (several of which were known carcinogens) contributed to his lung cancer. He and his wife brought suit on these issues at a district court in Georgia.

The district court judge following *Daubert* excluded plaintiffs' experts from testifying, because the testimony did not rise above "subjective belief or unsupported speculation," thus rendering it inadequate to present a "material issue of fact" for a jury. The judge granted a motion for a summary judgment and dismissed the suit. Because the experts could not testify, the Joiners' case was at an end.

Plaintiffs appealed to the Eleventh Circuit Court of Appeals, which took a careful look at the District Court Judge's exclusion of plaintiff's expert witnesses. The Eleventh Circuit Court of Appeals found that the trial court had misunderstood plaintiffs' experts' methodology and ruled that when exclusion of evidence had such an "outcome determinative" effect on a trial, an appellate court should more carefully review admissibility decisions. It held that the District Court Judge had abused her discretion in excluding evidence. The Eleventh Circuit, thus, reversed the trial court. However, defendants sought review of the Circuit Court's decision in the Supreme Court, which decided on the appropriate standard of appellate review of a district court's admissibility ruling.

The Supreme Court overturned the Eleventh Circuit's procedural ruling, holding that "abuse of discretion" is the proper standard for reviewing trial court decisions concerning the admissibility of evidence and holding that federal appellate courts may invalidate such rulings only if the lower court "abuses its discretion." A trial judge has not abused her discretion, if, "[w]here there are two permissible views of the evidence, the choice between them cannot be clearly erroneous." Thus, if a judge admits or excludes an expert critical to a case, she cannot be overturned on appeal unless the decision was clearly

113 See *Paoli, Inc. v. Getty Oil Co.*, 33 F.3d at 741-52.
mistaken. This is not an impossible appellate hurdle to overcome but a very difficult one. A litigant could fail to have a critical witness admitted, have the case dismissed as a consequence, and not prevail on the admissibility decision on appeal.120

In addition, a majority of eight judges,121 excluding Justice Stevens, then proceeded to do something it had not done in Daubert—examine the details of the scientific record. In a surprisingly elaborate discussion of plaintiff’s expert’s evidence, the court applied the newly articulated standard for appellate review of evidence and upheld the District Court’s review of expert testimony and its foundation as not being an abuse of discretion. It held that the district court had not abused its discretion in rejecting each piece of evidence relied on by Joiner’s experts as inadequate to support the conclusions that he contracted lung cancer from exposure to PCBs. It found that plaintiffs’ reliance on studies of infant mice exposed to PCBs failed to support the conclusion that PCBs caused lung cancer.122 The Court also concluded that the District court ruled properly in excluding individually each of the epidemiological studies as providing a reasonable foundation for expert testimony.123

Respondent Joiner had argued, contra the district court, that the weight of the evidence methodology was reliable and that this court’s review had violated the Daubert principle that “the focus, of course, must be solely on the principles, not on the conclusions they generate.” To this a majority of the Supreme Court responded:

He claims that because the District Court’s disagreement was with the conclusion that the experts drew from the studies, the District Court committed legal error and was properly reversed by the Court of appeals. But conclusions and methodology are not entirely distinct from one another. Trained experts commonly extraplate from existing data. But nothing in either 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 ... 124

120 If we take the language seriously — where there are two permissible views of the evidence, the choice between them cannot be erroneous — this poses a difficult problem for a judge faced with testimony that is on the frontiers of scientific research, precisely where legitimate scientific disagreements are highly likely. The abuse of discretion standard suggests that judges’ admissibility decisions would not be reviewed no matter which choice was made, even if they consistently decided one way only, for example, always for plaintiffs or always for defendants. If both views are respectable, a judge should not choose between them.


Justice Stevens was sufficiently concerned about this section of the ruling that he dissented. He would have left review of the admissibility decision to the appellate court, which is closer to the evidence. In addition, he pointed out that the Supreme Court’s ruling did not remain faithful to Daubert’s insistence that the focus be “solely on principles and methodology, not on the conclusions they generate.”125 Moreover, because Joiner’s experts utilized a “weight of the evidence” methodology to support their conclusions, “[t]hey did not suggest that any one study provided adequate support for their conclusions, but instead relied on all of the studies taken together (along with their interviews of Joiner and their review of his medical records).”126 Because the focus of the trial court’s ruling “was on the separate studies and the conclusions of the experts, not on the experts’ methodology (‘Defendants . . . persuade the court that Plaintiffs’ expert testimony would not be admissible . . . by attacking the conclusions that Plaintiffs’ experts draw from the studies they cite’), the evidence assessment by the court of appeal was “persuasive.”127 Moreover, Stevens argued, both defendants and federal agencies utilize similar methodologies in drawing inferences from studies about the carcinogenicity of substances. Finally, “using this methodology, it would seem that an expert could reasonably have concluded that the study of workers at an Italian capacitor plant, coupled with data from Monsanto’s study and other studies, raises an inference that PCBs promote lung cancer.”128

The final Supreme Court case developing new law was Kumho Tire v. Carmichael. The admissibility issue concerned an engineer who, by training and experience, claimed to be able to determine, by inspection and a methodology he developed, after the fact of a tire failure whether a tire that had shredded and caused a car accident had been defective. The trial court excluded the engineer’s testimony but the Eleventh Circuit Court of Appeals overturned the trial court.129 The Supreme Court in turn reversed the Eleventh Circuit, holding that the Daubert factors may apply to all expert testimony, and that the abuse-of-discretion standard “applies as much to the trial court’s decisions about how to determine [scientific] reliability as to its ultimate conclusion.”130 Thus, the Court held that a judge has discretion both to decide how to conduct an admissibility review as well as making the actual admissibility judgment with both subjected to the abuse of discretion standard of review.

Interestingly, the unanimous court, perhaps aware that district courts might be overly zealous in their review of evidence, noted that trial courts need
discretionary authority "to avoid unnecessary 'reliability' proceedings in ordinary cases where the reliability of an expert's methods is properly taken for granted... as well as to avoid 'unjustifiable expense' and delay as part of their search for 'truth' and the 'just[ic] determin[ation]' of proceedings."131

The Kumho Court noted that the aim of admissibility should be to ensure "that an expert... employs in the courtroom the same level of intellectual rigor that characterizes the practice of an expert in the relevant field."132 A judge may exclude expert testimony that falls "outside the range where experts might reasonably differ, and where the jury must decide among the conflicting views of different experts, even though the evidence is 'shaky.'"133 I return to this guidance in Chapter 7.

The Admissibility Picture after the Daubert Trilogy

The picture of how trial court judges should review expert testimony following this trilogy of cases seems to be the following. Trial judges have a heightened duty to review expert testimony and its scientific foundation to determine whether "the reasoning or methodology underlying the testimony is scientifically valid and... whether that reasoning or methodology properly can be applied to the facts in issue." (emphasis added).134

Moreover, the Court seemed to have endorsed the view that the newly articulated admissibility rules were more liberal (would result in admitting a wider range of evidence) than the rejected Frye rules,135 but this has been put into doubt by the implementation of the decisions by other courts. Trial courts must focus on the reasoning and methodology of an expert, not her conclusions, unless there is too great a "gap" between the methodology and conclusion as Joiner reasoned. Judges are authorized to review all experts under their heightened duty of review, but have considerable discretion to decide both how to review their testimony and how to determine its admissibility. A trial judge has not abused her discretion, if "[w]here there are two permissible views of the evidence, the choice between them cannot be clearly erroneous."136

Thus, if a judge admits or excludes an expert critical to a case, she cannot be overturned on appeal unless the decision was clearly mistaken. This discretion is needed to ensure that courts can avoid both unnecessary delay as well as unjustifiable expense in conducting these proceedings.

Finally, a trial judge should seek to assure that an expert "employs in the courtroom the same level of intellectual rigor that characterizes the practice of an expert in the relevant field,"137 and may exclude expert testimony that falls "outside the range where experts might reasonably differ, and where the jury must decide among the conflicting views of different experts, even though the evidence is 'shaky.'"138 That is, although there is an emphasis on experts evaluating evidence and reasoning about it in the courtroom as they would in their own fields, this decision suggests that trial and appellate judges also should be alert to admitting experts whose testimony is within "the range where experts might reasonably differ."139

Subsequent to these decisions, the Federal Rules of Evidence have been amended to reflect the decisions and codify the changes. Rule 702 now reads

If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise, if (1) the testimony is based upon sufficient facts or data, (2) the testimony is the product of reliable principles and methods, and (3) the witness has applied the principles and methods reliably to the facts of the case.140

The proposed rule requires that expert testimony must be "based upon sufficient facts or data," the testimony itself should be the product of "reliable principles and methods," and the expert should apply the principles and methods "reliably" to the facts of the case. In some respects the amended Rule 702 more clearly separates issues for courts to consider, distinguishing between the facts or data, the testimony based on them and the application of the testimony to the facts of the case. (There may be difficulty in finding "reliable principles and methods" for expert inferences, as I consider in later chapters.)

Comments on the amended Rule also indicate various "factors" to assist judges in reviewing testimony for reliability (building on the original Daubert decision).141 There is some ambiguity as to what these "Daubert factors" should

131 Kumho Tire Co. v. Carmichael, 526 U.S. at 153 (Justices Scalia, O'Connor, and Thomas joined the majority, but also issued a concurring opinion cautioning courts not to perform admissibility reviews "inadequately" (526 U.S. at 159).
132 Kumho Tire Co. v. Carmichael, 526 U.S. at 152.
133 Kumho Tire Co. v. Carmichael, 526 U.S. at 153.
134 Daubert, 509 U.S. at 592–593.
135 Kumho Tire Co. v. Carmichael, 526 U.S. at 153.
137 Kumho Tire Co. v. Carmichael, 526 U.S. at 152.
139 Kumho Tire Co. v. Carmichael, 526 U.S. at 153.
140 Advisory Committee on Evidence Rules, "Proposed Amendment: Rule 702" (December 2000) (emphases in original and indicate new material).
141 Additional considerations are referenced in Advisory Committee on Evidence Rules, "Proposed Amendment: Rule 702" (December 2000): (1) Whether experts are "proposing to testify about matters growing naturally and directly out of research they have conducted independent of the litigation, or whether they have developed their opinions expressly for purposes of testifying." Daubert v. Merrell Dow Pharmaceuticals, Inc., 43 F.3d 1311, 1317 (9th Cir. 1995). (2) Whether the expert has unjustifiably extrapolated from an accepted premise to an unfounded conclusion. See General Elec. Co. v. Joiner, 522 U.S. 136, 146 (1997) (in some cases a trial court "may conclude that there is simply too great an analytical gap between the data and the opinion proffered"). (3) Whether the expert has adequately accounted for obvious alternative explanations. See Cleary v. Burlington N.R.R., 29 F.3d 499
be applied: the testimony itself, studies on which the testimony is based, or inferences experts make from underlying studies. These are each different. Some of them seem to apply most naturally to studies or tests on which expert testimony is based and less well to expert inferences. Others are more general admonitions.

Several observations by the Committee are of interest. The Committee found that, while the rejection of expert testimony is “the exception rather than the rule...the trial court’s role as gatekeeper is not intended to serve as a replacement for the adversary system.” It also reiterated the Daubert Court’s admonition about the importance of “vigorous cross-examination, presentation of contrary evidence, and careful instruction on the burden of proof” as means to attacking “shaky but admissible evidence.” Finally, following the Third Circuit Court of Appeals, it notes that proponents do not have to demonstrate to the judge by a preponderance of the evidence that the assessments of their experts are correct, they only have to demonstrate by a preponderance of evidence that their opinions are reliable. The evidentiary requirement of reliability is lower than the merits standard of correctness.

All this is salutary, as I argue in Chapter 7. However, there remain some issues in understanding the change in law and in how courts are applying the decisions from Daubert as I consider in what follows.

THE AFTERMATH OF THE BENDECTIN LITIGATION

Critiques

Although in Daubert the Supreme Court held for plaintiffs, they lost on remand in the Ninth Circuit Court of Appeals, and so did virtually every other Bendectin (9th Cir. 1994) (testimony excluded where the expert failed to consider other obvious causes for the plaintiff’s condition). Compare Claas with Ambrosini v. Maybahn, 712 F.3d 129 (D.C. Cir. 1996) (the possibility of some uneliminated causes present a question of weight, so long as the most obvious causes have been considered and reasonably ruled out by the expert). 4. Whether the expert “is being as careful as he would be in his regular professional work outside his paid litigation consulting.” Sheehan v. Daily Racing Form, Inc., 104 F.3d 940, 942 (7th Cir. 1997). This is quite similar to the “intellectual rigor” consideration from Kumho Tire Co. v. Carmichael, 119 S.Ct. 1167, 1175 (1999). (5) Whether the field of expertise claimed by the expert is known to reach reliable results for the type of opinion the expert would give. See Kumho Tire Co. v. Carmichael, 119 S.Ct. 1167, 1175 (1999). 142

As a result critics of the tort law have frequently used this litigation as an exemplar that there are substantial problems with the tort law.

Direct monetary costs to Bendectin defendants might have been “in the range of $100 million” with plaintiffs’ firms spending “tens of millions.” Moreover, the tort system tends to impose a tax of 50–70 percent on every dollar transferred from defendants to plaintiffs. However, in the Bendectin litigation, this tax might have approached 100 percent, because little money changed hands despite two thousand cases and twenty years of litigation.

Indirect costs included the withdrawal of Bendectin from the market because of litigation pressures and the alleged deterrence of research and innovation in the pharmaceutical industry, which has substantial capacity to improve public health and welfare.

Finally, more strident critics have argued that “there is no statistically significant association between Bendectin and birth defects,” and that “Bendectin is safe for both the mother and the unborn child...” In the extreme, some claimed that the number of birth defects actually increased since Bendectin was removed from the market, because violent nausea might cause injuries to developing embryos. 149

Correctives

In his extensive study, Bendectin and Birth Defects, Michael Green argues that the allegations about the tort law are far from established. “Many commentators disagreed with several aspects of the [initial] crisis account, its causes, and its consequences.” There was a substantial increase in tort cases filed from 1974–1985, with asbestos, a quite potent toxicant, making up 31 percent of the cases, but significantly, “in absolute terms, the incidence of claiming is well below the incidence of injurious events that might justify a claim.” Other scholars agree; the best studies available about the tort law indicate that only a small percentage of persons wrongfully injured by a doctor or a company’s products ever approach a lawyer seeking redress for their injuries. For example, when there are known tortiously actionable injuries, only about 2–3 percent

143 509 U.S. at 595
144 In re Puelll R.R. Yard PCB Litigation, 35 F.3d 717, 744 (3d Cir. 1994)
become lawsuits, a fairly common finding.\textsuperscript{152} In federal law, only a small part of the total tort filings, the tort law area was hardly the fastest growing area of litigation, with general non-tort filings growing faster.\textsuperscript{153}

More specifically on the Bendectin litigation, Merrell Dow withdrew Bendectin from the market because of the litigation. This has “no doubt deprived some pregnant women of relief from the nausea and vomiting of pregnancy, . . . [but] the loss of Bendectin is not nearly as tragic as some of the critics’ semi-hysterical claims have made it out to be.”\textsuperscript{154} One study found that Bendectin “relieved morning sickness in only 10% more of the women than took a placebo. . . . [and for nausea alone it provided] benefit to 23 percent more of those receiving Bendectin than those who were given a placebo.” For vomiting there was only a 7 percent difference.\textsuperscript{155} (These are “relative benefit” ratios of 1:1.1, 1:1.23, and 1:1.07. Compare them with defense and some court claims that human epidemiological studies must exhibit “relative risks” greater than 2.0 before an expert can rely on them for expert testimony.)

Moreover, the absence of Bendectin from the market was not necessarily a bad thing as this “avoided the significant overuse of the drug that occurred in the 1970s . . .” Green notes that physicians and women should be cautious in exposing a developing embryo to many such substances, since this is one of the most biologically vulnerable periods in a human life as an embryo grows from one cell to billions in a short period of time.\textsuperscript{156}

On the scientific claims, Green concludes that some of the most strident critics have gone much too far in asserting the safety of Bendectin in absence of scientific studies showing adverse effects. For example, one critic claimed that “. . . overwhelming scientific evidence [shows] that Bendectin is safe for both the mother and the unborn child.”\textsuperscript{157} Green counters, “The range of risk [of shortened limbs from exposures to Bendectin] that is consistent with the scientific evidence is small, but it still exists.”\textsuperscript{158} Nonetheless, that risk is too small to permit a plaintiff to satisfy the proof conditions in a toxic tort case.

Is the tort system a social drag on the economy? “[W]e simply do not know. . . . Moreover, there is good reason to be skeptical that Bendectin signals much of anything about the net social impact of tort law on the pharmaceutical industry.”\textsuperscript{159} A Rand study “observes that liability effects on innovation cannot be observed or quantified.”\textsuperscript{160} At most tort liability might shift some research funding from research on “modest drugs” to research on those that would represent “a major break-through and the promise of huge profits.”\textsuperscript{161}

Consequently, “[r]ather than being emblematic, the Bendectin litigation may be idiosyncratic [or even aberrational] in assessing the role of the tort system’s impact on pharmaceutical technology and innovation.”\textsuperscript{162} Although in the end Green would add the Bendectin litigation to the negative side of the mass toxic tort ledger, he is cautious in “overemphasizing its impact.”\textsuperscript{163} Bendectin’s negative impact must be balanced against the litigation concerning the Dalkon Shield IUD, asbestos, MER/29, thalidomide, alachlor, atrazine, formaldehyde, and perchloroethylene, all justified tort claims.\textsuperscript{164} The tobacco litigation, which had not developed extensively at the time he completed his book, should be added to the positive side of the ledger as well.

CONCLUSION

The review of the Daubert trilogy of cases that modified the law on the admissibility of expert testimony and its scientific foundation sought to show where in the legal process this occurs. However, much more is required to analyze the impact of these changes on the legal system, the scientific community, and ordinary citizens whose lives are affected by the tort law. To evaluate these changes, we need a deeper understanding of some of the legal implications as well as a better understanding of the science that will be needed in such cases. These are subjects for the chapters that follow.

\textsuperscript{152} Michael J. Saks, “Do We Really Know Anything about the Behavior of the Tort Litigation System – and Why Not?” Pennsylvania Law Review 140 (1992): 1184–1185. (Saks cites numerous studies showing how few legitimately injured persons actually file suits and fewer still proceed to trial.)

\textsuperscript{153} Federal government suits for recovery of overpayments to individuals or firms, social security cases, and contract litigation all increased faster than tort cases, yet these areas were not “in crisis.” (Saks, “Do We Really Know Anything,” 1200–1201.)

\textsuperscript{154} Green, Bendectin, 336.

\textsuperscript{155} Green, Bendectin, 336.

\textsuperscript{156} Green, Bendectin, 337 (quoting a standard textbook on the effects of drugs on the fetus).

\textsuperscript{157} Green, Bendectin, 330 (quoting Howard Denman, “Improving Litigation Against Drug Manufacturers,” 413, 427–428).

\textsuperscript{158} Green, Bendectin, 330 (emphasis added).

\textsuperscript{159} Green, Bendectin, 339.

\textsuperscript{160} Green, Bendectin, 339.

\textsuperscript{161} Green, Bendectin, 340.

\textsuperscript{162} Green, Bendectin, 341.

\textsuperscript{163} Green, Bendectin, 341.

\textsuperscript{164} Berger, “Eliminating General Causation,” 2135.
In *Daubert*, the Supreme Court correctly saw that lower courts had reviewed the admissibility of expert testimony and its foundation on the basis of a principle — the *Frye* “general acceptance” test — that had been superseded by the more liberal admissibility guidance of the Federal Rules of Evidence. At the same time, when it sought to articulate guidance for this activity, it heightened the gatekeeping duties of judges. However, in doing this, it entered intellectual territory that is not readily accessible to judges with their typical training.

The Supreme Court did not mention and seemingly disregarded its own decision of a decade earlier in *Barefoot v. Estelle*. This decision had held that cross-examination and jury assessment of witnesses’ credibility and reliability were sufficient to protect a criminal defendant in a death penalty case against dubious and unreliable expert testimony that was widely criticized by the expert’s own profession.¹ By the time *Daubert* was decided in 1993, instead of merely rejecting *Frye*, as Chief Justice Rehnquist argued in dissent, and going beyond the plain language of the Federal Rules of Evidence, it created a “reliability” screen for expert testimony.² This contrasted with *Barefoot v. Estelle* and with much of the previous application of the *Frye* test. In many jurisdictions, the *Frye* test only applied to generic tests, studies, technological devices, and scientific procedures that provided the foundation of scientific testimony, not to scientific opinions or the inferences of scientists.

In this chapter, I sketch some issues the court created by entering the intellectual terrain of epistemology, philosophy of science, and the nature of causal inferences. None of these issues is easy, but the Court took them up anyway. The easy part of the *Daubert* opinion was the rejection of the *Frye* test. A much more difficult issue for the courts and for the rest of us is understanding the import of the decisions and how they should guide the admissibility of expert testimony and its scientific foundation. Subsequently, the Federal Rules of Evidence were amended to reflect the Supreme Court’s opinions and to guide federal judges; this is also introduced.

**OBVIOUS LESSONS**

First, the Court gave federal trial judges a heightened “gatekeeping” duty to review expert testimony and its scientific basis. This gatekeeping duty has loomed much larger than the amount of space the Court devoted to discussing and characterizing it. Only two places in the majority opinion *Daubert* did the Court use the word “gatekeeping.”³ The Court has clearly been understood as authorizing trial courts to serve as “nontrivial” “gatekeepers” of expert testimony.

The screening responsibility conveyed by the gatekeeping requirement has become quite substantive. Experts are no longer merely reviewed, as they were previously in federal courts and as they continue to be in some state jurisdictions, (a) to see that they are properly qualified by knowledge, skill, experience, training, or education to testify about technical issues they are asked to address; and (b) to ensure that their scientific opinions are based upon generic techniques, tests, studies, and scientific procedures that are viewed as reliable by the scientific community. Subsequent to the *Daubert* decision if an expert opinion is not judged to be sufficiently reliable, judges may exclude it.

The Court’s language suggests that judges must ensure that expert testimony purporting to be scientific must indeed be based upon scientific reasoning and methodology.⁴ The aim of this requirement seems to be to ensure that legal decisions will comport more closely with what is known scientifically or what can be reasonably inferred from the science that is relevant to the legal issues. However, at least some commentators have suggested that the courts were really struggling with a much more basic issue — how to preclude charlatans or perhaps even “liars” from testifying.⁵ This was a perceived concern about torts in the late 1970s and early 1980s — perhaps some experts could be found to testify on nearly any subject and say nearly anything that their employers needed to be said if they were paid for testifying. Indeed, it is likely that some experts on both sides of tort cases testified as their employers wanted. And there are a

---


² Feigman et al., *Modern Scientific Evidence*, 12 (suggesting that the meaning of the Federal Rules of Evidence was far from “plain,” and that *Daubert* substantially changed past practice).


⁴ “Faced with a proffer of expert scientific testimony, then, the trial judge must determine at the outset ... whether the expert is proposing to testify to (1) scientific knowledge that (2) will assist the trier of fact to understand or determine a fact in issue. This entails a preliminary assessment of whether the reasoning or methodology underlying the testimony is scientifically valid and of whether that reasoning or methodology properly can be applied to the facts in issue.” *Daubert*, 509 U.S. at 592–593 (emphasis added).

⁵ Gottesman, “From Barefoot to Daubert to Joiner,” 753.
variety of incentives and more subtle relationship processes that might select for experts to testify as their employers wish. Courts also might have been concerned about experts who expressed "unsupported speculation," where they had insufficient knowledge about which they are testifying or perhaps had mere "subjective beliefs" that were not as fully grounded in scientific studies as might be reasonably required. In order to address these issues the Court sought to ensure that experts base their testimony on what is known or what can be reasonably inferred from them.

The principle from Daubert is that expert testimony that is not "reliable," that is, that probably not grounded "in the reasoning or methodology" of science should be excluded. According to Daubert, when courts determine the admissibility of evidence, they should ask whether the evidence was more likely than not based upon scientific reasoning and methods. If the answer is no, the evidence should be excluded. If the answer is yes -- that the evidence probably resulted from scientific reasoning -- then it should be admitted. The courts are vague about "scientific reasoning." I consider it in several chapters that follow.

Although courts have considerable latitude in screening evidence, if the questions are framed properly, the answers to them appear to be more determinate than some appellate opinions suggest. The reason for this is that scientists routinely utilize certain kinds of evidence to come to their conclusions concerning toxicity: epidemiological evidence, if it is available, as well as animal, short-term toxicity, structure-activity, and mechanistic studies. If a particular scientist relies on such scientifically relevant evidence and evaluates it as do other respectable scientists, but assigns somewhat different weights to the studies, the scientist more likely than not has utilized scientific reasoning and methods in making inferences from the data, even if her conclusions do not necessarily accord with other experts' conclusions. Thus, it may be more difficult than courts have suggested to show that an expert's reasoning is probably not reflective of respectable scientific reasoning on a particular toxicological issue. This is especially the case when respectable scientists have a range of views on an issue, as they often do. Scientific experts frequently disagree, even when evaluating the same evidence and engaging in quite good science. I return to these issues in Chapters 4, 5, and 6.

The amended Rule 702 appears to be an improvement on the original Daubert trilogy. The modified Rule 702 permits expert testimony provided "(1) the testimony is based upon sufficient facts or data, (2) the testimony is the product of reliable principles and methods, and (3) the witness has applied the principles and methods reliably to the facts of the case." Judges accordingly have three tasks. They must assess whether experts have relied on the kinds of studies scientists typically would utilize and whether such studies (taken together) constitute "sufficient" evidence for expert testimony. (For courts to evaluate the scientific "sufficiency" of such evidence may be difficult [Chapter 4].) They must assess whether the testimony (one might say the "inferences") from the data are based on reliable principles and methods. (This, too, is not easy.) And they must judge whether the testimony is properly applied to the facts of the case. Rule 702 appears to be an improvement on the Supreme Court decisions because it distinguishes different tasks for courts; this may assist their analysis. However, it also may indicate a further shift in substantive legal policy, about which some have already expressed concerns.

According to Daubert (and echoed in Rule 702 comments) trial court review is to be guided by four nondefinitive factors, if they are appropriate for the review in question. The Kalamo Tire case emphasizes the flexibility in their use. Judges must rule on the admissibility of litigants' evidence based on the considerations that are pertinent to the facts of the case in question. Moreover, to what are the various Daubert factors to be applied? There seems to be some confusion on this point. Do they apply to the underlying tests, studies, or technologies for generating information, as is often suggested, or do they apply to the inferences experts make from the studies? Two of the original four factors seem more naturally to apply to the underlying studies -- testability (falsifiability) and error rate -- and some commentators recommend this. Sometimes these two factors are applied to scientists' inferences, but they seem much less at home there, much more difficult to utilize and scientists themselves seem

---

7 Daubert, 509 U.S. at 588.
8 Daubert, 509 U.S. at 589-590.
9 Not all judicial rulings on scientific evidence raise issues of toxicology. Some merely deal with correctly assessing circumstantial evidence involving exposures to toxic substances. These are not addressed here, but for examples see Moore v. Ashland Chemical, Inc., 151 F.3d 269 (5th Cir. 1998) (considering the adequacy of circumstantial and scientific evidence that exposure to a spilled solvent caused respiratory tract disorders); Wright v. Willamette Indus., Inc., 91 F.3d 1105 (8th Cir. 1996) (concerning circumstantial and scientific evidence that exposure to formaldehyde-impregnated wood dust caused respiratory disorders); and Zuchowicz v. United States, 140 F.3d 381, 389-391 (2d Cir. 1998).

10 Advisory Committee on Evidence Rules, "Proposed Amendment: Rule 702" (December 2000) (emphasizes original and indicate new material).
11 Margaret Berger expresses concern about testimony being based on "sufficient facts or data," since this may well indicate a significant shift in legal policy and give federal courts a more substantive role than they should have in reviewing the mere admissibility of expert testimony for cases originating in state courts. (Berger, "Upsetting the Balance Between Adverse Interests," 323. The concern is that federal courts could intrude on state rights to jury trial by means of an admissibility hearing.)
12 "The conclusion, in our view, is that we can neither rule out, nor rule in, for all cases and for all time the applicability of the factors mentioned in Daubert, nor can we now do so for subsets of cases categorized by category of expert or by kind of evidence. Too much depends upon the particular circumstances of the particular case at issue... . Engineering testimony rests upon scientific foundation, the reliability of which will be at issue in some cases... . In other cases, the relevant reliability concerns may focus upon personal knowledge or experience" (Russo v. Carmenich, 526 U.S. 137, 150).
unlikely to use them for this purpose. The other two — general acceptance and peer review — appear to apply to "more fundamental activity[ies] of scientific community." Of the factors identified by subsequent courts and noted by the Advisory Committee on Evidence Rules two seem to go to assessing the credibility of the expert (whether experts propose to testify about matters growing naturally out of their own research, and whether an expert is being as careful in legal testimony as he or she would in professional work outside the courtroom). One concerns the field itself (whether it is known to be reliable). Two concern scientific inferences from studies (whether there is "unjustifiable extrapolation from an accepted premise to an unfounded conclusion" and whether an expert has adequately accounted for obvious alternative explanations).

The Daubert and Joiner rulings reveal a tension between different Court concerns. One "emphasizes that the Federal Rules are designedly permissive with respect to expert testimony." The other draws attention to the "gatekeeper" role of the trial court. The attention to the gatekeeper role emphasizes that they have a heavy responsibility to exclude unreliable evidence; that they are to take this gatekeeper responsibility more seriously than perhaps they did in the past; and that the exclusionary rules should be used more aggressively with respect to expert testimony, because such testimony can have an undue impact on the jury. The Kamho Tire decision may ameliorate the strong gatekeeping role that some courts appear to have utilized, but it is not entirely clear. For example, Associate Justice Breyer, who wrote the opinion, clearly envisions that some cases involving experts will require little review, while other will require a more extensive assessment.

The trial court must have the same kind of latitude in deciding how to test an expert’s reliability, and to decide whether or when special briefing or other proceedings are needed to investigate reliability, as it enjoys when it decides whether or not that expert’s relevant testimony is reliable. . . . Otherwise, the trial judge would lack the discretionary authority needed both to avoid unnecessary "reliability" proceedings in ordinary cases where the reliability of an expert’s methods is properly taken for granted, and to require appropriate proceedings in the less usual or more complex cases where cause for questioning the expert’s reliability arises. Indeed, the Rules seek to avoid "unjustifiable expense and delay" as part of their search for "truth" and the "just[ification]" of proceedings.

As already noted, the Court states that experts’ testimony may be excluded when it falls "outside the range where experts might reasonably differ," and where the jury must decide among the conflicting views of different experts, even though the evidence is "shaky." The language from Kamho Tire is to be commended, but it is not clear that it has penetrated the federal court system, because it appears that tensions between the "substantive gatekeeping" and the liberal admissibility remain. The presence of tensions between different court goals invite lower courts to make admissibility decisions in different directions: toward more or less permissive admissibility decisions, and it has resulted in some contrary decisions between courts and between circuits. For example, recall the Parodel cases from Chapter 1 as well as other disagreements.

There is a more pessimistic view as well. By formulating their standard of appellate review as they did in Joiner and Kamho Tire, the Court has sent the message that trial courts will be upheld on admissibility decisions unless their decisions are "manifestly erroneous." An implicit message might be that they do not want appellate courts to be bothered to review admissibility decisions, unless there is some quite mistaken decision. Thus, there will be less oversight of district courts by appellate courts than there would have been had Joiner been decided differently.

The other side of this issue is that trial courts have greater responsibilities on their own to ensure that their decisions are reasonable and fair, as appellate courts have limited authority to correct any lower court mistakes. Ultimately, if trial courts have too much difficulty in addressing the admissibility of scientific testimony, appellate courts may have to intervene to a greater extent simply to ensure basic fairness between litigants.

Social science evidence about federal law is beginning to suggest that trial courts are excluding more experts and the decisions are strongly asymmetrically against plaintiffs. For example, as already noted, since the Daubert decision, of the cases ending in summary judgments before trial, the rate has more than doubled with 90 percent of the terminated cases going against plaintiffs.

19 Kamho Tire v. Carmichael, 526 U.S. at 153, referring to Daubert, 509 U.S. at 596, 113 S.Ct. 2786.
21 Dixon and Gill, Changes in the Standards for Admitting Expert Evidence in Federal Civil Cases.
contrast, courts could recognize the myriad kinds of evidence that are available and that can be integrated to provide evidence that a product causes human harm. Moreover, if the above pragmatic barriers are common, as they appear to be, courts should permit testimony as early as possible in the history of knowledge of the toxicity of a substance. This will facilitate compensation for wrongfully injured parties and strengthen deterrence messages to firms who manufacture and market products that can harm fellow citizens.  

Thus, courts have alternatives in how they conduct admissibility reviews. How have they chosen and what reasons have they given for some of their decisions? This is the topic of Chapter 6.

146 Green, Bendtson, 192, 208.

6

Science and Law in Conflict

When scientific evidence is centrally needed to assist legal decisions, the Supreme Court, *inter alia*, aimed at increasing the chances that legal decisions would more closely follow or at least not be at great odds with the relevant science. This is not as easy a task as it might have seemed. For one thing, generic tensions between the law and science hamper easy pursuit of this goal. In addition, given the complexity, subtlety, and near inscrutability of some scientific evidence, judges understandably, but regretfully often struggle with scientific studies and reasoning. These difficulties increase the more complex and subtle the evidence becomes.1 Moreover, judges may have disagreements with one another about how scientific evidence should be reviewed, given its complexity.2 Finally, the chances that admissibility decisions will result in mistaken judgments are even greater given some of the pragmatic problems reviewed in the previous chapter. That is, even when judges review quite good evidence, there are numerous opportunities for errors. When there are myriad pragmatic barriers to obtaining good evidence about the toxicity of substances and less than optimal evidence is available, the potential for stresses and strains increases. Failures to attend successfully to these issues pose threats to the legitimacy of the law as an institution.

The current chapter focuses on some of these issues; the next chapter suggests a partial corrective to them. This chapter first reviews some tensions between science and the law that can affect how well they can function together. Different standards of proof central to each, different time frames within which each operates, different concerns about the distribution of mistakes, and different approaches to uncertainty, simplicity, and complexity create a context

2 This is a problem Judge Glikiewicz notes in *Stevens v. the Secretary of Health and Human Services*, 2001 WL 387418 (Fed. Cl.). He was concerned that federal magistrates' decisions in vaccine injury cases were utilizing different conceptions of scientific evidence in deciding cases and proposed a common framework.
in which judges can err. Such tensions on top of the pragmatic barriers already considered increase courts’ burdens.

Next, I review some district and appellate court cases in which the admissibility of scientific evidence has been at issue and assess the reasons courts have given for their decisions in light of how scientists typically address similar evidence. The examples suggest that some courts have had difficulty in reviewing and understanding scientific studies and reasoning, or have adopted quite different standards for assessing scientific evidence than scientists themselves would. These problems are of concern, because it appears that at least some courts are struggling with the task of admitting and excluding evidence. To the extent they do, this undermines the aim of making the law more consistent with the science pertinent to the cases and affects the just resolution of legal issues. Recognizing the shortcomings of simple and restrictive guides can serve as a corrective to past errors and point the way to improving how the tasks could be approached better and more sensitively (Chapter 7).

Before the Daubert decision, a number of commentators were concerned that some courts might have permitted into trial scientific testimony that in Justice Breyer’s words (quoting the physicist Wolfgang Pauli) might not have been “good enough to be wrong.” Published district and appellate court opinions now suggest that some judges in reviewing scientific evidence are issuing decisions on scientific testimony and its foundation contrary to how most scientists themselves would assess the same evidence. Some courts excluded experts who had utilized kinds of evidence on which scientists would routinely rely for making causal inferences about human harm, whereas others have required that testimony must be based on evidence that scientists might find quite valuable, but not necessary for causal inferences. They appear to have had difficulties with the scientific relevance of some evidence and even more had difficulties appreciating how complex pieces of evidence can fit together to provide a scientific explanation that is more likely than not reliable.

It is important to note, however, that although judges might have arrived at these judgments on their own, litigants on the other side of the case may have contributed to their views. Litigants often urge that courts should demand higher certainty, human studies or other stringent requirements. In some respects we should not be surprised by such arguments. We saw how tempting this could be from a scientific point of view (Chapter 5). Moreover, some have adopted it as a more conscious strategy. “Doubt is our product,” has long been used by the tobacco industry to avoid responsibility for the adverse effects of tobacco exposure. Other industries have followed suit, including the lead, vinyl chloride, and asbestos industries, \textit{inter alia}. Thus, some defense lawyers and experts appear to have engaged in “junk science” of precisely the kind the Supreme Court aimed to exclude. This newer junk science is more insidious because it often has a patina of acceptability to it that can be difficult to pierce to its misleading core.

When courts rule on evidence and reasoning in ways that are at odds with the scientific community, these rulings are often propagated throughout the legal system and then perpetuated via a written record in the form of precedents or merely from “following the Joneses” of sister courts. This multiplies mistakes of science and decreases the possibility of justice throughout the legal system.

**GENERIC TENSIONS BETWEEN SCIENCE AND THE LAW**

**Tension in Goals**

A common aim of science is epistemic: Scientists first seek to describe and, then to explain and understand accurately the phenomena they study. Their accumulating data and describing phenomena are the easier parts of research. Making inferences from the descriptions and data to explain and understand the phenomena are more difficult.

An overriding virtue of science might be seen as pursuit of truth about the world or the small portions of it under study, with an emphasis on pursuit of this truth for its own sake – not typically a means to something else. This last point needs qualification, of course, as much contemporary discussion of science emphasizes the value of scientific discoveries for developing beneficial technologies and other products that will improve human life. On the one hand, scientists might study quarks, which at the moment appear to have little pertinence for human institutions, or, on the other hand, they might try to understand the causes of AIDS or how a drug controls cancer, both of which can have substantial consequences for humans. An appealing virtue underlying scientific inquiry, however, is pursuit of truth about a subject for its own sake.

This sketched reminder about scientific research is most apt for what we might consider theoretical science where pursuit of the truth for its own sake to answer empirical questions seems most at home. The picture is more complicated for scientific fields that have more varied goals such as toxicology.

Some toxicologists aim to recognize, identify, and quantify the hazards from toxic substances to humans and the environment, some to create new drugs or

---

4 Brown & Williamson - Smoking and Health Proposal, Doc. No. 332506, at http://tobaccodocuments.org/bw/332506.html. See also David Michaels, "Doubt Is Their
pesticides, whereas some engage in more basic research in order to understand the mechanism(s) of acute and chronic illnesses caused by toxicants. This last is the theoretical focus of the field. Toxicology is representative of many fields that courts need to assist them in addressing the adverse effects of human exposures to substances— they have both theoretical and more pragmatic or applied aspects. Failure to distinguish between these different parts of a field may lead to some confusions. One that can be of importance in the law is to conflate procedures and implicit standards of proof used for theoretical advances in a field with the epistemic procedures and standards that would be more germane for pragmatic areas of the same field. The procedures and standards may not be identical. For example, Michael Gallo writes that toxicology "is both a science and an art"... with the science being the observation and data gathering that is typical of the discipline and the art being utilization of that same data to "predict outcomes of exposure in human and animal populations." His very use of different terms for these activities suggests some substantive differences between the areas they designate. He adds that toxicological theories have "a higher level of certainty than do [predictions]" of toxic effects from a group of animals to humans or to another group of other animals, thus suggesting that the different facets of the field are subject to somewhat different standards.

Typically, scientific research embodies systematic procedures for developing explanatory answers to empirical questions that scientists have posed and are trying to answer. Indeed, part of what gives science its honorary connotation in at least some contexts is the fact that the systematic procedures for answering questions are paradigmatic of some of the best ways of addressing and answering empirical questions. As David Goodstein of the California Institute of Technology puts it, "The things that science has taught us about how the world works are the most secure elements in all of human knowledge"... although he distinguishes between "science at the frontiers of knowledge (where not all is understood) and "textbook science that is known with great confidence." The time scale of scientific research differs from that of the law, as one might suppose for an intellectual area in which pursuit of truth for its own sake is a leading virtue. A scientist pursuing research questions for their own sake cannot anticipate when discoveries will be made, satisfying explanations reached and understanding achieved. In addition, when experiments can take considerable time to conduct, the problem is especially difficult, or progress must depend on the efforts of many researchers, this will add to the time before scientific understanding is achieved.

Moreover, the end products of scientific research are relatively open-ended, because of complexity, scientists' inability to predict the outcome of experiments and when understanding will be reached, and researchers using noninductive inferences to explain the phenomena. Even comparatively settled conclusions are open to revision on the presentation of new data, theories, or discoveries that might overturn existing explanations in favor of better ones. The open-endedness of science and the practices that accompany it can affect the communication between science and the law, as we have seen. Scientists, thus, may hedge even well-supported claims in which they strongly believe. Judges need to recognize this practice because it can mislead courts and disadvantage meritorious litigants who have respectable evidence but who must carry the burden of proof to establish a claim.

In scientific endeavors, there is a collaborative aspect that has several features. Typically, research results are cumulative—one scientist or research group builds on the progress of others; personal collaborations across space and more impersonal collaboration across time contribute to the development of fields. In developing a theory or insight into how the world works, scientists typically assimilate the results of others, augment that with their own contributions and then develop or refine a new view. In addition, progress in science typically results in developing a consensus about the subject of study. The new results are much better secured when an appropriate subfield of scientists have been persuaded of the view.

Finally, as we discussed earlier, scientists for the most part do not attend to the distributive consequences of their results or mistakes that can result from research outcomes. They give less attention to false negatives than to false positives. Much toxicity research, conducted in workplaces, does not attend to adverse effects for more varied subpopulations.

By contrast with science, torts is a body of law that seeks to provide the means by which individuals who have been harmed by the actions of others may receive compensation for the injuries they have suffered. As a means to this end it provides a forum to adjudicate disputes between parties where one is claiming that the other wrongly harmed her.

10 Gallo, "History and Scope of Toxicology," 3.
11 Gallo, "History and Scope of Toxicology," 3.
12 As indicated (Chapter 3), the use of systematic procedures for finding answers to empirical issues does not necessarily mean that there is a univocal "scientific method" or that such systematic procedures are the same from field to field, or subdiscipline to subdiscipline.

14 That is, it is not part of basic scientific research to address how scientific mistakes might affect other institutions, and there appears to be little concern within science with how false positives and false negatives might affect the institution of science (apart from avoiding false positives). This is very different from the law, which has a self-conscious concern with such matters.
Justice is the leading virtue of the law. Adjudication must be done in such a way that the parties are treated fairly in the process and that, when a decision is rendered, justice has been done between the parties. Other normative goals also shape and guide legal dispute resolution, such as, inter alia, efficiency, administrative cost, wealth distribution, and morality. The law is, thus, “normative to the core.”15 This is something unlikely to be said of science, even though there are normative elements manifested in scientific inquiries and the institution at large.16

It is more difficult to point to readily identifiable systematic procedures that are characteristic of legal institutions in the same way that some systematic empirical investigations are characteristic of science. Of course, in litigation the adversarial presentation of facts, theories and interpretations of the law are central to getting at the legal truth of the issues and resolving disputes between parties. Adversarial procedures can attain the truth about an issue, but they can also mislead or create a much messier picture.

Although the time frame of science is open-ended, the law imposes several generic time constraints on the resolution of disputes. It seeks to ensure that disputes are addressed within a reasonable period of time and then resolved in a timely and conclusive manner. Thus, a plaintiff’s allegations of harm must be brought within specified periods of time fixed by statutes of limitations or they will not be considered.17 Once allegations are timely filed, the legal system imposes on itself some time constraints within which the issues must be taken up in court proceedings. After the legal issues have been litigated, typically they are considered settled once and for all by res judicata – “an issue that has been definitively settled by judicial decision.”18 Res judicata rules recognize the fact that “the purpose of a lawsuit is not only to [allow enough time] to do substantial justice but to bring an end to controversy.”19 The latter provides legal judgments that have stability and certainty so that “the parties and others may rely on them in ordering their practical affairs . . . but also so that the moral force of court judgments will not be undermined.”20 Although contemporary law seeks to provide sufficient time for legal issues to develop, it may not permit enough for the requisite scientific research to be completed to ensure full documentation of any adverse effects from suspected toxicants. Meanwhile, res judicata rules typically foreclose revisiting issues even if new evidence becomes available at a later time.21

The finality of law, Peter Schuck argues, creates another legal “bias” that can in turn generate tensions with science.

Because much law must be predicted, understood, and applied by many ordinary people with limited resources, simplicity is often a compelling legal virtue. Law cannot afford to be as nuanced as the realities it seeks to shape; it necessarily draws lines and creates categories that force many legal decisions into a binary mold; one is either in or out of the category, and it matters a great deal which.22 By contrast, scientists, trying to understand the natural world of biology or toxicological relationships, must take their subject as they find it, with all its subtlety, complexity, and uncertainty about effects. Even when scientists seek to model the biological world, their resulting products often have considerable complexity. Vern Walker has more extensively argued a similar point in a forthcoming paper. He notes not only the differences between the two areas of endeavor but also how legal pressures for more simple and understandable legal guides tend to push scientists in their testimony away from the needed complexity of their discipline.23 As we will see later, there are a number of temptations for judges to adopt comparatively simple rules to guide the review of evidence. However, if they aim to remain faithful to the science, they will have to approach this subject much more subtly and sensitively. More subtle approaches toward the science will serve both accuracy and fairness between litigants (Chapter 7). Courts will need to be sensitive to the simplicity-complexity tension between these two institutions in order to do justice to both, but also will need to review the scientific foundations of expert testimony in a way that does credit to the complexity of the science.

Finally, although the distributive effects of scientific research usually receive little attention, the distributive effects of legal rules and decisions are of preeminent importance. Distributive considerations are central to many issues in the law, but consider only two: that there is justice between parties and that admissibility procedures treat litigants equitably (discussed in the next section).

17 A statute of limitations is a “statute establishing a time limit for suing in a civil case, based on the date when the claim accrued (as when the injury occurred or was discovered). The purpose of such a statute is to require diligent prosecuting of known claims, thereby providing finality and predictability in legal affairs and ensuring that claims will be resolved while evidence is reasonable, available, and fresh.” Black’s Law Dictionary, 8th ed., Bryan A. Garner (St Paul, MN: West Publishing, Thomson Business, 2004), 1450–1451.
18 Black’s Law Dictionary, 1336–1337.
19 James and Hazard, Civil, 532.
20 James and Hazard, Civil Procedure, 532.
21 In some circumstances courts have crafted rules to permit toxic tort issues to be reopened if a disease actually develops (after permitting recovering under causes of action short of actual harm); see Hagerty v. Ldl Marine Services, Inc, 788 F 2d 315 (1986).
Tensions between Scientific and Legal Epistemic Practices

In the tort law, evidentiary procedures with which adversaries must comply might be considered part of the requirements that collectively aim at not exceeding a tolerable balance of mistakes — of legal false positives/false negatives — between parties and at achieving other nonepistemic institutional goals, such as serving justice, providing fair procedures to litigants, and the like. Although tort law does not endorse casual rejection of the status quo, the standards of proof embedded in its burdens of persuasion do not appear as demanding as scientific standards of proof, and there is a different concern with mistakes that might result. 24

For example, there is a normative evenhandedness between litigants that is central to the tort law as revealed by the ultimate standard of proof to which plaintiffs must persuade a jury, the preponderance of evidence standard. This is a long-standing norm from the traditions of civil law and the Supreme Court has endorsed it. 25 Unlike the criminal law, which embodies procedural rules that tend to prevent the wrongful conviction of innocent persons, the tort law more equally balances the concerns of avoiding mistakenly holding defendants accountable and mistakenly denying plaintiffs recovery. There is substantial legal history supporting this view. In Speiser v. Randall, Justice Brennan noted:

There is always in litigation a margin of error, representing error in factfinding, which both parties must take into account. Where one party has at stake an interest of transcending value — as a criminal defendant his liberty — this margin of error is reduced as to him by the process of placing on the other party the burden of... persuading the factfinder at the conclusion of the trial of his guilt beyond a reasonable doubt. 26

Justice Harlan developed this theme in a concurring opinion in In re Winship:

The standard of proof influences the relative frequency of these two types of erroneous outcomes. If, for example, the standard of proof for a criminal trial were a preponderance of the evidence rather than proof beyond a reasonable doubt, there would be a smaller risk of factual errors that result in freeing guilty persons, but a far greater risk of factual errors that result in convicting the innocent. Because the standard of proof affects the comparative frequency of these two types of erroneous outcomes, the choice of the standard to be applied in a particular kind of litigation should, in a rational world, reflect an assessment of the comparative social disutility of each. 27

Note that the two types of erroneous outcomes possible are a factual outcome that favors the plaintiff when the facts warrant an outcome for the defendant or an “erroneous factual determination” for the defendant when a correct understanding justifies a judgment for the plaintiff. Justice Harlan then discussed the preponderance of the evidence standard:

In a civil suit between two private parties for money damages, for example, we view it as no more serious in general for there to be an erroneous verdict in the defendant’s favor than for there to be an erroneous verdict in the plaintiff’s favor. A preponderance of the evidence standard therefore seems peculiarly appropriate for, as explained most sensibly, it simply requires the trier of fact to “believe that the existence of a fact is more probable than its nonexistence...” 28

Moreover, recently the Supreme Court in Santosky v. Kramer, adopted the standard set forth in Addington that “in any given proceeding, the... standard of proof... reflects not only the weight of the private and the public interests affected, but also a societal judgment about how the risk of error should be distributed between the litigants.” 29 It then added that the preponderance of the evidence requires litigants to “share the risk of error in roughly equal fashion.” 30

The same standard of proof is utilized in the vaccine “off table” injury cases and is articulated as follows:

Petitioners must only demonstrate more probably than not (50% and a feather) that the vaccine can and did cause the injury alleged. 31

Thus, the tort law, based on the “more likely than not” standard of proof, is “indifferent as between a plaintiff’s erroneous recovery [a legal false positive] and a defendant’s erroneous non-liability [a legal false negative],” reflecting important non-epistemic values, such as the risk of injustice between parties. 32 The distribution of the risk of mistakes just reviewed rests on an interpretation of the ultimate standard of proof in the tort law that must be established to a jury’s satisfaction.

It is barely possible that admissibility decisions might be seen as embracing a different standard for risks of mistakes, but that would be odd. Instead, a reasonable presumption in the law appears to be that judges should be at least

31 Stevens v. the Secretary of Health and Human Services, 2001 at 32.
32 Santosky v. Kramer, 455 U.S. at 755 (1982) (adopting the standard set forth in Addington v. Texas, 441 U.S. at 423 (1979); Green, “Expert Witnesses,” 687. Legal and factual false positives (or false negatives) are not identical, of course. However, when there are a few pieces of evidence that are critical to a litigant’s case, a factual error will more likely translate into a legal mistake.
as evenhanded in admitting and excluding experts as the civil law is in its ultimate adjudication of civil litigation. The argument for this point would be that although both the tort law and the criminal law have different standards of proof to which a jury must be persuaded, judges in both cases are presumed to be equally evenhanded in admitting and excluding evidence in the two areas of law. That is, the legal risk of mistakes in admissibility does not follow the risk of mistakes as represented by the ultimate standard of proof and the only reasonable presumption is that judges must be equally evenhanded toward each side.\textsuperscript{33} Thus, it seems that the litigants should share the “risk of error” in admissibility “in roughly equal fashion.”\textsuperscript{34} Judges, hewing to this standard, must exercise care in treating scientific evidence evenhandedly in admissibility decisions because of the mistake norms of the law.

Science has different commitments that can pose tensions and even undesirable consequences for the law. The goals of understanding phenomena and adding carefully to the knowledge status quo have led the scientific community to adopt selective evidentiary procedures in order to ensure that new (typically theoretical) knowledge is securely added to the field. As I argued in Chapters 4 and 5, in general, scientists are typically quite demanding in preventing factual false positives, that is, their procedures are designed to minimize study results that show that a substance has a toxic property when in fact it does not. At the same time, scientists seem to have a lesser concern to prevent false negatives. Although this is appropriate for theoretical scientific research, it risks problems in other institutional contexts, such as, in public health institutions or in either the regulatory or the tort law that have different mistake norms.

Scientific epistemology is blind to the law’s evenhanded norms; it is not indifferent between preventing these two different kinds of factual mistakes. Thus, courts need to be sensitive to these differences when they make admissibility rulings — erring in ways that neither favor plaintiffs nor defendants as they review expert testimony and its foundation. Judges will need to exercise special care when reviewing expert testimony resting on scientific studies that, because of the epistemic norms implicit in science, do not protect as well against factual false negatives as against factual false positives. Such errors can have substantial effects on the legal interests of the litigants. If they fail to recognize such issues they risk being unfair to litigants.

To further illustrate this, the special magistrates in vaccine-injury cases are sensitive to this issue insofar as it pertains to the preponderance of the evidence burden of proof, as Judge Golkiewicz puts it in his Stevens’ opinion:

\textsuperscript{35} Another possibility is that the review of evidence might follow the stringency of the standards of proof. This would increase the screening of the government’s evidence in criminal cases because of the high burden of proof it faces, but the screening of evidence in civil litigation would continue to be evenhanded because its ultimate standard of proof is as well.

\textsuperscript{34} Santosky v. Kramer, 455 U.S. at 755.

Of course, the court’s reliance on the [Institute of Medicine’s scientific] committees’ reports in no way suggests that a petitioner must demonstrate causality under the same strict scientific principles employed by the panel members. Petitioners [analogous to plaintiffs in tort suits] must only demonstrate more probably than not (50% and a feather) that the vaccine can and did cause the injury alleged; petitioners need not prove their case to a scientific certainty.\textsuperscript{35}

However, there is a further point to which we will return later. In ruling on the admissibility of expert testimony that rests on a few individual studies, courts must exercise care so that the scientific concern to prevent falsely positive results does not greatly increase the possibility of falsely negative results of studies to the detriment of plaintiffs. As we saw in Chapter 4, this can be a particularly vexing issue when statistical studies are too small to equally reduce random mistakes. This example is particularly dramatic and mathematically certain, but it only serves to highlight a much more general problem that judges can face in admissibility rulings. They also must exercise caution not to fall prey to insisting on high standards of proof from science and demand those for the law; this, too, would distort legal norms.

**Critical Stresses**

Which of the tensions just reviewed might pose the greatest problems for the law-science interaction? One obvious difference between the two institutions is in the time frames of scientific inquiry and the legal resolution of disputes, respectively. There are several time-sensitive issues that can make a difference.

1) Typically, scientific curiosity must have been pricked to lead scientists to pursue a problem. Scientific inquiry tends to be driven by curiosity and question answering, whereas the law imposes various time-constraints on bringing legal cases within statutes of limitation. A scientist must have a research problem in which she is interested and have the funding and other resources to pursue it before the research will be done.\textsuperscript{36}

Any research needed in legal settings may or may not be available within the limitations imposed by the law. If research has not been conducted by the time injuries have been noticed and the toxicity of substances suspected, it may be difficult to provide results to assist the legal issues. This occurred in the toxic tort suit described in *A Civil Action* shortly after the case was settled, the U.S. EPA issued a report indicating that trichloroethylene did reach the water wells of Woburn, Massachusetts, which plaintiffs had alleged caused leukemia.

\textsuperscript{35} Stevens v. the Secretary of Health and Human Services, at 32.

\textsuperscript{36} Of course, scientists might well conduct research on more practical problems not primarily driven by curiosity, provided sufficient funding were available.
in several children. The author, Jonathan Harr, notes, "On the face of it, the verdict appeared to stand for an example of how the adversary process and the rules and rituals of the courtroom can obscure reality."  

2) Research itself takes time. Even if investigating the toxicity of a new substance is on a scientist’s research agenda, research may not be quick to conduct. A reasonable battery of animal testing to provide a variety of data points about the toxicity of a substance can take up to six years to complete. Epidemiological research also can take time, depending on the availability of data.

3) Biological systems impose their own constraints. For example, epidemiologists must study exposed individuals for a sufficient duration to ensure that both the induction and latency periods have elapsed so that the studies are not falsely negative because researchers have violated these fundamentals. This is particularly a problem for diseases such as cancer that have long latency periods. In general nature does not always reveal her secrets quickly. Courts must be sensitive to such issues to ensure fairness between litigants.

4) Once a legal case has been adjudicated, it cannot be reopened even though scientists might have uncovered compelling new evidence that could change the outcome. The generic issue recently received the attention of the U.S. Supreme Court in *Dow Chemical Company, Monsanto Company, et al., v. Stephenson and Isaacson.* Although earlier a district court had overseen a settlement between Vietnam veterans who alleged they had contracted a variety of diseases from exposure to Agent Orange, new scientific evidence has now strongly suggested that a broader class of individuals had contracted disease than were part of the earlier settlement. Will the veterans who are alleging new kinds of diseases be permitted to pursue their legal claims?

A second difference between science and the law is that there are funding constraints on scientific inquiry that influence the development of a research agenda. Even if a scientist suspects that a substance adversely affects humans, because research is expensive, she must have sufficient funding to carry out the studies. She must find a funding source and needed personnel. This is not always easy. Some problems are not of interest to national funding agencies. Some problems might be of interest to the firm that created and distributed the substance, but, as we have seen, testing can often only invite legal problems, so this source of funding may be closed. Victims likely do not have resources for


38 Similar concerns apply to animal studies, although because animals have shorter lifetimes and faster metabolism research can be done more quickly on them than on humans.


the funding and plaintiffs’ attorneys’ funding will be criticized as suspect (the same critique should apply to defendants).

A third and most significant feature that distinguishes between the institutions is their respective standards of proof. We have discussed these differences earlier and suggested why judges should exercise great care not to substitute one for the other.

A fourth difference is in their approach to the complexity of the subjects before them. Scientists recognize the need to be sensitive to the subtlety and complexity of their subject (hence some of their hedging), whereas courts need to make the law more or less accessible to the citizenry and tend to simplify.

A fifth and related difference concerns approaches to uncertainty. Uncertainty is endemic to science; indeed, part of the integrity of scientific claims is that researchers exercise considerable care in saying what they do and do not know about the subject under consideration (or they assign probability judgments to claims). Moreover, they usually do not need to make action-guiding decisions based on their uncertain knowledge (other than how to pursue research further). They acknowledge and are quite comfortable with varying degrees of certainty and uncertainty about their subject. By contrast, although the law recognizes and has decision rules for addressing uncertainty (e.g., legal presumptions as well as burdens and standards of proof), if there is too much uncertainty for a particular decision rule, the party with the burden of proof on the issue loses (recall Chapter 2). In a trial context, courts do not have the luxury of avoiding a decision in the face of uncertainty. Moreover, it is usually forced into a binary choice— for one party or the other—despite uncertainty. The mere presence of uncertainty in the law may make it easier to decide against the party with the burden of proof. And, it leads to deliberate strategies to emphasize or exaggerate the uncertainties facing that party.

A sixth critical distinction is in the distributive concerns between the two institutions: (1) differences in standards of proof can also produce substantive tensions between law and science on distributive issues because the two institutions have different approaches to the distribution of mistakes. Thus, judges must ensure that scientific inattention to false negatives does not distort the legally mandated equal concern with the risk of errors between plaintiffs and defendants in admissibility decisions or in the ultimate outcome of a trial; and (2) human toxicity studies may have been conducted independently of legal concerns, thus ignoring issues of biological variability that can arise between different subpopulations of exposed persons: infants, children, the elderly, women, pregnant women, and diseased individuals. In Chapter 7, I consider specific examples of this issue.

40 See Michaels and Monforton, "Manufacturing Uncertainty," S39-S48, for a discussion of the use of this strategy going back to the 1950s.
JUDICIAL RESPONSES TO THE SCIENCE-LAW INTERACTION

The tensions just described create a context in which courts can understandably, regrettably err. At the same time they can be exacerbated or ameliorated by judicial reviews. On the one hand, for example, courts’ failure to recognize the less than textbook evidence that is often available in legal venues can lead them to impose admissibility requirements that are simply too stringent for the practical decision making of the law. This might well put at risk the possibility of justice between parties. In addition, a tendency for simple and restrictive rules may deny parties some of the more complex evidence they might need in support of their cases. This would undermine one of the main goals of the Daubert trilogy – having legal decisions better supported by the pertinent scientific evidence.

On the other hand, how judges review evidence can aggravate or reduce tensions between the two institutions. If courts adopt some of the more extreme views toward the prevention of false positives from the scientific community or require numerous kinds of evidence (more than needed) for expert testimony, they can exacerbate tensions. By failing to observe the “mistake norms” of the law and adopting some of the more extreme mistake norms of science, courts can upset the legally mandated balance of mistakes between plaintiffs and defendants. This would end plaintiffs’ cases for legally and scientifically mistaken reasons. This could invite criticism from the scientific community as courts impose conceptions of scientific evidence and reasoning in opposition to scientific norms. By contrast, when courts are too lenient toward experts and scientific evidence, this opens the possibility that courts will permit scientific evidence and testimony that would not be within the boundaries of acceptable scientific disagreement, invites criticisms from the scientific community from a different direction, and creates opportunities for final court decisions to be at odds with existing scientific evidence. Thus, there are risks of mistakes on both sides of admissibility decisions.

However, there can be an asymmetry from the scientific community’s responses to admissibility errors. On one side, at the time of the Bendectin decision some courts had admitted expert testimony at odds with the prevailing science about the substance in question; this in turn might have invited and probably did invite a wider critical response from the scientific community. Legal decisions at odds with the known scientific evidence (if it is indeed known) creates a target that is easy to publicize. On the other side, when judges exclude evidence that should be admitted, this is comparatively invisible – resulting from an admissibility hearing, not a public trial – and probably receives little or no attention. These decisions do not become such an inviting target simply

because they are below the public radar screen. Pressures from the scientific community that might serve as correctives to judicial admissibility mistakes tend to be asymmetrical: more strongly arrayed against decisions that mistakenly admit flawed scientific evidence, but likely much less or nonexistent against decisions that mistakenly exclude good evidence and reasoning. This may have begun to change with some papers by legal scholars and scientists in special issues of journals or books that are about to appear concerning the disregarding of reasonable scientific evidence.42

I do not argue that the generic tensions between science and law have created the problems about to be considered, but they set the stage for mistakes and invite them. Judges, like the rest of us, can become trapped by our own context or by idealized views of a tool we need to use, in this case science. The earlier section sought to highlight some differences between science and the law, so that those who must work with both can better see how to use them more compatibly.

The remainder of the chapter conducts a philosophic analysis of some reasons courts have given for excluding evidence in existing cases.43 The cases considered merit review for several reasons. First, they provide some windows into the effects of the doctrinal changes instituted by the Supreme Court.44 Second, they provide some, but not fully comprehensive, information about whether the Court’s reform is making law and science more or less compatible. Fragmentary evidence to date indicates that although it may be that courts are excluding some scientific experts whose reasoning and methodology exceed the boundaries of respectable scientific disagreement in their testimony, other courts are overreaching and excluding experts whose testimony appears to be within the boundaries of respectable scientific disagreement – they are


43 The analysis is philosophical in the sense that I consider various reasons judges have given for admitting, or more commonly, excluding evidence. To what extent are these reasons consistent with those scientists would utilize? The scientific comparison is based on reasons that consensus scientific committees would provide for their conclusions.

44 It is, of course, difficult to have a comprehensive view of this issue because evidence of mistaken judicial admissibility decisions is asymmetrical – when evidence is excluded which precludes a plaintiff’s case from going forward to trial, the judge must write an opinion and it is subject to appeal (but whether or not an appellate court in fact reviews it is another matter). When evidence is admitted there is usually little or no record of the judicial reasoning, unless that decision is appealed.

excluding experts for reasons that are at odds with good scientific practice and whose testimony should be heard by juries. Third, how would such decisions affect the balance of interests between litigants in toxic tort cases?45

THE RISK OF SIMPLIFIED ADMISSIBILITY RULES

The Daubert opinion stressed the need for a flexible set of criteria to determine the admissibility of scientific evidence. Nevertheless, that decision left open the possibility for the use of overly simple admissibility rules because judges have great latitude in reviewing scientific evidence without appellate review. Many of the cases presented below suggest that courts often utilize relatively simple rules as guides for reviewing expert testimony and its foundation. Because of the evidentiary complexity in science, there is a risk that it may overwhelm judges who, as a result of their education, may not be prepared for the subtle and difficult task of evaluating and weighing the various kinds of evidence and scientific reasoning for the context in question. Simple rules may be endemic to and valuable in the law as Schuck and Walker suggest. And they are tempting in order to ease the task of court administration. However, to remain faithful to the letter and spirit of Daubert, judges should avoid using such guides. Their reviews should be as subtle and sensitive as the evidence demands. This will ensure that their reviews remain faithful to the science and better ensure fairness to litigants.

SPECIFIC CONCERNS FROM COURT DECISIONS

Some courts have demanded that experts base their testimony on particular kinds of evidence, even though scientists would not insist upon such evidence in order to make a toxicity judgment about a substance.46 Other courts exclude testimony based on evidence that scientists routinely rely upon to draw inferences. When either occurs, it will frustrate the laudable aim of the law utilizing the pertinent science. Moreover, it is a reasonable conjecture that there is no minimum kind or amount of evidence for a judgment that substance S causes or contributes to an injury to humans and there is no required or privileged explanation that S causes or contributes to a disease. There are various explanatory paths to such conclusions and different kinds and patterns of evidence that can be assembled to come to such judgments, several of which I will present in Chapter 7.47 If courts exclude evidence or expert testimony needed to show that a particular explanation is plausible, they may make an explanatory mistake — litigants utilizing such an inference are not permitted to utilize certain arguments to argue for an explanation of the harm.

If courts err, their mistakes can also have an abiding legal impact because they appear as a written opinion. Appellate decisions serve as precedents for the court in question and for lower courts in the same jurisdiction. Appellate or district court opinions can function as supportive reasons, but not as precedents, if the court is at the same level in a judicial hierarchy, but in a different circuit or jurisdiction (thus, the reason cannot act as an explicit precedent). “Following the Joneses” of other jurisdictions, however, can perpetuate mistaken evidentiary understandings through the legal system.48

Demands for Particular Kinds of Evidence

Ideal Evidence Is the Enemy of the Good

Shortly after Daubert was decided, a few courts insisted that experts testifying must base their testimony on what might be considered ideal evidence. Wade-Greaux v. Whitehall Labs illustrates this issue.49 The plaintiffs argued that a mother’s use of Primatene Tablets and Primatene Mist, over-the-counter asthma medications sold by the defendant “caused TiaNicole Wade-Greaux to be born with true malformation of her upper limbs and other skeletal defects.”50 The trial court held that plaintiffs, in order to have their scientific evidence admitted, had to show that their claims about causation were supported by “repeated, consistent epidemiological studies; ... an animal model that duplicates the defects resulting in the human from the exposure; ... a dose/response relationship between the exposure and the effect on the experimental fetus; and ... the mechanism of teratogenicity of the agent should be understood and make biologic sense.”51 As we discuss later, most of the court’s necessary conditions are scientifically problematic. A court’s requiring that all of these conditions be satisfied for admissibility presents even greater difficulties.

48 For example, in Case v. Ohio Medical Products, Inc., 877 F. Supp. 1380 (N.D. Cal. 1995), the district court judge did not permit plaintiff to rely upon case studies and this court’s reasons have been perpetuated and perpetuated throughout the legal system (but not a matter of legal precedent), very likely resulting in mistakes.
Although this decision is not a leading one, and by now is probably seen as an outlier (but the trial court was he[...] the reliability of an advocate's evidence and reasoning, not its probabilistic value compared with the defendant's. This should be considered in the context of a directed verdict.

Second and more seriously, the court seems to take a literal textbook approach to admitting scientific evidence. That is, because standard methodology references suggest that there must be epidemiological, animal, and other evidence in support of the claim of causation, courts should require all of this evidence before plaintiff's experts can testify in a trial. Such multiple sources of evidence may be the best and ensure the most certain route to a correct scientific judgment.

Summary judgment was granted because the plaintiff's evidence fell far short of the court's articulated criteria, and the decision was upheld without comment on appeal. See Wade-Greaux v. Whitehall Lab., Inc., 874 F. Supp. at 1476–1486, and Wade-Greaux v. Whitehall Lab., Inc., 46 F.3d 120 (3d Cir. 1994) (unpublished table decision). Nonetheless, the announced criteria seem particularly problematic.

See Daubert, 509 U.S. at 595. ("The focus, of course, must be solely on principles and methodology, not on the conclusions that they generate.") The court listed various factors for determining whether an agent is a human teratogen, noting that those factors (which included epidemiological evidence, animal models, and a demonstrated dose-response relationship) were generally accepted by the community of teratologists, taught in medical schools throughout the country, and included in highly esteemed treatises on teratology. See Wade-Greaux v. Whitehall Lab., Inc., 874 F. Supp. at 1450–51.

In its findings of fact on epidemiology, the court stated:

Absent consistent, repeated human epidemiological studies showing a statistically significant increased risk of particular birth defects associated with exposure to a specific agent, the community of teratologists does not conclude that the agent is a human teratogen.

...[These] statistical findings are, standing alone, insufficient to permit a conclusion that a particular agent is teratogenic because the scientific community also requires confirmatory evidence from experimental animal studies.

Wade-Greaux. 874 F. Supp. at 1453 (citations omitted). 'Toxicologist Arthur Furst appears to adopt a similar view (see Arthur Furst, "Yes, But Is It a Human Carcinogen?" Journal of the American College of Toxicology 9 (1990): 12), but it is more appropriate for the context within which he works. He was addressing a society of toxicologists and asking what criteria a scientist should require to be satisfied before being certain on substantive scientific grounds that a substance is a carcinogen. Courts assessing the admissibility of scientific evidence are operating in a much different context with different evidentiary rules and different guidance from the Daubert decision.'

This approach, however, poses several difficulties. It is overly restrictive for what constitutes an appropriate explanation of plaintiff's causation claims. The court appears to require that certain categories of evidence must be present, even though not all of this evidence would be required for a scientist or consensus scientific body to come to a reasonable scientific conclusion that substances are toxic to humans.

Moreover, such demanding evidentiary standards should not be needed to survive an admissibility review. A court adjudicating a tort claim need not be persuaded to a scientific certainty that a substance is a teratogen. Such certainty would substitute the most demanding scientific standards of proof for tort law admissibility. Thus, the Wade-Greaux court appears to have required a more constraining evidentiary explanation than is necessary, and demanded the best or most certain evidence when reasonable evidence is the most that should be needed to survive the reliability requirement for admissibility.

Daubert only requires that for testimony to be admissible it must be more likely than not reliable and fit the facts of the case. The more specific amended Rule 702 requires that testimony must be "based on sufficient facts of data," the testimony must be the "products of reliable principles and methods" and "reliably applied to the facts of the case." Courts should assess a litigants' evidence as a whole for reliability, but it need not possess the highest degree of certainty, be the best evidence or even be more likely than not correct. The evidence utilized by and conclusions of consensus scientific committees serve as counterexamples to the Wade-Greaux court's demands (Chapter 7). Moreover, often scientists do not have such evidentiary luxury (except perhaps in pure research for its own sake). In many circumstances, they can come to correct conclusions with evidence well short of this ideal.

Such requirements are additionally odd if we recall that a decision on admissibility is a preliminary review of whether an expert's opinion rests on a sufficiently reliable basis to support the factual basis of plaintiff's cause of action and assist a jury. In Wade-Greaux the ideal became the enemy of the good and possibly of the admissible; courts need to guard against this.

59 See Daubert v. Merrell Dow Pharm., Inc., 509 U.S. 579, 589. ("Of course, it would be unreasonable to conclude that the subject of scientific testimony must be "known" to a certainty; arguably, there are no certainties in science."") See also the more recent Advisory Committee on Evidence Rules proposed Amendment: Rule 702, Committee Note: "At the court stated in In re Paoli R.R. Yard PCB Litigation, 35 F.3d 717, 744 (3d Cir. 1994), proponents "do not have to demonstrate to the judge by a preponderance of the evidence that the assessments of their experts are correct, they only have to demonstrate by a preponderance of evidence that their opinions are reliable. . . . The evidentiary requirement of reliability is lower than the merit's standard of correctness.")
Demands for Human Epidemiological Evidence

Other courts have approached the Wade-Greaux mistake, but the error is subtler. These courts have demanded that expert testimony must be supported by human epidemiological studies. This was particularly the case with decisions made prior to and immediately following Daubert, but some courts continue to demand human data as a necessary foundation for expert testimony. For example, in the leading Agent Orange opinion, Judge Weinstein stated that “[a] number of sound epidemiological studies have been conducted on the health effects of exposure to Agent Orange. These are the only useful studies having any bearing on causation” (emphasis added). Similarly, in Lynch v. Merrell-National Laboratories, Division of Richardson-Merrell, Inc., the First Circuit Court of Appeals noted that nonepidemiological studies used “singly or in combination, do not have the capability of proving causation in human beings in the absence of any confirmatory epidemiological data.” Other courts hearing Bendectin cases came to similar conclusions. Courts hearing toxic tort cases involving other substances have concurred as well. Moreover, the influence of an epidemiological threshold continues to the present with courts often citing the Brock Bendectin case. Even when epidemiological studies are not explicitly a necessary condition, the context may suggest “that epidemiological evidence is a necessary prerequisite for a plaintiff to prevail.”

However, it is simply a mistake to think that epidemiological studies are necessary for scientists to form reasonable views about toxic effects in humans.

63 Although this statement is ambiguous as to whether it is a claim about the particular case or a more general criterion for admissibility, a number of courts appear to have taken his remarks as announcing a general criterion.
64 Lynch v. Merrell-National Lab., Div. of Richardson-Merrell, Inc., 830 F.2d 1190 (1st Cir. 1987).
66 See Brock v. Merrell Dow Pharm., Inc., 874 F.2d 307, 312 (5th Cir. 1989) (the court concluded that a Bendectin plaintiff must prove a statistically significant study before satisfying her burden of proof on causation); cert. denied, 494 U.S. 1046 (1990); Richardson v. Richardson-Merrell, Inc., 857 F.2d 823, 825, 831 n.59 (D.C. Cir. 1988), cert. denied, 493 U.S. 882 (1989) (noting that “epidemiological studies are of crucial significance”).
68 See Chambers v. Exxon, 81 F. Supp. 2d 661, at 664. (“Epidemiological studies are necessary to determine the cause and effect relationship between an agent, in this case exposure to benzene, and a disease, CML [chronic myelogenous leukemia]” (citing Brock v. Merrell Dow Pharmaceuticals, Inc., 874 F.2d 307, 311 (5th Cir. 1989), modified by 874 F.2d 307, 311 (5th Cir. 1989).)

Because of limited evidence, consensus scientific bodies in fact frequently utilize various kinds of nonepidemiological evidence in combination to conclude that a substance is a known or probable carcinogen to humans. Some courts have resisted the impulse to enshrine epidemiological studies as necessary before an expert can testify to causation and there appears to be a more recent and salutary trend away from an epidemiological threshold.

The attraction of courts to epidemiological studies is understandable, as we saw in Chapter 4. Epidemiological studies are the most direct evidence of human harm. Well-designed and well-done epidemiological studies (much is built into these ideas), with sufficiently large samples that are sufficiently sensitive to detect the adverse effects in question and with a sufficient duration can greatly assist in identifying toxic effects of product. Unfortunately, such ideal studies appear to be the exception rather than the rule. Other kinds of human data, such as clinical trials (even rarer), if there are any, well-substantiated human case studies, and molecular and toxicity data from humans, also can be important evidence of a substance’s human toxicity. And there are other kinds of scientifically relevant evidence that support judgments of a substance’s human toxicity.

There are both scientific and legal reasons against requiring epidemiological studies as the foundation for expert testimony in a toxic tort case. There are various scientific problems, limitations, shortcomings, and weaknesses that affect their usefulness, especially in toxic tort suits. For many substances, epidemiological data simply are not available. Recall from Chapter 4 that there is often too little human experience with a toxicant to fully establish its toxicity, exposures tend to be crude, and epidemiological studies are notoriously insensitive. For example, there is good epidemiological evidence for only about half or less of the known or likely human carcinogens assessed by national or international scientific bodies. How widespread a problem this is for other toxicants is difficult to know. Special Magistrates in the Vaccine...
Injury Compensation Program have judicially noticed this point. Because it is morally wrong to deliberately expose people to a toxicant for study purposes, any studies must be conducted when accidental exposures occur. However, this makes it difficult to have accurate exposure information and accurate results.

Even when epidemiological studies can be conducted, they tend to be expensive. And, if studies are too small or of too short duration, they may fail to detect adverse effects even if the substance in fact causes or contributes to them (Chapter 4).

Surprisingly also, human studies that are conducted too long after exposure can also result in underestimations of risks. For example, studies conducted on retired workers who were exposed to benzene during their working years show a lower relative risk than would be expected in such circumstances. Finally, it is a reasonable conjecture that for many exposures typical of toxic tort suits, conducting an appropriate epidemiological study is simply not an option, for example, too few may be exposed, exposure data is not likely to be present, the diseases may be rare, or the studies will be quite expensive.

In addition, because epidemiological studies are not controlled experiments, their reputation in the scientific community may not be nearly as strong as the standing they appear to have with federal judges, especially those who have said that human epidemiological evidence is the only or the best kind of evidence in toxic tort suits. An epidemiologist at the National Cancer Institute describes their relative strength as follows:

There is a perception in the scientific community that epidemiological evidence, observational in nature and prone to confounding and bias, is weaker evidence than animal model studies and other types of laboratory-based toxicological studies. The perception is decades old.

This perception is likely to be mitigated by the recent development and successes of molecular epidemiology and multidisciplinary programs in disease prevention that develop and use evidence from many different scientific disciplines.

Epidemiological studies are, however, just one kind of evidence; there are others (Chapter 4). Moreover, an insistence on epidemiological evidence obviously privileges “holistic” human body evidence over other quite good evidence on which scientists rely.

There are legal problems with requiring epidemiological evidence. To insist that plaintiffs, for example, alleging that exposure to a substance such as ethylene oxide caused brain cancer, must have epidemiological studies on which to base testimony by scientific experts is to require a study that is extremely difficult to do. Causes of rare diseases, of which brain cancer is one, can be quite difficult to detect without a large, expensive study that is sufficiently sensitive to detect the effect. In most cases, there are not persons appropriately exposed to serve as the basis of a study.

Special Restrictions on Epidemiological Studies

Some courts and commentators have gone further and have required that the epidemiological studies must satisfy additional conditions before they can provide a foundation for expert testimony in toxic tort cases. Some have insisted that such studies be “statistically significant.” Others have insisted that the studies find a relative risk of at least two between the exposed and control populations. Still others have suggested that all or most of Hill’s factors must be satisfied by epidemiological studies before they should be admitted. Such considerations help interpret studies. However, if they become necessary conditions on scientific evidence that forms the foundation of expert testimony, this poses problems.

**Statistical Significance Requirements:** Both before and after the Daubert decision, several courts and numerous commentators have insisted that epidemiological studies must be “statistically significant,” a not unreasonable, but possibly misleading, requirement. This requirement means that studies must

---

75 Stevens v. Secretary of the Department of Health and Human Services, at 14. (“Unfortunately, ... epidemiology and [toxicological] footprints are rarely available — such (Unfortunatly, ... epidemiology and [toxicological] footprints are rarely available — such


78 Douglas L. Weed, M.D., Ph.D., Dean, Education and Training, Chief, Office of Preventive Oncology; Director, Cancer Prevention Fellowship Program, Division of Cancer Prevention, National Cancer Institute, personal communication, January 2003.

79 Jasanoff, Science at the Bar, 125.


have less than some low probability, for example, typically less than .05, that a statistical association between exposure to a substance and a disease is not the result of random chance. Some judicial rulings suggest that if studies do not satisfy this condition, expert testimony based on them should be rejected as evidence in toxic tort cases. Thus, statistical significance is treated as something like a bright-line rule that epidemiological studies must satisfy before they can be part of the foundation for expert testimony. However, many scientists, although recognizing its importance, do not necessarily regard statistical significance as decisive in judging whether epidemiological evidence can contribute to causal judgments of harm.

A variety of considerations show demands for low statistical are problematic. First, although some courts suggest that studies should be statistically significant at the .05 level, scientists take a much more nuanced approach toward the exact level of statistical significance they utilize. Colleagues with whom I have collaborated would tolerate a wider range of statistical significance for interpreting studies, for example, higher than .05, but would understand properly what they do and do not show about the data. After all, statistical significance rules out only one possible explanation of a positive study result – chance distributions in the sample population under study. And there seems no necessity to adopt a law, for example, .05, value for litigation. If there is a .10 chance of a false positive, is this too much higher than a .05 chance? Yet courts will be pressured to utilize the lower number.

Second, statistical significance compared with confidence intervals was an issue between amici in the Daubert case, but recent discussions suggest that the cutting edge of the field seems to be moving away from tests of significance for two reasons. Tests of significance are a kind of decision rule, useful for certain purposes but not others. Moreover, tests of significance reveal less about the underlying data than other presentations of the evidence.

---

89 Rothman, Modern Epidemiology, 120–121.
be informative and there is nothing sacrosanct about a 95 percent confidence interval. 90

The important evidentiary point is, "What does the evidence show about the phenomena?" Thus, Rothman argues for not using any automatic procedure to decide the issue, but more informationally rich means of exhibiting the data. Judges should be similarly willing to tolerate more relaxed confidence intervals or rules concerning statistical significance (if they insist on that) because scientists do. (This will make their review tasks more difficult, but the Supreme Court has given them the job.)

Third, if scientific results are excluded merely because they are not statistically significant, decision makers risk excluding important evidence and the decision might result in "far greater inaccuracy." 91 Greater inaccuracy can result because demanding tests of significance asymmetrically prevent false positives, but permit more false negatives (Chapter 4). And, "[p]reemptorily reject all studies that are not statistically significant would be a cursory and foolish judgment, particularly if there are multiple studies tending to show a consistent effect." 92 Thus, court decisions might be more accurate on factual grounds if a wider range of epidemiological data were admitted. 93

In addition, Rothman points out, when tests of significance dominate the interpretation of epidemiological data, they also can be quite misleading in another way. For example,

[In a review of 71 clinical trials that reported no 'significant' difference (P > 0.05) between the compared treatments, Freiman et al. [1978] found that in the great majority of such trials the data either indicated or at least were consistent with a moderate or even reasonably strong effect of the new treatment. In all of these trials, the original investigators interpreted their data as indicative of no effect because the P-value was not 'statistically significant.' The misinterpretations arose because the investigators relied solely on 'significance' testing for their statistical analysis rather than on a more descriptive and informative analysis. On failing to reject the null hypothesis, the investigators in these 71 trials inappropriately 'accepted' the null hypothesis as correct, resulting in a probable type II error [false negative] for many of these so-called 'negative' studies. 94]

90 Rothman, Modern Epidemiology, 119–120.
91 Green, "Expert Witnesses," notes one reviewer who identified seventy-one epidemiologic studies that failed to satisfy statistical significance, but concluded that the "studies were consistent with a moderate or strong effect of the treatment under investigation" (685) (citing Jennie A. Freiman et al., "The Importance of Beta, the Type II Error and Sample Size in the Design and Interpretation of the Randomized Control Trial: Survey of 71 'Negative' Trials," New England Journal of Medicine 299 (1978): 690).
94 Rothman, Modern Epidemiology, 117–118 (emphasis added).

There are also policy reasons to be concerned about stringent statistical significance rules. Recall the inverse relationship between preventing false positives and preventing false negatives (Chapter 4). 95 Thus, one cannot, without using very large sample sizes, have very low false positives, very low false negatives, and a study that will detect comparatively small relative risks, for example, of around two or three. For studies with small samples trying to detect the causes of comparatively rare diseases, researchers will be unable to detect some outcomes of scientific interest and perhaps of social import.

Moreover, determining which factual error one should risk in legal decisions is a policy matter. 96 If courts adopt a legal concern strongly to protect against false positive mistakes they are committed to the view that in the law it is more important to protect against false positives than to protect against false negatives. Mathematically, this favors defendants and makes it more difficult for plaintiffs to provide probative evidence.

In addition, when judges demand low statistical significance, they are also implicitly insisting that one explanation for a positive test result must be ruled out with a very high degree of confidence (95 percent), namely, that the positive result was a statistical anomaly. Thus, an interpretive tool of scientific inquiry, and a seemingly neutral one, can have quite unintended asymmetric effects in another institutional context and adversely affect the social outcomes of those institutions. 97

Moreover, such a screening rule seems at odds with the evenhandedness with which torts treats plaintiffs and defendants as revealed by its ultimate burden of proof and reasonable admissibility procedures. 98 The law in embracing scientific mistake norms as expressed in rigid statistical significance rules undermines its own mistake norms. Such rules will systematically disadvantage the party seeking to establish a fact with statistical evidence, typically the plaintiffs. This consequence is even more worrisome if, as some suggest, juries and judges accept statistical evidence much less critically than other kinds of evidence. Statistical evidence is then given greater credibility, and likely imposing particular hardships on plaintiffs. 99 None of this is to suggest that judges should

96 The issue is somewhat more complex than this even. One must make tradeoffs between preventing false positives, preventing false negatives and being able to detect comparatively low relative risks. With samples that are smaller than an "ideal size" to ensure all three, something must give. See Chapter 4 and Cranor, Regulating Toxic Substances, 36–40. See also Green, "Expert Witnesses," 691–692 (providing an example showing that the chances of a false negative can easily be nearly 10 times the chances of a false positive) and Cranor, Regulating Toxic Substances, 71–78.
97 As we have seen, there are other mistakes that can be artifacts of studies, such as biases in the design of studies and confounders, but because these are somewhat less likely to result in asymmetric errors that can adversely affect the law, I do not review them.
99 See Green, "Expert Witnesses," 693.
accept all epidemiological studies. Rather, they must become better consumers of scientific evidence so that their admissibility decisions concerning scientific evidence and reasoning do not distort the law.

We would do well to heed both a scientist and a legal scholar on the issue of statistical evidence. Sir Austin Bradford Hill, whom we met earlier, posed some fundamental questions scientists should ask themselves regarding causation: “[Is there any other way of explaining the set of facts before us, is there any other answer equally, or more likely than cause and effect?” He then proceeded to sum up the views of many working scientists concerning statistical significance:

No formal tests of significance can answer those questions. Such tests can, and should, remind us of the effects that the play of chance can create, and they will instruct us in the likely magnitude of those effects. Beyond that they contribute nothing to the “proof” of our hypothesis.\(^\text{100}\)

Michael Green, an academic lawyer concerned about the inordinate influence of the Bendectin and Agent Orange cases, concludes a discussion of statistical significance as follows:

[The art of teasing out causal inferences in the absence of a mature epidemiologic record is far too complicated for courts seriously to review the methodologies and analyses involved. Making the ultimate causal inference requires an assessment not only of the quality of the epidemiology but the biological plausibility, based on what is understood about the mechanisms of toxicity. Indeed, one of the lessons of the Bendectin cases is that the courts are not truly engaging in greater scrutiny of experts’ opinions; rather, they are adopting a few relatively simple screening devices. ... Especially as the available universe of evidence gets thinner, inadmissibility decisions have significant risks.\(^\text{101}\)]

I concur. Rigid and low statistical significance rules perhaps simplify the job of screening epidemiological studies. They also risk ignoring salient scientific evidence, encourage less accurate and less nuanced decision making, systematically disadvantage plaintiffs, and, thus, risk upsetting the balance of interests between plaintiffs and defendants.

**Relative Risk Rules:** Still other courts have required that epidemiological studies must find a relative risk of at least two. Carruth and Goldstein found thirty-one cases in which relative risks greater than two (RR > 2) were discussed. In twenty-nine cases, courts addressed whether RR > 2 is a threshold for proof of causation, with twelve saying that RR > 2 “was required to support a reasonable inference of causation”; whereas fourteen indicated that it was not.\(^\text{102}\) In addition, twenty-one of the opinions discussed whether RR > 2 is a threshold for the admissibility of an expert opinion on causation, with ten saying that “RR > 2 is required” and eleven that it is not. Thus, nearly half of the opinions that discuss this issue require it for ultimate proof of causation and about one-third make it a threshold for the admissibility of expert testimony. For the discussion here, I focus on the use of RR > 2 in support of expert testimony for admissibility.

Before discussing this suggested admissibility constraint, what is attractive about it? A judicial rule that makes having a RR > 2 necessary for expert testimony seems to hope for a certain ideal. It seems to require an evidentiary foundation for testimony that is objective, directly pertinent to human harm, is free or at least freer from subjective scientific judgment by the experts than studies that provide more complex and less direct evidence of human harm, and finally, is one that seemingly carries the legal burden of proof on its face in the statistics of the study.

It is seemingly objective because it is a scientific study, not simply an expert’s opinion. It is directly pertinent to assessing human harm in a way that animal or some other kinds of evidence are not because any revealed adverse effects are on humans. It appears to limit the latitude of experts in their judgment about the probability of causation because scientific judgment superficially appears not to enter into reporting and using the results of such studies. Finally, it seems to provide a scientific answer to either the standard of proof that must be met for general or specific causation. It appears that the statistics alone indicates that a randomly chosen diseased person in the exposed group “more probably than not” had his or her disease caused by the substance, thus plausibly satisfying the ultimate tort law standard of proof on specific causation. Similar considerations might be invoked in an argument for general causation.

Many pre- and post-*Daubert* courts and a number of commentators have endorsed the idea that an epidemiological study must reveal a relative risk of at least two in order for testimony based on such evidence to be admissible.\(^\text{103}\) Prominent among these is the Ninth Circuit Court of Appeals. It utilized such

\(^{100}\) Hill, “The Environment and Disease,” 15. 19

\(^{101}\) Green, “Expert Witnesses,” 693–694.


\(^{103}\) See, e.g., Daubert v. Merrell Dow Pharm., Inc., 43 F.3d 1311, 1320–21 (9th Cir. 1995), on remand from 509 U.S. 579 (1993); see also In re Joint E. & S. Dist. Asbestos Litig., 758 F. Supp. 199, 203 (S.D.N.Y. 1991), aff’d, 52 F.3d 1124 (2d Cir. 1995), and the Supreme Court of Texas in Merrell Dow Pharmaceuticals, Inc. v. Havner, 953 S.W.2d 706 (1997). Carruth and Goldstein reported that twelve of twenty-nine cases discussing a relative risk greater than two argued that RR > 2 “was required to support a reasonable inference of causation,” whereas opinions in fourteen cases indicated that RR > 2 was not required for proof of causation (Carruth and Goldstein, “Relative Risk Greater than Two,” 200–201). Commentators who argued early on for requiring a relative risk even greater than two are Bert Black and David E. Lilienfeld, “Epidemiologic Proof in Toxic Tort Litigation,” Fordham Law Review 52 (1984): 732, 769. More recently, Joe G. Hollingsworth and Eric Lasker argued for the same
a consideration to decide the Daubert case on remand without returning it to the trial court of origin.

California tort law requires plaintiffs to show not merely that Bendectin increased the likelihood of injury, but that it more likely than not caused their injuries. . . . In terms of statistical proof, this means that plaintiffs must establish not just that their mothers' ingestion of Bendectin increased somewhat the likelihood of birth defects, but that it more than doubled it — only then can it be said that Bendectin is more likely than not the source of their injury. Because the background rate of limb reduction defects is one per thousand births, plaintiffs must show that among children of mothers who took Bendectin the incidence of such defects was more than two per thousand.104

The Texas Supreme Court concurred in Merrell Dow Pharmaceuticals, Inc. v. Havner:

The use of scientifically reliable epidemiological studies and the requirement of more than a doubling of the risk strikes a balance between the needs of our legal system and the limits of science.105

There are several problems with courts requiring relative risks greater than two for admissibility rulings. Some of the issues arise because of what one might consider pragmatic issues in a study, whereas others reflect fundamental methodological problems in epidemiology. Finally, they also pose legal problems.

Consider the pragmatic issues. First, as epidemiological studies are ordinarily conducted to identify public health problems from environmental or workplace exposure, once a causal relationship is suspected between exposure and disease, if preventive actions can be taken for those who continue to be exposed, they ordinarily will be. An epidemiological study is essentially a "snapshot in time,"106 but it is as early a snapshot as might be helpful for public health purposes. Consequently, for diseases with any significant latency period, there will likely be additional diseases that were not detected by an early study and thus, the early study will suffer from incomplete accrual of diseases.107 This is likely to underestimate the severity of the disease causation. If courts insist on relative risks of at least two in published studies as a foundation for expert testimony, they will exclude some evidence of disease effects in humans. Moreover, if it is in a study that did not allow for complete accrual of disease, they exclude possible evidence of a higher rate of disease.

Second, in many cases epidemiological studies are done in occupational settings. Those who are employed tend to be healthier than the general population, producing a "healthy worker effect." Thus, "comparing exposed workers to the general population tends to understate relative risk" because the general population has a greater variability and wider range of susceptibility to diseases than healthy workers.108 Courts must be sensitive to such effects and realize they have limited applications to children, the diseased, the elderly and so on.

Third, when a population at risk has been studied, if there is an elevated disease rate, there typically would be efforts to reduce exposures and, thus, disease. If a follow-up study is then conducted on populations for which remedial measures have been taken, disease rates will be lower, and lower relative risks will be likely. Thus, follow-up studies can easily underestimate the potential substances have for causing adverse effects. Courts should inquire about the context of the studies.

Fourth, a study typically reports an average relative risk from exposure to a toxic substance, which may disguise higher or lower relative risks, on the one hand, resulting from higher or lower exposures respectively, or, on the other hand, disguise more or less sensitive individuals. For example, higher exposures might result in relative risks of four, whereas lower exposures might reveal relative risks of only 1.7. The weighted average of the different relative risks might yield an overall average relative risk less than 2. As scientists become increasingly aware of sensitive subpopulations, for example, resulting from genetic or other susceptibilities,109 they may discover that average relative risks inadequately reveal real risks to sensitive subgroups. (The causal basis for this was considered in Chapter 5 under "weak and strong" causal contributions.)

The tort law clearly protects sensitive subgroups, those with "eggshell skulls,"110 but overly stringent admissibility rules risk frustrating this aim. Thus, admissibility rules that preclude studies with overall relative risks less than two might prevent the admissibility of studies which included relative risks greater than two for individuals exposed to higher levels of a toxic substance or relative risks greater than two for particularly vulnerable groups. The automatic exclusion of testimony based on such evidence unfairly disadvantage both biologically sensitive plaintiffs and those subject to greater exposures than the study average. Unless courts (or legislatures) make policy decisions to exclude such subpopulations from tort compensation, epidemiological studies

---

104 Carruth and Goldstein, "Relative Risk Greater than Two," 207.
105 Carruth and Goldstein, "Relative Risk Greater than Two," 207.
106 Carruth and Goldstein, "Relative Risk Greater than Two," 207.
107 Carruth and Goldstein, "Relative Risk Greater than Two," 207.
108 Carruth and Goldstein, "Relative Risk Greater than Two," 208 (noting that a 1.5 relative risk in an exposed worker population might well reflect a doubling of disease rate compared with the appropriate unexposed population).
110 See W. Page Keeton et al., Prosser and Keeton on Torts, 291–292.
with relative risks less than two can provide scientifically relevant evidence and should not automatically be excluded from trial (or experts excluded for using it as part of their testimony).

The preceding point is supported by a striking example from radiation epidemiology that suggests it is a serious mistake to exclude expert testimony based on epidemiological studies with relative risks less than two. Ionizing radiation has long been a known carcinogen. It causes cancer in many human organs. One study shows that atomic bomb survivors from Hiroshima and Nagasaki contracted leukemia and multiple myeloma, as well as cancer of the esophagus, stomach, colon, other segments of the digestive system, urinary tract, lung, lymph nodes, and a number of other sites.\(^{111}\)

Contrary to widespread belief,\(^{112}\) the radiation exposures for many in these study populations were relatively low. What is striking about these findings is that epidemiological studies show with (interestingly) 90 percent confidence that all malignant neoplasms taken together except leukemia have a relative risk of less than two. Individual cancers that have a relative risk less than two include stomach, other parts of the digestive system, lung, and some other sites. Only leukemia, multiple myeloma, urinary tract, and colon cancer have relative risks greater than two from radiation exposure.

Such findings would be problematic for courts that have ruled inadmissible testimony based on epidemiological studies with relative risks less than two. Because ionizing radiation is one of the best-known carcinogens and one that scientists are certain causes cancer, courts, faced with plaintiffs who have been exposed to radiation, must either rule inadmissible testimony based on epidemiological studies for all neoplasms for which there is not a relative risk of greater than or equal to two, or must admit the evidence and then engage in the sensitive and complex task of assessing and weighing testimony based on evidence that exposure to ionizing radiation caused the cancer in question. Clearly, given the substantial evidence and degree of certainty about this carcinogen, the latter course seems much more defensible on both scientific and legal grounds. Several courts have judicially recognized the carcinogenic potential of radiation.\(^{113}\) Moreover, courts should apply this lesson to studies of other toxins.

There is also a subtler theoretical point about relative risks of the two. Sander Greenland and Jamie Robbins in a series of papers have argued that it is difficult to determine how much disease to attribute to an exposure. Thus, it is difficult to utilize epidemiological studies as a very good guide to the increase in disease causation.\(^{114}\) For example, exposure to a toxicant might cause new cases of disease, that is, diseases in individuals who would not have contracted the disease at all, or it might accelerate the onset of disease in individuals who would have contracted the disease in any case, but not as early. Thus, exposure may not only be an on-off switch for the disease in question, that is, turning it on; it may also function as an accelerator of disease that would have occurred in any case, only at a later time in a person's life. Both contributions to disease should be attributed to the exposure. However, epidemiological studies ordinarily would only record the new cases of disease. This theoretical point constitutes an additional reason to be skeptical about treating relative risks of two as any kind of important cutoff for legal purposes. (How much practical import their point has is less clear since pragmatically it will be quite difficult to estimate the acceleration of disease in those who would have contracted it in the natural course of events.) Robbins and Greenland, thus, argue that utilizing epidemiological studies as evidence for attributing disease to exposure risks under-reporting disease rates and some other approach should be taken.\(^{115}\)

Moreover, studies showing a relative risk of two or greater may simply not be available, as the Stevens' judge notes for epidemiological data in vaccine injury cases.\(^{116}\) Judge Golickiewicz argues that the best evidence in vaccine cases are epidemiological studies or vaccine "footprint" evidence ("dispositive clinical or pathological markers").\(^{117}\)

Unfortunately, few petitioners are afforded this evidentiary luxury since epidemiology and footprints are rarely available—such is the nature of science. This lack of direct evidence leaves petitioners no other recourse than to corroborate their causation claim with circumstantial evidence.\(^{118}\)


\(^{115}\) Carruth and Goldstein, "Relative Risk Greater Than Two," 206 (also citing Greenland and Robins's work).

\(^{116}\) Stevens v. Secretary of the Department of Health and Human Services, at 14.

\(^{117}\) Stevens v. Secretary of the Department of Health and Human Services, at 14.

\(^{118}\) Stevens v. Secretary of the Department of Health and Human Services, at 14.
Finally, there is a much more serious ethical and policy issue. A demand for human epidemiological studies with relative risks greater than two commits the tort law to the view that other people must have suffered serious diseases or death from a similar exposure before a plaintiff at the bar even has a chance at a jury trial in an attempt to receive just compensation for injuries suffered. In short, a necessary condition of admissibility of expert testimony is death or serious diseases suffered by others in order for an expert to have an epidemiological study on which to rely. Moreover, although the tort law is not on the front line of protecting the public’s health, deterrence of harmful conduct or products is a substantial part of the justification of torts.

For these reasons, one should be skeptical on both scientific and legal grounds of legally requiring a relative risk of two for the admissibility or the sufficiency of epidemiological evidence. The American Law Institute in modifying the Restatement of Torts on tort law causation has now also come out against requiring RR > 2.119

Sample Size and Duration of Studies: Sample size and duration of epidemiological studies are topics, which, although not explicitly utilized by judges (or recommended by commentators) as restrictions on expert testimony based on epidemiological studies, nonetheless merit a brief caution because both can be shortcomings of a study that can adversely affect their legal value. Recall that epidemiological studies that are of too short a duration may fail to reveal an existing relative risk because the induction and latency periods of the disease is longer than the study. Similar problems attend studies that are based on too small a sample. Both could produce mistaken “no effect” results (Chapter 4). Such outcomes would be particularly likely for studies that sought to identify cancers, which tend to have much longer induction and latency periods than many diseases.120 A longer study would likely be more sensitive to lower risks.121

Finally, there are a further point related to duration of exposure. Studies that are conducted too long after exposure has ceased can also underestimate risks; the “observed relative risk may not be the maximum risk experienced by the cohort.”122 Benzene researchers have found that the relative risks from benzene exposure decrease the greater the time that has elapsed between the last exposure and the time of study. The data suggest that leukemia may have a relatively short latency period or that the most susceptible individuals “succumbed to the disease early, leaving a less susceptible population at risk for the majority of the follow-up period.”123 These findings show how critical the timing of a study can be. If studies conducted quite long after last exposure are utilized as the basis for estimating the risk of disease from contemporaneous exposure, this would substantially underestimate risks. Consequently, epidemiological studies simply cannot be taken at face value for what they show without an inquiry into their duration, context of study, or sample size.

Extrapolation from Women to Men and Middle-Aged to Old and Young Persons: Recently, a federal court case in Washington state produced a new allegation from defendants about why even an exhaustively peer-reviewed, well-conducted, statistically significant epidemiological study showing a relative risk greater than two for women between the ages of eighteen and forty-nine should not provide evidence for adverse effects of the same substance for either children, those older than the study group, or some groups of men.124 This study showed that women using phenylpropanolamine (PPA)-containing appetite suppressants had a 16.85 times greater risk of suffering a hemorrhagic stroke when compared with those not using it, and a 3.13 relative risk of hemorrhagic stroke for any use of PPA (also in cough or cold products). There were an insufficient number of men using these products to provide any data on increased risks. Defendants challenged plaintiffs’ evidence on a number of grounds, only a few of which I mention.

Defendants objected that extrapolating from the eighteen- to forty-nine-year-old age group was not good science.125 They argued that because those older than forty-nine tend to have more strokes than those under forty-nine, extrapolation to older individuals should not be permitted. Quoting a defense scientist, however, Judge Barbara Rothstein rejected this argument, noting, “There are no drugs I’m aware of that get safer the older you get.”126

Defendants also called attention to differences between the women studied and extrapolations to men and children. However, Judge Rothstein found “wide support” in the scientific and medical literature for extrapolations to children, anticipating that toxic effects will be “as great, if not greater in children.”127 And,
defendants "failed to introduce any evidence that plaintiffs' experts' reasoning was not scientifically valid."128

The issues just discussed about extrapolations are not problems with judicial rulings. Quite the contrary. We see defendants trying to introduce some skepticism about plaintiffs' excellent scientific studies and a knowledgeable judge properly rejecting it. This judicial opinion shows plausible sources for possible judicial mistakes concerning scientific evidence, namely, one party to the litigation suggests some skepticism about the science, which a less astute judge might not have noticed (although defense arguments in this case were implausible).

**Using "Hill's Factors" for Excluding Evidence:** Subsequent to Daubert courts have reviewed experts' "methodologies." One such methodology is an expert's use of Hill's considerations as a guide to inferring causation from association. However, some courts have merely noted that these are some considerations that experts use to guide their inferences,129 whereas other courts have made more rigid uses of it.130 Some commentators, like courts, have merely noted the helpfulness of Hill's considerations.131 Others have urged the courts in more rigid directions.132 They appear to suggest that failure to satisfy some or many

---


129 See, for example, *Amoroso v. National Railroad Passenger Corporation*, 137 F. Supp. 2d 147 (E.D. N.Y. 2001) ("Epidemiologists generally look to several additional criteria to determine whether a statistical association is indeed cause. These criteria are sometimes referred to as the Bradford Hill criteria . . . .") and *Magistrini v. One Hour MARTINIZING DRY CLEANING*, 180 F. Supp. 2d 584 (2004). (These factors, first set forth by Sir Austin Bradford Hill, also have been referred to as "viewpoints," emphasizing that one or more of the factors may be absent even where a causal relationship exists and that no factor is a sine qua non of causation (593).)

130 *Merrell Pharmaceuticals v. Haver*, 953 S.W. 2d 706, 724 (It must be reiterated that even if a statistically significant association is found, that association does not equate to causation. Although there may appear to be an increased risk associated with an activity or condition, this does not mean the relationship is causal. There are many other factors to consider in evaluating the reliability of a scientific study including, but certainly not limited to, the sample size of the study, the power of the study, confounding variables, and whether there was selection bias (724).) *In re Breast Implant Litigation*, 318 F. Supp. 2d 879 (Colorado D.C. 1998) (The Bradford-Hill criteria are then used "to establish scientific cause and effect . . . . such criteria as the temporal sequence of events, the strength of the association, the consistency of the observed association, the dose-response relationship, and the biologic plausibility of the observed association.") *In re Joint E. & S. Dist. Asbestos Litig.*, 827 F. Supp. 1014 (S.D.N.Y. 1993).


132 An early work, which appears to insist on using Hill's criteria, is Bert Black and David Lilienfeld, "Epidemiologic Proof in Toxic Tort Litigation," *Fordham Law Review*, 52 (1984): 732–785, at 764 (arguing that all of Koch's hypothesis should be satisfied). See also Bernstein, "The Admissibility of Scientific Evidence," 2166, 2168. Bernstein's remarks are ambiguous between the claim that all of Hill's criteria must be met for a study to be admissible (a view that is clearly at odds with good scientific practice) and the claim that if none (which would include the temporal criteria) of the criteria are met a study is not admissible. The second contention would be correct while the first I would sharply disagree with, as did Hill himself. See Hill, "The Environment and Disease," 19. Furthermore, Bernstein argues that if "proffered epidemiological evidence meets some but not all of the criteria a judge would do well to consult with a court-appointed epidemiological expert to assist her in judging the reliability of the evidence." Bernstein, 2168–2169. Based on Hill and contemporary epidemiologists' views, judges should be hesitant to rule epidemiological studies inadmissible on such grounds since the absence of any single criteria (with the exception of temporality) is consistent with causation.

133 Hill, "The Environment and Disease," 19.


136 See *In re Joint E. & S. Dist. Asbestos Litig.*, 827 F. Supp. at 1038. The court stated: "While none of the Sufficiency Criteria is decisive by itself in determining the sufficiency of a plaintiff's epidemiological evidence in the context of a Rule 56(b) motion, sufficient epidemiological evidence will necessarily satisfy several of these criteria. More significantly, when epidemiological evidence fails to satisfy any of the Sufficiency Criteria, it cannot be relied on to support a jury verdict in the face of a motion for judgment as a matter of law."

---
turned Hill's considerations into criteria for judging the admissibility of expert testimony based on epidemiological evidence.

Subsequent epidemiological articles and the Federal Judicial Center's Reference Manual also reject the court's reasoning. 137 Rothman notes that although strength retains "some meaning as a description of the public health importance of a factor...[i]t is devoid of meaning in the biologic description of disease etiology" because whether an association is "weak" or "strong" depends on the "prevalence of complementary component causes in the same sufficient cause... [T]his prevalence is often a matter of custom, circumstance or chance, and is not a scientifically generalizable characteristic." 138 Other epidemiologists echo Hill concerning the plausibility and coherence factors. Susser notes, "Coherence is an ultimate and yet not a necessary criterion for causality.... But coherence supports existing inference and theory." 139 He continues: "[I]ncoherence may also have a more general explanation, in which instance it will generate a new theory. As Lilienfeld has said: 'the finding of a biologically implausible association may be the first lead to this extension of knowledge.'" 140

Some of their sharpest criticisms are saved for those who would emphasize "specificity." Some have noted that although there may be:

a tendency toward clustering of specific clinical features and other manifestations among patients afflicted with a particular cause of disease... and [it is possible to] find diseases in which there is very high association of a particular cause with a particular effect[,]... the majority of causal agents that are chosen as criteria for constructing disease entities are associated with a great diversity of clinical, pathological, and biochemical patterns. 141

Others are more critical.

Arguments that demand specificity are fallacious, if not absurd. There can be no logical reason why any identifiable factor, and especially an unrefined one, should not have multiple effects.... By now it is evident that the associations of health disorders with smoking depend on a variety of mechanisms,

some causal and some not. Specificity enhances the plausibility of causal inference, but lack of specificity does not negate it. 142

In sum, in the cases just reviewed the courts have made more stringent use of Hill's considerations than he and other leading epidemiologists would. Hill's considerations constitute questions scientists should ask about epidemiological studies to assist in their judgments about the strength of the evidence, but they must serve an explanatory function within a nondeductive inference to the best explanation (discussed in Chapter 4).

The Unfortunate Consequences of "No Effect" Studies

Although this point is more difficult to document, it appears that judges may take "no effect" epidemiological studies at face value and conclude that because a study fails to show an effect between exposure and disease, this demonstrates that the exposure does not cause the disease. 143 Frequently, courts are presented with epidemiological studies that show no relation between exposure and disease. Because plaintiffs must establish the basis of causation, if there is no epidemiological evidence for a causal relationship between exposure and disease, they must establish causation in some other way, provided it can be done. For example, they may have to use animal studies, case reports, structure-activity relationships, other molecular studies, and the like. If courts are reluctant to allow such evidence for causation, this precludes plaintiffs from going forward.

However, courts sometimes appear to go further and suggest that because epidemiological studies do not show evidence of an effect, there is not one. That is, they appear to assume that no effect studies compared with other kinds of evidence show there is no adverse effect. Such a view would commit multiple mistakes. That is, courts may believe that negative epidemiological studies trump all other kinds of evidence. In only a few cases do courts note that a "no effect" study might fail to show there is evidence of no effect; usually they are silent on this issue.

First, a defendants' evidence compared with plaintiffs' evidence should be irrelevant for an admissibility decision in which a court considers only the evidence on one side of the case. Second, comparing evidence between parties is applicable only for assessing whether plaintiffs' evidence is legally sufficient to create a material issue of fact for a jury to consider. Third, and most seriously, courts that make such an assessment risk a serious error on the science.

138 See Rothman, Modern Epidemiology, 43. This subtle point concerning a common model of causation was discussed in Chapter 4.
140 Susser, "Judgment and Causal Inferences," 77.
143 See Chambers v. Exxon Corp., 81 F. Supp. 2d 661, at 665-666 (M.D. La. 2000) (holding that expert testimony that benzene exposure causes chronic myelogenous leukemia ("CML") was inadmissible for lack of scientific reliability, in the absence of an epidemiological study that conclusively established a statistically significant risk of contracting CML from exposure to benzene (emphasis added)).
There are two possibilities for no effect studies. A study might only show no evidence of an effect or it might show the much more difficult point that there is evidence of no effect (this second possibility is not likely except in very special circumstances as I discuss below). On the first, one might say that there is not any human evidence of an adverse effect. On the second, there is more affirmative evidence of no effect.

These two judgments are clearly different as an analogy suggests. Suppose that I am looking over a large plain at some distance. If someone asks me if there are people down on the plain, not having seen any I might say “There is no evidence that there are.” However, this does not mean that humans are not present. I might not have had very good evidence that no one was on the plain. I might have poor eyesight. I might have used a telescope that had insufficient magnification to reveal people at such a distance. I might not have investigated the issue in any other way. I could better claim that there was evidence that no one was on the plain if I had verified that my telescope had sufficient magnification to reveal people if they were there, and used it carefully in examining the terrain, but still was unable to see anyone. However, even a high-resolution telescope might still be insufficient if I looked at the wrong time, for example, during the daytime and people only moved about on the plain at night. If I wanted to be highly certain that there was no one on the plain, then I might need to undertake a number of other investigations, such as leaving my vantage point and going to investigate much closer.

Thus, judges need to recognize the difference between “there is no evidence of an effect” and “there is evidence of no effect.” Even if they did this, however, neither result should legally contravene admissibility, as that is concerned only with plaintiffs’ evidence. Even if defendants believe they have no evidence of an effect, plaintiffs might well have evidence of an effect based on other kinds of data. For example, cancer researchers recently observed that epidemiological evidence that exposure to human papilloma virus caused cervical cancer lagged other kinds of evidence by five to seven years.\(^{144}\) The question then becomes, do plaintiffs’ experts have a reasonable enough foundation for their testimony to be admitted?

For judging the legal sufficiency of plaintiffs’ evidence on a directed verdict the “no effect” studies defendants have proffered is pertinent, but even here they are quite limited. If plaintiffs have offered some evidence of an effect based on other kinds of studies, simply because defendants have no evidence of an effect in particular epidemiological studies does not show there is evidence of no effect; it should not trump plaintiffs’ evidence. Defendants’ evidence could trump plaintiffs’ other evidence of an effect, only if defendants’ evidence could sufficiently establish that there was evidence of no effect from human studies and it so overwhelmed plaintiffs’ evidence that no reasonable jury could find for plaintiffs. However, this is an extremely difficult showing to make.

Scientists are extremely careful to avoid inferring from the fact that there is “no evidence of effect” that there is “evidence of no effect,” because of its invalidity and because of the enormous consequences that attach to such an inference. Thus, a consensus scientific body, such as the International Agency for Research on Cancer, requires that some highly specific and detailed conditions should be met before making such an inference. Even then they apply only to the disease end point of interest and not to unrelated diseases (the conditions are indicated in the footnote).\(^{145}\) Judges should become quite skeptical of no effect studies. They only show that there is evidence of no effect when the highly specific conditions IARC discusses are satisfied.\(^{146}\)

**Demanding Mechanistic Evidence**

The Fifth Circuit Court of Appeals in *Black v. Food Lion, Inc.*, decided subsequent to *Kumho Tire*, reversed a trial judge for admitting medical testimony that

---

144 Carbone et al., 5518–5519.

145 For example, the International Agency for Research on Cancer addresses the issue in the following way:

When several epidemiological studies show little or no indication of an association between an exposure and cancer, the judgment may be made that, in the aggregate, they show evidence of lack of carcinogenicity. Such a judgment requires first of all that the studies giving rise to it meet, to a sufficient degree, the standards of design and analysis described above. Specifically, the possibility that bias, confounding or misclassification of exposure or outcome could explain the observed results should be considered and excluded with reasonable certainty. In addition, all studies that are judged to be methodologically sound should be consistent with a relative risk of unity for any observed level of exposure and, when considered together, should provide a pooled estimate of relative risk, which is at or near unity and has a narrow confidence interval, due to sufficient population size. Moreover, no individual study nor the pooled results of all the studies should show any consistent tendency for relative risk of cancer to increase with increasing level of exposure. It is important to note that evidence of lack of carcinogenicity obtained in this way from several epidemiological studies can apply only to the type(s) of cancer studied and to dose levels and intervals between first exposure and observation of disease that are the same as or less than those observed in all the studies. Experience with human cancer indicates that, in some cases, the period from first exposure to the development of clinical cancer is seldom less than 20 years; latent periods substantially shorter than 30 years cannot provide evidence for lack of carcinogenicity. (International Agency for Research on Cancer, *Monographs on the Evaluation of Carcinogenic Risks to Humans*, Preamble, section 8, Studies of Cancer in Humans. Rev. Aug. 18 2004. Available at: http://193.51.164.11/monoeval/StudiesHumans.html.)

146 See *Ambrosini v. Labarague*, 101 F.3d 129, at 136 (1996), for a case in which the court recognizes the need to consider statistical power in order to assess the plausibility of a negative epidemiological study. (The court pointed out that plaintiffs’ expert “explained that an epidemiologist evaluates studies based on their ‘statistical power.’ Statistical power, he continued, represents the ability of a study, based on its sample size, to detect a causal relationship. Conventionally, in order to be considered meaningful, negative studies, that is, those which allocate the absence of a causal relationship, must have at least an 80 to 90 percent chance of detecting a causal link if such a link exists; otherwise, the studies cannot be considered conclusive.”)
plaintiff’s fall in defendant’s grocery store had caused her to develop fibromyalgia. This is a syndrome characterized by “generalized pain, poor sleep, an inability to concentrate, and chronic fatigue.” The physician had followed approved methods for identifying fibromyalgia. However, the appellate court found that because there is no known etiology for fibromyalgia (which the expert conceded), it held that it was scientifically illogical for expert to conclude that the disease must have been caused by the fall merely because she had eliminated other possible causes. The court argued that the district court should have determined “whether [the diagnosing physician] tied the fall at Food Lion by some specific train of medical evidence to Black’s development of fibromyalgia,” then continued into more troubling territory:

The underlying predicates of any cause-and-effect medical testimony are that medical science understands the physiological process by which a particular disease syndrome develops and knows what factors cause the process to occur. Based on such predicate knowledge, it may then be possible to fasten legal liability for a person’s disease or injury.

In this case, neither Dr. Reyna nor medical science knows the exact process that results in fibromyalgia or the factors that trigger the process. Absent these critical scientific predicates, for which there is no proof in the record, no scientifically reliable conclusion on causation can be drawn. Dr. Reyna’s use of a general methodology cannot vindicate a conclusion for which there is no underlying medical support. (emphasis added)

Such a requirement seems much too optimistic for many diseases and, thus, too demanding as a requirement for the admissibility of expert testimony. We have seen the importance of mechanistic information. Moreover, the court in Black v. Food Lion might well have made the correct admissibility decision in the particular case (I take no position on that). However, its reasons are mistaken, its rationale is too general and its view that mechanistic evidence is needed suggests a misunderstanding of various ways the scientific community can identify causal relationships.

Although the court’s general point is correct — understanding the mechanisms of disease can greatly assist causal inference, and scientists would like such data — it is a scientific mistake to make it a necessary condition of good causal inferences and a legal error as well. Scientists can and do make causal inferences without such evidence, and, unfortunately, frequently it is often not available. Recall the point from Chapter 5 about the widely used and well-studied over-the-counter drug, aspirin. Its therapeutic and adverse effects have been well documented. However, as recently as 1991 the mechanisms by which aspirin produces them were not understood.

Benzene constitutes another example. Consensus scientific committees have documented for more than two decades that benzene causes leukemia and aplastic anemia at very low exposure levels, for example, between one part per million and ten parts per million, perhaps as low as .1 ppm. Yet research scientists as of this date do not understand the physiological processes and mechanisms by which benzene causes these diseases.

Moreover, mechanistic evidence is asymmetrical — when it is present it can greatly strengthen a causal inference, but when it is absent it does not necessarily undermine the inference. Thus, it is not a scientifically necessary condition of causal inference.

The legal error is twofold. First, by requiring mechanistic evidence, courts demand something that often does not exist and something that is more than science may be able to deliver in the short run (and sometimes even in the very long run, e.g., recall scurvy). In addition, if experts must understand the “physiological predicate by which a particular disease syndrome develops” this will ensure that very few experts will be permitted to testify when such understanding is absent, thus, excluding nearly all litigants from court.

This is another instance in which a court has conflated particularly helpful evidence with necessary evidence for causal inference. When judges insist on ideal or very good evidence, they risk overlooking other good and quite reasonable evidence that could serve admissibility purposes. Again, the ideal becomes the enemy of the good.

Second, the courts’ words suggest that understanding the physiological process by which a disease develops is something that is known in every case or known easily or known in many cases in which scientists ascribe a disease to a particular cause. Although physicians and toxicologists have substantial understanding of diseases and their grosser causes, it appears to be comparatively rare that they understand the detailed step-by-step physiological processes and biological mechanisms by which the disease develops. It is only in the most well-studied diseases, even those that have been studied for decades, that such basic physiologic and mechanistic understanding is present.

In addition, often such understanding is irrelevant. For public health purposes, once the causation pattern between exposure and disease is well established, many researchers may have little incentive to investigate further the physiological route by which that exposure results in that particular disease.

147 Black v. Food Lion, Inc., 171 F.3d 308, at 309 (5th Cir. 1998).
148 Black v. Food Lion, Inc., 171 F.3d at 314.
149 For another decision that expresses such a view see McClain v. Metabolife International, Inc., 2005 WL 477861 (11th Cir. (Ala.), at 14–15.
150 Thagard, How Scientists Explain Disease, and discussion in Chapter 4.
152 David A. Eastmond, Director, Environmental Toxicology, University of California, Riverside, personal communication.
They have established to the satisfaction of their scientific community that a particular exposure contributes to a particular disease. Moreover, they now would know enough to take steps to prevent the exposure from occurring, which would prevent the disease in question. Continuing to study the physiological processes by which diseases occur is in the interests of researchers only if it appears to hold out the promise of making a wider contribution to the field or to understanding basic biology or toxicology. If the mechanism of biological action were always required before protective or compensatory legal actions were taken, few substances would have been addressed by regulatory or tort law.

The Mistaken Exclusion of Evidence

The Denigration of Animal Evidence

Some courts have excluded animal evidence as pertinent for experts making causal judgments about the effects of toxicants on humans or excluded it unless it was accompanied by epidemiological evidence. Yet this is a kind of data on which toxicologists and other scientists routinely rely for making toxicity assessments. Although animal studies do not provide mathematically certain means by which to infer the causal effects of a toxicant on humans, they are scientifically good and relevant evidence for identifying substances as human toxicants.154

Judge Weinstein's decision in In re Agent Orange Product Liability Litigation,155 has influenced a number of courts to exclude animal studies per se from evidence in toxic tort suits.156 He argued that, "[T]he studies on animal exposure to Agent Orange, even Plaintiffs' expert concedes are not persuasive in this lawsuit. . . . There is no evidence that plaintiffs were exposed to the far higher concentrations involved in [the animal studies] . . . ."157 Moreover, he argued that because the animal studies involved different biological species, they were not helpful to the case.158 He said the studies "are of so little probative force and are so potentially misleading as to be inadmissible. . . . They cannot be an acceptable predicate for an opinion under Rule 703."159 Although Judge Weinstein placed an emphasis on "this lawsuit," his opinion has been widely interpreted as excluding reliance on animal studies, unless they are accompanied by epidemiological evidence.160 Even when there has been no absolute legal barrier to the use of animal studies, they have faced undue skepticism from courts. In a Bendectin case the Fifth Circuit Court of Appeals in Brock regarded animal studies of "questionable applicability to humans," especially in the absence of some reference to epidemiological studies. Moreover, the court used as support for its legal rationale a highly controversial regulatory case from 1983 in which it objected to animal studies that show a risk of cancer from urea formaldehyde foam insulation.161 Courts' difficulties with animal studies, thus, have considerable history; perhaps it is time for them to better understand the science.

Apart from what may be errors in an understanding of toxicology, a generic exclusion of animal studies is mistaken on two counts. First, it appears contrary to the Daubert Court's emphasis on the consideration of scientific evidence that is relevant to expert testimony concerning causation. Clearly, scientists consider them quite relevant evidence in making toxicity judgments. Second, if one believes that it is appropriate for courts to consider the science that is available for assessing the toxicity of substances and their effects on human beings, then such evidence should ordinarily be permitted to be part of the foundation of expert testimony. Recall that such evidence is much more likely to be available than human studies, for instance.

For toxicologists, the fact that there is information from "other" biological species is both scientifically relevant and probative evidence. Moreover, contrary to Judge Weinstein, such studies have considerable probative force even if they might not always be as direct evidence of human harm as thorough, well-designed epidemiological studies with sufficiently large samples conducted for a sufficiently long duration to detect any toxic effects.

Animal evidence can and does have considerable explanatory power for toxicologists (Chapter 4). It also can, importantly, supplement or cast doubt on human data. For example, animal data might rule out a positive epidemiological study as having little biological plausibility. For a more specific example, it has been quite difficult to duplicate in animal studies the adverse effects seen in humans from exposure to electromagnetic fields. For some scientists, this casts doubt on some fairly consistent human epidemiological studies.162 By

---

154 See Chapter 4.
157 In re Agent Orange, 611 F. Supp. at 1241.
158 In re Agent Orange, 611 F. Supp. at 1241.
159 In re Agent Orange, 611 F. Supp. at 1241.
similar to humans in the way they react to the chemical in question, and because none involved studies the federal government had relied on as a basis for concluding the chemical was a probable health hazard [as was true in this case].

The Paoli opinion is on much firmer scientific ground than some of the opinions noted earlier.

However, even the Third Circuit's view of the pertinent evidence may not be fully accurate. That is, given what toxicologists know and how they view the evidence that is pertinent to making causal judgments, courts should be open to a wider range of toxicological evidence than even the Third Circuit suggests. Animal studies should not be excluded even in the face of no effect epidemiological evidence to the contrary (as indicated earlier). Such evidence might or might not be ultimately legally sufficient, given evidence on the other side of the case, but that is a separate matter. Moreover, animal evidence typically is scientifically pertinent to judgments of whether toxic substances cause human harm. As Judge Golkiniewicz in Stevens points out, animal evidence can provide "biological plausibility" of an effect or perhaps suggest doubt about a human study because it cannot be replicated in an animal model. Moreover, that expert opinion based partly on animal studies should be excluded; and Saabho Rubników v. Witco Chem. Corp., 242 N.J. Super. 86, 576 A.2d 7, 15 (1990) (under New Jersey law reversing trial court's exclusion of expert testimony, which was partly based on animal studies that PCBs caused cancer). In Villari v. Terminix Int'l Inc., 692 F. Supp. 568, 579 (E.D. Pa. 1988), Judge Pollak explained that:

While it may be true that defendant can offer tests and experiments that do not support the findings of plaintiff's expert, the defendant cannot deny that animal studies are routinely relied upon by the scientific community in assessing the carcinogenic effects of chemicals on humans. Even defendant's own expert acknowledges that animal experiment studies are built on "prudent presumptions," although he concludes that they should not be admitted.


The Paoli court's own citation provides helpful support. See, e.g., In re Bendectin Prod. Liab. Litig., 732 F. Supp. 744, 749 (E.D. Mich. 1990) (experts in the field think it is reasonable to rely on nonepidemiological studies to link Bendectin to birth defects); Hagen v. Richardson-Merrell, 697 F. Supp. 334, 337 (N.D. Ill. 1988) (defendants did not adequately demonstrate "a scientific basis" for their argument).

63 See Ball et al., "Alternatives to Using Human Experience," 356.
64 At the time, governmental epidemiological studies did not show adverse long-term health effects from exposure to Agent Orange. ("These epidemiological studies alone demonstrate that on the basis of present knowledge, there is no question of fact: Agent Orange cannot now be shown to have caused plaintiffs' numerous illnesses." In re Agent Orange Prod. Liab. Litig., 611 F. Supp. 1223, 1241.)
66 For a summary of some of this work, see Committee to Review the Health Effects in Vietnam Veterans of Exposure to Herbicides, Institute of Medicine, Veterans and Agent Orange (1996), 14 (Some subsequent studies that show that people exposed to the herbicides used in Agent Orange in occupational or environmental exposure had increased risks of various cancers and other diseases and that at least Vietnam veterans with very slight exposure to Agent Orange "could have risks approaching those in occupational and environmental settings"); and Brief Amici Curiae of the Lymphoma Foundation of America, Carl F. Capraro, Devra Davis, Peter L. Detus, Brian G. Duey, Alan H. Lockwood, David Ozenoff, Arnold J. Schecter, and David Walling in support of Respondents, in Dow Chemical Company, Monsanto Company v. Stephenson, and Isaacson (Supreme Court of the United States, 2002).
68 The Paoli court's own citation provides helpful support. See, e.g., In re Bendectin Prod. Liab. Litig., 732 F. Supp. 744, 749 (E.D. Mich. 1990) (experts in the field think it is reasonable to rely on nonepidemiological studies to link Bendectin to birth defects); Hagen v. Richardson-Merrell, 697 F. Supp. 334, 337 (N.D. Ill. 1988) (defendants did not adequately demonstrate..."
animal evidence appears to be available for a wider range of substances than human evidence.

An interesting counterexample to a claim about the irrelevance of animal evidence is provided by the scientific detective story considered in Chapter 4 concerning dimethylnitrosamine poisoning. Recall that this was a scientific and legal case in which animal evidence and a small number of human case reports combined with other circumstances, revealed the cause of two deaths and led to the criminal conviction of the person who was responsible for poisoning them with dimethylnitrosamine. Data about mechanism, carcinogenic doses, and lethal doses came from animal or in vitro studies, not human epidemiological studies.172 If some of the rules concerning the nonadmissibility of animal evidence in tort cases had been applied to exclude the evidence in that criminal case, a criminal would have gone free. More important, the scientists used all of the toxicological evidence they had available to them.

What is needed to establish causation in a tort case is an explanation based on scientifically relevant studies that is more probably than not true connecting the defendant’s actions to the plaintiff’s injuries. However, providing an appropriate explanation does not automatically require the use of only human epidemiological evidence. For admissibility only a respectable scientific explanation that is supported by sufficient data, which is reliably applied to the facts of the case are needed. Thus, judicial and commentator insistence that the explanations must have certain necessary components is mistaken. Moreover, there are instances in which animal evidence conjoined with short-term test and structure-activity relationships might well be sufficient to show more probably than not that a substance is a human carcinogen.173 We return to some of these issues in Chapter 7.

**Discriminating among Animal Studies**

More recently, in *Allen v. Pennsylvania Engineering*, both a district court and the Fifth Circuit Court of Appeals rejected plaintiff’s evidence and held that the fact that ethylene oxide caused brain tumors in *rattus* could not be evidence for the claim that ETO could cause brain tumors in humans. The reason: ETO did not correspondingly cause brain tumors in phylogenetically similar mice.174 There is generic plausibility to such an argument. However, the court did not probe further and, thus, left a problematic decision. There tends to be general concordance of toxicity effects between two phylogenetiically close species such as rats and mice, but there is no necessity to it. Substances may well be more toxic in one species than another, and still be toxic to humans.175 The underlying principle is that different species may show different toxic effects to a greater or lesser degree. Rats may be more or less susceptible to a given toxicant than mice.176 For example, both the human carcinogens Direct Black 38 and Direct Blue 6, two benzidine based dyes, are carcinogenic in rats but not in mice under the same experimental conditions and routes of exposure.177 MPTP – 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine – a chemical causing a Parkinson’s disease–like condition in humans, induces a similar neurotoxic effect in mice, but not in rats.178 Melphalan, a human carcinogen, is positive in rhesus monkeys and shows no effect in the phylogenetically similar rhesomologous monkeys.179 If substances are carcinogenic in two species, the probability that they are carcinogenic in humans is greatly increased; but a substance might be quite potent and harmful to humans even though it did not result in carcinogenicity in at least two rodent species. Frequently, substances have not been tested, or not adequately tested, in other species. Toxicologists try to design studies to have the best chance of detecting toxic results that are pertinent to humans.180

---

172 See Renate D. Kimbrough, "Case Studies," in *Industrial Toxicology: Safety and Health Applications in the Workplace*, ed. P. L. Williams and J. L. Burson (New York: Lifetime Learning Publications, 1985), 414, 417–420. Kimbrough notes that the evidence showing the toxicity of dimethylnitrosamine was based on studies in rats and then an amount lethal to adult humans was calculated from the results of those studies. See Kimbrough, "Case Studies," 417–420.

173 See, e.g., Kimbrough, "Case Studies," 417–420 (describing cases where case studies on cancer in animals were useful in determining cause of death in humans).

174 See *Allen v. Pennsylvania Eng’g Corp.*, 102 E3d 194, 197 (5th Cir. 1996).


176 See Gad, "Model Selection and Scaling," 841, 849, and Tables 5, 6.


180 The literature on these issues is substantial. See Gad, "Model Selection and Scaling," 813–840.
Even though the ETO toxicity results in rats were not duplicated in mice, could they be good evidence? The court should have inquired whether (or plaintiffs should have argued that) there was something important about the rats as a model for humans. Did they have more plausibility than might have been apparent? I return to this in Chapter 7.

Target-Site Arguments

Judges also will need to adjudicate a general criticism of animal evidence that because the target sites of cancer in animals are different from the target sites in humans, this cannot be evidence for carcinogenicity in humans. This is another view that is tempting, but not generally correct. For example, benzidine is a known human bladder carcinogen, but is not a bladder carcinogen in animal species (except for dogs), although it induces tumors in hamsters (liver tumors), rats (liver, ear duct, mammary, and intestinal tumors), and mice (liver tumors). Moreover, there appears to be no scientific agreement that there must be tissue concordance between animals and humans. Concordance in tumor sites, although considerably strengthening the evidence, is not essential, of chemical toxicity are largely identical in humans and animals. (Gad, “Model Selection and Scaling,” 813)


[The mechanisms of control of cell growth and differentiation are remarkably homogeneous among species and highly conserved in evolution. Thus far, there is evidence that growth control mechanisms at the level of the cell are homologous among mammals, but there is no evidence that these mechanisms are site concordant. Moreover, agents observed to produce tumors in both humans and animals have produced tumors either at the same (e.g., vinyl chloride) or different sites (e.g., benzene) (NRC, 1994).

As indicated earlier and in Chapter 4, the results of well-conducted animal tests can provide reliable evidence for the toxicity and carcinogenicity of chemical and physical agents. As previously noted, these results need to be evaluated sensitively for reproducibility and relevance to humans. In reviewing animal studies, some courts and commentators have suggested enshrining into law more stringent criteria for judging the validity of scientific inferences and explanations that are required in the science itself. Courts are not being faithful to the science. Moreover, the adoption of rigid rules places more stringent legal restrictions on litigants than scientists themselves adopt. These constraints will legally skew admissibility decisions, a point to which we return in the conclusion of this chapter.

Chemical Structure–Biological Activity Evidence

Many courts have routinely excluded expert testimony based on chemical structure–biological activity relationships as pertinent to assessments of causal effects of toxicants on humans, arguing that at best they form the basis of a hypothesis, not evidence, of causation. Indeed, structure–activity relationships have a number of well-known difficulties. One legal scholar whose comments have now been repeated by courts noted that a small difference in chemical structure can make a major difference in biological activity. There is truth to this point, but that is not the end of the story, as some courts seem to regard it. As already discussed (Chapter 4), properly understood chemical structure–biological activity evidence can contribute substantially to causation in certain contexts and in any case could be part of any scientifically integrated evidence of causation. In general, it is not the strongest scientific inference to argue from similarity in chemical structure to similarity in biological activity, as even specific similarities in chemical structure with minor dissimilarities elsewhere can result in fairly significant differences in biological effects. Recall, however, the Institute of Medicine’s and National Research Council’s view of the importance of structure–activity relationships from Chapter 4.

However, for chemical families with certain properties, there also are scientifically quite strong inferences. For example, molecules with chemical groups

---


186 Sanders, Bendictus on Trial. ("Molecules with minor structural differences can produce very different biological effects" [46].

that are known to interact with mammalian DNA or proteins provide strong, but not infallible, reasons for thinking that substances with chemical similarities have similar biological activity.\(^{187}\) For another example, scientists regard substances that bind to the Ah receptor (aryl-hydrocarbon receptor) as being sufficiently similar that they can assign a toxicity equivalence factor to judge the toxicity of different substances.\(^{188}\) Clearly, structure-activity relationships can assist scientists in assessing the toxicity of a substance to humans; the data needs to be evaluated sensitively. If substances are members of classes that have strong chemical and biological relationships as some of the examples just suggested, they have even greater evidentiary value.\(^{189}\)

How much structure-activity relationships can contribute to a given toxicity judgment depends on the particular structure-activity data, on the other evidence available and on how it all "fits together." Courts in assessing expert testimony should permit experts to utilize all the kinds of evidence on which scientists themselves rely in making their judgments of causation, including appropriate structure-activity evidence.

Moreover, there can be wider kinds of evidence at the molecular level that can be powerful evidence. For example, ethylene oxide is a small, direct-acting molecule (it does not need metabolic transformation) that can reach nearly any tissue and cause mutations. Most substances may not have this property, but courts need to allow for such possibilities.

### The Exclusion of Case Studies as Evidence

In Chapter 4, I discussed some of the circumstances in which case studies could be good evidence for causation or at least contribute to judgments of causation. However, case studies have tended to fare quite badly in toxic tort suits. As a co-author and I will report elsewhere,\(^{190}\) judges in fifteen of seventy-seven tort cases admitted case reports as a substantial basis of expert testimony for the purpose of proving causation.\(^{191}\) Eight of these cases involved adverse drug/nutritional supplement reactions (three Parodel cases, two cases involving psychiatric medications, two cases involving diet pills, and a phenylpropoanaline (PPA) case). In three cases, typing on a keyboard or lifting a heavy object was alleged to have caused muscular/skeletal injuries. Another nineteen of the seventy-seven cases remain on the list because they contain relevant comments regarding the uses of case reports. However, they do not make explicit rulings regarding the admissibility of case reports as evidence for medical causality assessments.\(^{192}\)

The forty-three remaining cases all found expert testimony based in part on case reports to be inadmissible. In two of these cases, the judges explicitly stated that expert testimony based on some case reports can satisfy Daubert criteria; but the particular case reports presented to the court were not reliable. Among the remaining forty-one cases rejecting case reports, many judges rejected case reports categorically, but several were ambiguous as to whether case reports can ever be admissible.\(^{193}\)

We also examined forty legal cases from litigation under National Vaccine Injury Compensation Program, a body of law that has somewhat different procedural rules. There are somewhat less formal procedures for introducing evidence. In addition, although the judges are not explicitly required to use Daubert procedures and standards, many of them indicate they are following it. Finally, the judges use the preponderance of the evidence standard of proof.

Vaccine injury cases also are somewhat different from normal tort cases in that the judge serves both as the adjudicator of the law and as the fact-finder; there is no jury to protect from mistaken evidence or expert testimony. Nonetheless, judges must still decide whether or not to resolve the factual issues based in part on, more rarely, almost totally on case reports.

Judges in vaccine injury cases are much more receptive to the use of case reports supporting causal judgments than are most of the judges in traditional torts. Although case reports were not accepted as causation evidence in every hearing, they were accepted quite frequently. What is most interesting is that none of the seven judges who decided the forty cases categorically rejects the scientific relevance of case reports. This highlights a notable inconsistency between the judges in vaccine injury cases and many federal tort judges in the ways that case reports are perceived for supporting causal judgments. And since both sets of judges are reviewing case reports for their probative value as evidence for medical causation, the inconsistency appears to be based on a basic disagreement between judges as to the reliability or scientific relevance of case reports. The judges in the vaccine injury cases are much more nearly correct on the scientific issues than are the majority of the Article III federal judges, who tend to reject case reports. The reason for this appears to be that

---

187 J. Ashby and R. W. Tennant, "Chemical Structure, Salmonella Mutagenicity and Extent of Carcinogenicity as Indicators of Genotoxic Carcinogenesis among 222 Chemicals Tested in Rodents by the U.S. NCI/NTP," Mutation Research 204 (1988): 17–115; David A. Eastmond, Chair, Department of Environmental Toxicology, University of California, Riverside, personal communication.


189 Faustman and Omenn, "Risk Assessment."


191 One of these cases is Cella v. United States, 998 F2d 419 (C.A. 7 (Ind.) 1993), an appellate case that appears to have been decided under the Frye test.

192 For example, several of these cases discuss the use of case reports for establishing enough evidence to constitute a "duty to warn."

193 The primary decision that judges are required to make in Daubert hearings is whether the testimony (based on certain evidence) offered to the court is admissible. Because some judges found the particular case reports presented to the court to be unreliable, they did not bother to evaluate whether case reports are categorically bad evidence.
they are much more familiar with case reports and how they can inform or be decisive for causation judgments.

Like structure-activity relationships, case studies are not mathematically certain guides to causal relationships. Some case studies are good evidence; others are not. Thus, not every positive case study provides evidence of causation. However, not every positive epidemiological study or every positive animal study provides evidence of causation as well. If tort judges are to carry out their gatekeeping duties well, they must learn to evaluate and sift the evidence at the bar in order to admit expert testimony based on biological evidence that is scientifically relevant evidence of causation. In Chapter 4, I reviewed clear examples of good case studies, well accepted in the scientific community, and discussed considerations that consensus scientific bodies use to judge which case studies are good ones.

Good case studies are less rare than the recitation of a few examples might suggest. The World Health Organization has found that about 17 percent of case studies they considered are the basis of certain or probable casual relationships between vaccine exposure and adverse reactions. Seventeen percent is a comparatively small percentage, but it is not negligible. Moreover, because courts are required to review carefully all the evidence litigants present that might assist the jury in coming to its decision, they need to consider whether the case studies at the bar are scientifically relevant evidence of causal relationships or not. Finally, even if one case report (or several) is (are) not sufficient by itself (themselves) to support a causal judgment, they can importantly contribute to other evidence that is available.

Sometimes a case report by itself can support a causation judgment. For example, the Special Magistrates from the Vaccine Injury Compensation Program utilize such evidence and sometimes found single case studies quite compelling. Consider the opinion in Stevens v. the Secretary of Health and Human Services. Judge Golikewicz points out that the special masters “have debated the utility of case reports” for causation inferences. Moreover, even some who initially opposed them

concluded that a single persuasive case report and a petitioner whose symptoms matched the case report’s facts adequately supported petitioner’s actual causation claim for a tetanus toxoid case GBS... Later, ...[the same special master] opined that a single case report may support the possibility that a vaccine can cause a certain injury, “[i]f sound medical and scientific principles have been applied in that one case and the matter has been published for peer review.”

The tetanus toxoid case referred to was example (2) from the case studies discussed in Chapter 4.

**Never Throw Evidence Away**

In order to evaluate possible explanations of a phenomenon, a scientist must consider all the scientifically relevant evidence. In the law, however, some courts have been excluding as irrelevant individual pieces of evidence that certainly appear to be relevant to scientific judgments. Particular courts appear to decide a priori that whole categories of evidence, for example, structure-activity relationships, case studies, or animal studies, scientifically cannot contribute to (or cannot support by themselves) a scientific inference of causation and proceed to eliminate them as well as the expert who relied on them. However, as discussed in Chapter 4, a scientist’s inference to a conclusion must be evaluated based on all the relevant evidence.

A number of consensus scientific committees have argued that scientists should, “Never throw evidence away.” For example, a large group of scientists and physicians involved in assessing adverse events from immunization proposed a method for assessing vaccine-caused adverse events that is based on the best available information, [such that] [t]he maximum use is made of all available information and nothing is arbitrarily discarded” (emphasis added).

Some scientific methodologists make the same point in even stronger terms:

A causality assessment method must respect Fisher’s fundamental rule of uncertain inference — never throw information away. That is, any fact, theory or opinion that can affect an evaluator’s belief that [a particular exposure] caused an adverse event E must be incorporated by the method into the "state of information" on which the assessment is based.

In the law, if a particular kind of evidence is scientifically relevant to scientists’ causality judgments, it should be available as part of a body of integrated evidence that a court then considers when reviewing scientific testimony for

---


195 Stevens v. the Secretary of Health and Human Services, at 13–15.

196 Stevens v. the Secretary of Health and Human Services, at 15 (quoting O’Leary v. Secretary of HHS, No. 90–17294, 1997 WL 254217, at 3).

197 Scientists and physicians involved in assessing adverse events from immunization propose a method for assessing vaccine-caused adverse events that is "based on the best available information, [such that] [t]he maximum use is made of all available information and nothing is arbitrarily discarded," Gerald M. Fischel, David A. Lane, John B. Livengood, Samuel J. Brofritz, John H. Menkes, and James F. Schwartz, "Adverse Events Following Immunization: Assessing Probability of Causation," Pedriatic Neurology 5 (1989): 287–290, esp. 290.

198 Fischel et al., "Adverse Events Following Immunization," 290.

199 Hutchinson and Lane, "Standardized Methods of Causality Assessment," 10 (emphasis added).
admission. The court procedure for judging overall admissibility might be the following: Presumptively, courts should not rule as irrelevant any *individual* scientific fact, theory, or opinion that plausibly can affect a scientist's belief that a particular exposure caused an adverse event. Courts should consider all the relevant evidence on which a scientist bases her inferences, then review whether it is integrated or "fits together in the right way" to support reliable expert testimony. Or, in the words of Rule 702, does the expert testimony rest on sufficient data (all data taken together), is it the product of reliable principles and is it reliably applied to the facts of the case? Specifically Rule 702 should be interpreted so that courts consider all an expert's scientifically relevant evidence to determine whether that body of integrated evidence sufficiently supports expert testimony.

Such evidence might ultimately be *inadmissible if the body of a litigant's evidence, taken as an integrated whole did not support "scientifically reliable" expert testimony or if it did not fit the facts of a case. Yet to date some courts appear to have violated the aphorism - "never throw evidence away" - excluding individual pieces of evidence that are scientifically relevant to a scientist's judgment about causation.

Requirements to Expert Testimony

In assessing expert testimony courts need to determine whether plaintiffs can "demonstrate 'the levels of exposure that are hazardous to human beings generally as well as the plaintiff's actual level of exposure.'"

In *Wright v. Willamette Indus.*, the court reasoned that "a plaintiff in a toxic tort case must prove the levels of exposure that are hazardous to human beings generally as well as the plaintiff's actual level of exposure to the defendant's toxic substance before he or she may recover." Arguing that courts and juries in toxic tort cases must "make more particularized inquiries into matters of cause and effect..." it goes on to argue that

At a minimum, we think that there must be evidence from which the factfinder can conclude that the plaintiff was exposed to levels of that agent that are known to cause the kind of harm that the plaintiff claims to have suffered. ... We do not require a mathematically precise table equating levels of exposure with levels of harm, but there must be evidence from which a reasonable person could conclude that a defendant's emission has probably caused a particular plaintiff the kind of harm of which he or she complains before there can be a recovery.

Despite this seemingly reasonable language, the court adds that

In this case, while the Wrights proved that they were exposed to defendant's emissions and that wood fibers from defendant's plant were in their house, their sputum, and their urine, they failed to produce evidence that they were exposed to a hazardous level of formaldehyde from the fibers emanating from Willamette's plant. Their experts' information on this subject was simply insufficient.

Although exposure is a needed factual predicate in an argument concerning injury from toxic substances, it is often a particularly difficult one to establish. Frequently, people are unaware that they are exposed. This occurred with radiation near the Hanford nuclear site in Washington state. Exposures can be quite accidental as they were in *Moore v. Ashland Chemical Co.* (chemicals spilled in the back of a truck, which the driver had to clean up in a very closed space), or otherwise difficult to document or measure. Victims do not carry monitors with them to document the extent of their exposures. Exposure information is difficult enough to provide for risk assessments by regulatory agencies (because of poor exposure records even in workplaces), and often there is a considerable demand for high degrees of quantification in that context. However, given the adventitious and often accidental exposures typical of toxic tort cases, it is an even more difficult factual issue to quantify with any degree of precision. Yet defendants may press strongly on this point because it is in their interest to do so. Courts should not expect or demand considerable precision or quantification of exposures.

The Federal Judicial Center's advice on this issue, endorsed by some courts, seems correct:

Only rarely are humans exposed to chemicals in a manner that permits a quantitative determination of adverse outcomes.... Human exposure occurs most frequently in occupational settings where workers are exposed to industrial chemicals like lead or asbestos; however, even under these circumstances, it is usually difficult, if not impossible, to quantify the amount of exposure.
Courts have adopted views ranging from the quite stringent to more flexible toward the amount of exposure that must be shown to satisfy the exposure requirement in typical toxic tort suits. The Wright court seemed especially stringent on this issue. Clearly, the Wrights had considerable exposure from formaldehyde impregnated wood dust and fibers—it was in their spum and urine—but the court was sufficiently unsure about the formaldehyde exposure to admit the testimony. Given the fortuitous nature of such exposures, it appears that those courts that are more flexible on the issue provide a fairer forum to the parties involved. More flexible views appear to have been endorsed by the Fourth Circuit Court of Appeals in Westberry v. Gislaved Gummi AB, as well as the Third Circuit Court of Appeals in Heller v. Shaw Industries.

Thus, although excellent exposure information would be very helpful in toxic tort suits, it is usually quite poor and often difficult to obtain. Courts need to be quite sensitive in reviewing this part of plaintiff’s case because a too-stringent and rigid approach will eliminate meritorious cases. The concern about exposure is heightened in light of (a) susceptible subpopulations and (b) the law concerning susceptible individuals. As we considered earlier, some individuals are more susceptible to toxic exposures than others. Thus, too demanding an approach toward explicit exposure data risks leaving susceptible individuals uncompensated for wrongfully inflicted injuries. Moreover, as a result of the eggshell skulls doctrine in torts, such persons are entitled to protection.

Lumping vs. Splitting Toxicological Evidence

Defendants often argue that if there are two related human diseases, judges should consider them separately and assess how much evidence is available showing that a potentially toxic substance causes each disease. For example, benzene is well known for being associated with a number of blood and bone marrow related diseases, such as acute myelogenous (AML), myelomonocytic (AMMoL), monocytic (AMoL), and chronic myelogenous leukemias (CML), as well as several others.

A common defense strategy is to separate these diseases into subtypes and ask how much particular evidence favors the claim that benzene exposure (and at what levels) causes each subtype of leukemia. Splitting the diseases into subcategories can separate some diseases that are more common from some that are much less common. Even if associations between exposure and disease have been documented for more common diseases, they may not have been as well documented or not documented at all for rare subtypes or relatives of the same generic diseases. Moreover, for reasons that seem not to be understood, some diseases are particularly rare in some countries or cultures—chronic myelomonocytic leukemia is much rarer in China than it is in other countries (or cultures). With respect to leukaemias Peter Infante, one of the major researchers on leukemia and for many years an official of the National Institute of Occupational Safety and Health, the research agency of the Occupational Safety and Health Administration, notes the problem with this approach:

One of the difficulties in determining whether a specific type of leukemia is associated with an elevated relative risk from formal epidemiologic study of benzene exposed workers stems from the fact that leukemia is a relatively rare form of cancer and further subdivision into specific types results in very little statistical power to evaluate such relative risk.

David Savitz and Kurtis Andrews, epidemiologists at the University of North Carolina at Chapel Hill, echo this view.

Dr. Infante argues that when rare subtypes of diseases have not been identified because epidemiologic studies are too insensitive, he would rely upon other kinds of evidence, such as case reports, animal studies, if they were available, and provide more generic modes of action arguments, to assist in determining whether exposure to benzene was associated with rarer subtypes of leukemia.

This is in keeping with standard scientific approaches in understanding disease causation.

Nonetheless, such scientific arguments did not persuade a federal judge, who accepted defense arguments for splitting diseases into rarer subtypes. In Chambers v. Exxon, the judge argued that Exxon produced “a number of scientifically performed studies which demonstrate no association between exposure to

207 Westberry v. Gislaved Gummi AB, 178 F.3d at 264 (4th Cir. 1999). ("Consequently, while precise information concerning the exposure necessary to cause specific harm to humans and exact details pertaining to the plaintiff's exposure are beneficial, such evidence is not always available, or necessary, to demonstrate that a substance is toxic to humans given substantial exposure and need not invariably provide the basis for an expert's opinion on causation.

208 167 F.3d 146, at 157 (3d Cir. 1999) (noting "that even absent hard evidence of the level of exposure to the chemical in question, a medical expert could offer an opinion that the chemical caused plaintiff's illness").


benzene and development of CML [chronic myelogenous leukemia]. He ruled Infante's argument "unreliable" and plaintiff's case was at an end. We have already seen that "no effect" epidemiological studies do not necessarily imply that there is evidence of no effect and that they might well be falsely negative. That is a greater problem when courts permit rare diseases to be split into even more rare subtypes, which may or may not be scientifically plausible. (In addition, even if there are good scientific reasons to split diseases, courts should not then insist on epidemiological studies (which may be much too insensitive to detect such rare subtypes), but permit other kinds of evidence to support causation inferences.)

Further Confusions about Weight-of-the-Evidence Arguments

The failure of courts to understand inference to the best explanation, beginning with the Joiner district court and endorsed by the U.S. Supreme Court, has led to several confusions about scientific arguments that must be addressed to better align legal outcomes with the needed science in each case.

Confusing the Form of the Argument with the Standard of Proof

Recall that the term "weight of the evidence" is often utilized by regulatory agencies when they assess the toxicity of substances. Recall also that "weight of the evidence" is simply another term for noneductive inferences. However, some courts, noting that regulatory agencies utilize this term and that they are interested mainly in assessing risks, not causal relations as required by torts, dismiss "weight-of-the-evidence" arguments in torts for two reasons: (1) because they are merely about risks, not retrospective causation; or (2) because the standard of proof may be lower. The view of the Fifth Circuit Court of Appeals in Allen v. Pennsylvania Engineering, is representative and sometimes repeated.

We are also unpersuaded that the "weight of the evidence" methodology these experts use is scientifically acceptable for demonstrating a medical link between Allen's RO exposure and brain cancer. Regulatory and advisory bodies such as IARC, OSHA and EPA utilize a "weight of the evidence" method to assess the carcinogenicity of various human beings and suggest or make prophylactic rules governing human exposure. This methodology results from the preventive perspective that the agencies adopt in order to reduce public exposure to harmful substances. The agencies' threshold of proof is reasonably lower than that appropriate in tort law, which "traditionally make[s] more particularized inquiries into cause and effect" and requires a plaintiff to prove "that it is more likely than not that another individual has caused him or her harm." There are at least two confusions in such reasoning. First, the term "weight-of-the-evidence argument" is merely another term for noneductive inferences. Weighing and evaluating the extent to which evidence supports one explanation compared with another is pertinent to assessing both risks and retrospective causation; it is not restricted to assessing risks. Taken literally the phrase "weight of the evidence" refers to which explanation has the stronger balance of evidence in support of it. Thus, one might think of it as referring both to a reasoning process, and to the kind and amount of evidence in support of one explanation compared with others. Courts appear to have confused the form of inference and the subject to which they apply – risks or retrospective causation. Such argument forms apply to both.

Federal or state regulatory agencies routinely rely on human, animal, and other kinds of studies to predict risks to humans. However, although some agency deliberations are predictive and preventive in nature, others search for retrospective causal effects, as in the tort law. The Food and Drug Administration (FDA) (e.g., for drugs and new food additives) and parts of the U.S. EPA (for pesticides) are required by law to evaluate substances before they enter the market (acting under so-called premarket approval statutes) and before there is any significant exposure. Toxicological evaluation of substances in these circumstances is more predictive and explicitly preventive in nature.

However, other agencies engage in less predictive assessments. A number of regulatory bodies, such as the Occupational Safety and Health Administration and other parts of the EPA, act under postmarket statutes. In this capacity, they must act as scientific investigators and reconstruct a causal explanation of what led to disease or death with a further aim of reducing the adverse effect. These scientific inquiries are much more like those needed in the tort law; thus, the conclusions and deliberations are quite pertinent to tort law inquiries.


217 Moreover, weighing and evaluating evidence of the relation between exposures and risks is not substantially different from weighing and evaluating evidence for evidence of a retrospective causal relation between exposure and harm. In fact, scientists regard both activities as like those of scientific detectives, as the evidence for risks from exposures is likely to be some kind of harm that the exposures have caused to people, animals, mammalian cells, organs, or DNA, or some combination of these—the same kind of evidence scientists would use for assessing causes of harm (personal communication, David A. Eastmond, Chair, Program in Environmental Toxicology).


219 Chan and Hayes, "Principles and Methods," 212.
Second, some courts seem to have confused the level of proof required to establish an explanation with the form of the argument (nondeductive arguments) utilized to explain an event (e.g., the Fifth Circuit). There are numerous problems with this view. The nondeductive form of argument does not determine the standard of proof used to judge the strength of an explanation for an event. Instead, the institution in question typically provides the standard of proof. Does the institution require that the best explanation be certain, beyond a reasonable doubt, highly probable, or only marginally probable compared with alternative explanations? Thus, a scientific inference to the best explanation might support the best inference by highly certain evidence, by highly probable evidence, by only 51 percent of the evidence, or by only some of the evidence. To illustrate: diagnostic arguments are used to discover the cause of disease for treatment purposes and to discover or explain some of the properties of black holes in physics. Different standards of proof might easily be used to judge their explanatory success in different scientific contexts. Consider another example. The criminal law might require a much greater weight of the evidence to show that a DNA sample from a crime scene matched the defendant’s (because of its “beyond a reasonable doubt” standard of proof) than might be required for an analogous showing in the tort law to establish paternity (because of its “more likely than not” standard of proof). In either case, more evidence (the weight of the evidence) should favor the better explanation than alternative explanations.

Thus, it would be a mistake for courts to reject a form of argument — weight of the evidence — because a different standard of proof might (or might not [I have not conceded that point]) be utilized in regulatory settings. The form of argument is correct: the degree of certainty by which a conclusion is judged is the most plausible of the explanatory alternatives is quite different. Moreover, agency personnel who write regulations do not see that the burden of proof used in support of them as a lesser standard than the tort law. In fact, at least some view agencies as seeking to ensure that regulations are supported by evidence beyond a reasonable doubt.220

Confusing the Likelihood of Causation with Statistical Evidence for It

A second major issue is that some courts also have confused the degree of certainty of causation with frequency of adverse effects indicated by a statistical study. This is evident in the demand for statistical studies showing a doubling of risks in exposed populations. A fundamental legal issue concerning causation (where general and specific causation are bifurcated) is what is the likelihood that a particular exposure can cause a disease, and what is the likelihood that a particular exposure in fact caused a particular plaintiff’s disease? In both instances, the issue is what is the likelihood of disease causation.

Statistical evidence is not necessarily needed to show probability of causation; it just happens to greatly facilitate the task. Chapter 4 presented five examples of disease causation based on singular (or a small number of) events. For each of them, scientists judged that exposure to the substance in question certainly or probably caused the adverse reaction, but there were no statistical studies showing causation. Moreover, accident investigators assess the causes of airplane accidents or space shuttle disasters without statistical studies of all such accidents; there are simply too few.

The conclusion: to the extent that courts demand statistical evidence for probability of causation, they have it backwards. What must be shown is that it is more likely than not that exposure to defendant’s substance can cause injuries like the plaintiff’s and more likely than not did cause plaintiff’s injury. Statistical evidence is merely one way, sometimes the only way, and often a particularly persuasive way to establish this claim (because it has a kind of objectivity to it whereas experts’ probability judgments may concern courts). However, there may be other ways to show this using inference to the best explanation arguments, depending on the evidence that is available in a particular case. The examples of good case studies as well as accident investigations show this quite clearly. The procedure for establishing the likelihood of causation by means of a nondeductive argument is to show by means of the evidence available that it is more likely than not that defendants’ substance caused plaintiffs’ injuries (compared with other explanations to account for plaintiff’s disease).

Although this is a less algorithmic way to show such claims compared with statistical evidence, it is how scientists and the rest of us make such judgments all the time.

General and Specific Causation

There are also some issues that merit clarification concerning general versus specific causation. In federal toxic tort cases (something not necessarily true in all state jurisdictions221), courts tend to insist that a plaintiff must first show that it is more likely than not that a substance can cause the disease in question (general causation), and then show the substance did cause the particular plaintiff’s injury (specific causation). Some courts (and some recent articles) insist on the particular order of showing — first, general causation and then specific causation. This makes sense in many cases, because there is no necessity to the order. Recall the case studies example concerning GBS (example 2, Chapter 4).


221 This is clearly the case with examples (1), (2), (4), and (5) in Chapter 4.

222 See, for example, Donaldson v. Central Illinois Public Service Company, 199 Ill.2d 63 (2002). ("Illinois law does not define causation in terms of 'generic' or 'specific' causation. Rather, our case law clearly states that in negligence actions, the plaintiff must present evidence of proximate causation, which includes both 'cause in fact' and 'legal cause'" (90.).)
In that instance, Institute of Medicine scientists explicitly concluded, “because [this] case by Pollard and Selby (1978) demonstrates that tetanus toxoid did cause GBS, in the committee’s judgment tetanus toxoid can cause GBS.” That is, because specific causation had been shown by the evidence available, by deductive reasoning they concluded the possibility of general causation followed as a logical consequence. It is not clear how often such inferences will be possible. However, there is nothing wrong with the inference in question concerning GBS, and there is no statistical evidence in support of it. Consequently, courts and commentators must exercise care in laying down hard and fast rules about the order of demonstration between general and specific causation.

Consider this point with respect to what is called “differential diagnosis,” yet another term for inference to the best explanation. Joe Sanders and Julie Machal-Faulks have argued that the majority of decided legal cases suggests that in differential diagnosis, a physician or scientist must “rule in” the possibility of a general causal relationship between an exposure and a disease before “ruling out” all other possible explanations. A number of courts have held that it is not enough in differential diagnosis for an expert to rule out all other explanations and then find that the remaining one—that defendant’s substance caused plaintiff’s disease—can be ruled in and the only explanation left standing.

There is some good sense to this suggestion, but again no necessity. A particular explanation must be a plausible one—ceteris paribus, it must be biologically plausible, typically must be consistent with what is known (although we have seen that these two considerations can be overemphasized), and satisfy other explanatory considerations. However, as the tetanus toxoid and vinyl chloride case studies show, given background information, particular exposures, and ruling out of other explanations, a previously unsuspected causal relationship can be revealed on the basis of a singular event or small number of events. Moreover, because a particular exposure did cause an adverse reaction, it follows that the exposure can cause that reaction. In such cases, because there was such a powerful inference to the best explanation from the evidence available, specific causation implies general causation. In the tetanus vaccine—GBS example, a particular explanation was ruled in by compelling singular circumstances and background information without general causation having been independently established. General causation was simultaneously shown by specific causation.

A similar inference occurred in the vinyl chloride case. When there is such evidence one rightly can say the scientists discovered general causation as a result of specific causation in the particular case(s) they were investigating. Such outcomes are probably unusual, but in the spirit of Daubert it seems courts should allow for this in their review of expert testimony and its foundation.

Statistical evidence can serve another purpose for medical and legal purposes by indicating the rate of disease in exposed populations compared with background rates. For public health purposes, if exposure to a vaccine or an anesthetic greatly increases disease rates above background, this is important information concerning whether or not to pursue a vaccination program or allow exposures to an anesthetic. If exposure only slightly increases disease rates above background, there may be no public health concern. For legal purposes, if courts have such statistical information, this can provide some evidence about the likelihood of causation. It is helpful but not a necessary condition (recall the discussion in Chapter 5).

The last two points raise a more significant one for which I will not argue, but for which I will offer a conjecture. Certainly courts, and the scientific world more generally, have become captivated by the idea that various kinds of statistical support are needed for scientific conclusions. Statistical studies seem especially important for biological research given the high degree of variability within populations. Moreover, the emphasis on statistical studies is neither surprising nor remarkable, given the methodological and other progress that has been made with statistical reasoning in science in the twentieth century. In addition, certain kinds of statistical evidence—in particular double-blind clinical trials with large numbers in the control and the exposed groups—can be paradigmatic of excellent scientific studies. The mistake, and it is a mistake, is to confuse a common example of good scientific reasoning for what is required of all scientific and factual reasoning. Austin Bradford Hill is especially sharp on this point. Harman and Wright both argue that induction by enumeration

---

223 Institute of Medicine, Childhood Vaccines, 85 (emphasis in original).
225 On a separate legal issue, if a defendant is subject to a negligence standard of liability, a plaintiff must show that defendant failed to exercise the kind of care a reasonable person in the circumstances would. Thus, in order to establish the defendant’s liability, the plaintiff would need to show that the harm was foreseeable. The foreseeability requirement would need some substantiation that an ex ante risk was being imposed on those exposed to a product or substance. Thus, having a disease rate elevated above background would be one kind of evidence to show foreseeability, but this is a separate issue from showing causation.
226 See his discussion at Hill, "Environment and Disease," 19.

Between the two world wars there was a strong case for emphasizing to the clinician and other research workers the importance of not overlooking the effects of the play of chance upon their data. Perhaps too often generalities were based upon two men and a laboratory dog while the treatment of choice was deduced from a difference between two bedfuls of patients and might easily have no true meaning. It was therefore a useful corrective for statisticians to stress, and to teach need for, tests of significance merely to serve as guides to caution before drawing a conclusion, before inflating the particular to the general.
is a special case of inference to the best explanation. The statistical evidence shows why one conclusion is the best explanation.227

Causal inference is needed in science. Inference to the best explanation is the reasoning process by which scientists infer that a causal relation holds between exposure and disease. There are a variety of kinds and patterns of evidence that could be used in inferences to the best explanation to conclude that a causal relationship more likely than not existed. Statistical studies of one kind or another are merely one means (one pattern of evidence, if you will) of supporting causal inferences, not the only one and not a necessary form of inference that every causal inference must have.

It is a reasonable conjecture that at least some courts have mistaken an important species of causal inference for the genus or necessary condition of all causal inferences. Instead, they should recognize that the fundamental issue is the causal relationship between exposure and injury and that there are a variety of kinds and patterns of evidence that would license an inference to such conclusions. A variety of patterns of evidence—utilizing all the kinds of evidence available in a particular case—can support causal inferences.

DEFENSE CONTRIBUTIONS TO THE ABOVE ARGUMENTS

This book constitutes an essay about institutions and how administrators of the law shape and mold it by their decisions. It is not an essay that aims to assign blame. Quite the contrary, the tasks the Supreme Court gave federal judges are complex and difficult for judges with their generic and typically nonscientific education. It does not prepare them well for such tasks. If they err, it seems that such mistakes could occur because of too little acquaintance with the relevant fields. I do not assume that this is a deliberate strategy to protect one party at the bar; this would be an abandonment of their responsibilities as fair administrators of the law.

However, in the adversary system there are at least two sides to every dispute judges must adjudicate. If courts have erred in their rulings, their ideas about an appropriate test for scientific evidence have probably not, like Athena, been born fully mature from the head of a Zeus-like judge. It is likely that judges have been influenced in their rulings (largely against plaintiffs) by adversaries on the other side. Moreover, evidence is beginning to accumulate that industries, trade groups, and their lobbying firms have tried to construct a view of science that permits them to reduce or avoid responsibility for the adverse effects of their products. Often these arguments have been used in regulatory settings, but similar points are suggested in tort courts.

I do not make this point a major theme of the book, however, because that would change the nature of the project, take considerable documentation (but some of that has been done), and be a distraction from the major institutional concern: wherever judges acquire their views about scientific evidence, they err in ways that profoundly affect the law.228

Nonetheless, there is evidence that defense attorneys and experts are constructing views of science that are contrary to those of consensus scientific committees and many (but perhaps not all) scientists. Much of this began with the tobacco industry. The tobacco industry was advised by Hill and Knowlton, a public relations firm, to emphasize three points about any relationship between smoking and lung cancer. "That cause-and-effect relationships have not been established in any way; that statistical data do not provide the answers; and that much more research is needed."229 The primary means by which such a claim would be defended would be to invoke an account of causation so stringent that there could be few causal relationships in the world. Moreover, because of court documents made public through litigation, researchers now know that their generic strategy was to create doubt about scientific evidence that their products caused harm. "Doubt is our product" as one document puts it.230

Other industries appear to have followed the lead of the tobacco industry in preventing information reaching the scientific literature or mischaracterizing it. Such strategies have been mounted on behalf of asbestos,231 vinyl chloride,232 lead,233 the general chemical industry (fighting the Delaney Clause),234 and benzidine dyes.235 It would be surprising that similar arguments were not used in toxic tort cases, as they have had some success in the larger area of public discourse.

In addition, there is a patina of plausibility to a claim that "more evidence is needed," because most scientific papers conclude with similar claims about the subject under consideration. Moreover, why would not judges (or regulators for that matter) not want to be more certain about scientific conclusions before incorporating them into the law? Such arguments rest on a certain ideal and are quite tempting, as we saw in Chapter 5. Yet they can be sufficiently misleading.

---

227 Harman, "Inference to the Best Explanation," and Wright, Practical Reasoning, 175–184.


231 Brodeur, Outrageous Misconduct, and Barry Castleman, Asbestos: Medical and Legal Aspects (New York: Pantheon Books, 1985).

232 Markowitz and Rosner, Deceit and Denial, 226–233, 300–301.

233 Markowitz and Rosner, Deceit and Denial, 300–301.


that they constitute their own version of junk science. Evidence is beginning to accumulate that similar strategies are being utilized in toxic tort cases.\textsuperscript{236}

In closing, I consider two broader issues suggested by the courts that can be addressed given the resources of this and previous chapters.

### THE INTELLECTUAL RIGOR TEST

The idea that courts should require of expert testimony the same “intellectual rigor” in court as in the field has become a frequently cited guide for judges to utilize in assessing expert testimony. However, whether the “intellectual rigor” idea is a desirable one or not depends on how it will be understood and utilized.

Surprisingly, there are at least two different concerns about it. On the one hand, the authors of Modern Scientific Evidence characterize it as “dangerous.” Their concern is that because it is discipline specific, experts from fields that lack sufficient rigor might not be reviewed carefully enough by judges.\textsuperscript{237} The intellectual rigor of the field might be insufficient. On the other hand, it tempts courts to emphasize the “rigor” too much and risks being overly stringent. This second concern seems more likely in toxic torts, because there seems to be little doubt about the quality of the pertinent scientific fields. If courts insist that an expert before testifying on causation in toxic tort cases must support his views by multiple kinds of tests and multiple kinds of evidence of the sort illustrated in textbooks or that were insisted upon by the Wade-Greene court, this would be at odds with much scientific practice and would erect nearly insuperable barriers for most plaintiffs. As we have seen, scientists tend to rely on a wide range of data in making inferences about the toxicity of substances, but without having the fullest panoply of evidence available in most cases. Scientists tend to be more flexible in reasoning about the toxic effects of substances than some idealized textbook or commentator views might suggest. Placing too much emphasis on “rigor,” an interpretation that some courts might already have done, would not conform to scientific practice and it would have undesirable effects on the tort law.

By contrast, if courts permit experts to testify on the basis of the weight of the evidence available to them, as, for example, toxicologists do when asked to make judgments about the likely causes of disease or physicians do in diagnosing the causes of disease, this would be a much better application of the intellectual rigor test.\textsuperscript{238} If courts emphasize the similarities between inferences experts

draw in the courtroom and in their out-of-court scientific profession with all the variety and sensitivity this involves, this would be a much more defensible interpretation.\textsuperscript{239}

Finally, courts should preserve the legal distinction between a preliminary review of the evidence to assess its reliability, and extremely rigorously arrived at scientific conclusions on the same issues that might make their way into textbooks. Judges— even those assessing the intellectual rigor of scientific reasoning and methodology—must merely conduct a preliminary review of an expert’s reasoning and methodology for it reliability and relevance to the facts of the case.

### PURSUIT OF TRUTH AND JUSTICE IN TORTS

Recall that in his concurring opinion in Joiner Justice Breyer noted two goals that the Federal Rules of Evidence should serve and that should guide judges in reviewing the admissibility of scientific evidence and expert testimony: truth and justice—two of the great concerns of human institutions (Chapter 3, 88-89).\textsuperscript{240} He explicitly seeks to ensure that the powerful engine of the tort law is directed toward the “right substances,” the ones that in fact cause harm, but also does not “destroy the wrong ones.” This is a concern that courts, after the results of a jury verdict, should yield a verdict, based on a correct (or reasonable?) scientific view about the toxicity of a substance and not mistakenly have eliminated beneficial, but (comparatively) harmless, products. He emphasizes the social good that has come and can come from the products of our technological society. He suggests that admissibility and trials be conducted so that substances that are part of beneficial products and that in fact do not cause adverse effects in humans do not become mistakenly judged legally as human toxicants. This would be a mistaken legal verdict for plaintiffs against defendants.

Whereas he also calls attention to the tort law’s concern for justice and court procedures for finding truly harmful substances by means of its legal proceedings, his greater emphasis—that courts should avoid false positives—does several things that can be problematic. It reinforces the asymmetric scientific norm to prevent false positives. However, rather than needing support, this scientific norm about mistakes probably merits some countervailing attention in the law. To legally reinforce the scientific norms about mistakes further burdens plaintiffs.

Moreover, there is another possible legal mistake that appears to receive less attention: legal false negatives—that is, a mistaken legal outcome finding that

\textsuperscript{236} Michaels and Monfort, "Manufacturing Uncertainty," S44.

\textsuperscript{237} Feigman et al., Modern Scientific Evidence, 48.

\textsuperscript{238} Kasirer and Cecil, "Inconsistency in Evidentiary Standards for Medical Testimony," 1382–1387. Note also a very recent decision in which the court recognizes that to be admissible, an expert’s "analyses, inferences and extrapolations connecting the science to the witness’s conclusions must be of a kind that a reasonable scientist or physician would make in a

\textsuperscript{239} This is a point we considered in Chapter 5.

\textsuperscript{240} Joiner v. General Electric Co., 118 S.Ct. at 520.
plaintiffs are not harmed by exposure to a substance when in fact they are. He is silent about this possible mistake. Does he undervalue admissibility procedures that would prevent errors that disadvantage plaintiffs? Does he inadvertently undervalue the importance of justice for plaintiffs, despite his articulated point about justice?

These topics raise a more difficult and subtle point about truth and procedures for arriving at it. Many of us might be tempted to believe that factual truth about the toxicity of substances should precede or be a prerequisite to determination of just compensation for plaintiffs. Why should we not insist on the scientific truth about toxicity in torts?

Although this question is easily posed, its answer is not so straightforward. For one thing, the correctness of an expert's conclusion is not the goal of an admissibility hearing, only the reliability of his or her testimony is. For another, we may be tempted to think that the truth about the toxicity of substances is easily knowable and relatively quick to obtain. Both claims about the science are not true for many kinds of toxic substances (Chapter 5).

To get at the truth about toxic substances for legal purposes, as a community we must utilize, first, the procedures of different areas of biology and toxicology, and, second, legal procedures. However, while we are in the middle of scientific and legal debates, we must rely on scientific and legal procedures to guide us.

What are the biases or tendencies of different procedures during the period before the science is fully settled and against the background of various practical hurdles to establishing scientific claims? How will these affect the law? What are the tendencies of scientific procedures? Is the process biased toward one outcome or another? These are, thus, questions of process tendencies both in science and in the science-law interaction in toxic tort cases. How will the process biases manifest themselves when science is used in the law?

Consider some of the evidentiary tendencies in scientific practices leading to consensual truth concerning a toxic substance. First, if a scientist is given a substance and asked if it is toxic, she might respond that she does not know. Or perhaps, invoking the memory of Paracelsus who is responsible for the aphorism that "the dose makes the poison," she might note that at some concentrations it will be toxic and at some concentrations it will not be, but she would not know what those were without investigation. In short, she would be agnostic about such claims in the absence of evidence.

Second, as I argued elsewhere and reviewed in Chapter 4, if a scientist attempts to study a substance's toxic effects with epidemiology, unless she has a

very large sample population, conceptually she will be forced to choose toward which kind of mistake her study has a predisposition—a false positive or a false negative. That is, if her study shows an elevated risk from exposure, is it truly an elevated risk or an anomaly of the study? If her study yields a no effect result, is it truly no effect or falsely negative? And beyond that, can it provide any evidence at all that there is no harm from exposure? The smaller the samples used in the study, the greater these problems will be.

If our scientist tries to answer the question with animal studies there are additional problems. One is the extrapolation from high-dose results in animals (needed to have a study of any sensitivity) to low-dose effects in animals. Another is the extent to which positive animal studies provide evidence of likely human harm, although as I argued in Chapter 4 scientists understand such studies so as to provide reasonable estimates of these effects. Even though scientists routinely utilize animal studies, skeptics have intellectual space to challenge the inferences. Our scientist, again, might guard against false positives and might withhold judgment about human harm based on positive animal studies. However, withholding judgment—even if there is sufficient evidence for the legal conclusion—creates the possibility that there may be enough uncertainty to undermine the case.

Third, if human studies are positive, scientists may be tempted to search for positive animal studies to help confirm human evidence and to help provide a model for the disease process. For understanding the disease, even multiple human results may be insufficient. More study is needed. If animal studies are positive, but human studies are not, does this show there is no human effect or that the human studies were insensitive, poorly designed, too short, or just a statistical aberration? Human and animal studies may not agree—one or the other may be positive, the other negative. This could cloud the picture and further delay judgment about human effects.

Fourth, even if both human and animal results tend to agree, some scientists might still want to understand the biological mechanisms of action in order to eliminate mere statistical associations and in order to understand the disease process better. The ideal would be to understand the disease process "all the way down" from external exposure, to exposure at the target organ, to metabolic activation and distribution throughout the body, even to molecular effects resulting ultimately in disease or death at the cellular level. Moreover, with diseases such as cancer there may be multiple molecular changes before the disease is manifested clinically. However, a demand for mechanistic understanding is not likely to be met for most substances (Chapter 4 and above).

There also may be demands for multiple kinds of evidence in support of expert testimony. This, too, appears to be rarely met even for scientific purposes. It will be rarer yet in toxic tort law with its time frames. To permit such demands

---

241 In addition, recall that some scientists begin with more controversial assumptions. If judges were to be convinced by defendants to adopt the presumption that substances caused no human health harms until overwhelming evidence established evidence to the contrary, this would raise quite high barriers to establishing appropriate toxicity claims.

242 Cranor, Regulating Toxic Substances, 31–47.
These arguments are correct—that scientific processes are biased by the confirmation of prior beliefs and presuppositions. More specifically, in the early stages of inquiry, the scientist's position is likely to be that which is consistent with the pre-existing body of evidence. As the scientist progresses in the research, making the process more difficult, the existing knowledge structure makes it more likely that the scientist will encounter evidence in line with their own prior beliefs and that of others. This is particularly true in the courts, where judges are not only experts in their field but also have pre-existing opinions on the issues. The result is often the confirmation of existing positions, resulting in a tendency toward the right substances and decisions that are in line with the existing knowledge and beliefs. As the scientific community progresses, it becomes increasingly difficult to change a scientist's views, and this tendency toward confirmation and support for existing beliefs becomes more pronounced.

In summary, the judicial process is similar to the scientific process in that both are based on the confirmation of prior beliefs and presuppositions. However, the judicial process is more likely to result in the confirmation of existing positions due to the difficulty in changing judges' views and the pressure to maintain consistency with existing knowledge and beliefs. This tendency toward confirmation and confirmation bias is particularly pronounced in the courts, where judges are not only experts in their field but also have pre-existing opinions on the issues. The result is often the confirmation of existing positions, resulting in a tendency toward the right substances and decisions that are in line with the existing knowledge and beliefs. As the scientific community progresses, it becomes increasingly difficult to change a scientist's views, and this tendency toward confirmation and support for existing beliefs becomes more pronounced.
testimony that are falsely positive and negative studies at issue in expert testimony that are falsely negative (as well as those that simply fail to show evidence of no adverse effects). It is easy to argue for pursuing truth through application of the Federal Rules of Evidence and interpretations of Daubert, but it matters importantly how this is done and what mistakes judges tolerate or try to prevent by reviewing expert testimony.

CONCLUSION

If courts are to accurately review scientific testimony and its foundation, to fairly perform their admissibility responsibilities and to enhance the possibility of justice in torts, they must become knowledgeable, sensitive, thoughtful consumers of scientific studies and inferences. This will not make their tasks easier; in all likelihood, it will be more difficult to conduct reviews of evidence well. Judges will first need to avoid using overly simple and restrictive guides for reviewing evidence. The court decisions considered as examples in this chapter illustrate a tendency to look for comparatively simple heuristics and procedures to assist their work, something that is less true of scientists as they study the complex biological world. The evidentiary picture in science is more complicated than such simplified rules can reasonably accommodate. However, if the law is to be more compatible with science, courts will need to conduct admissibility reviews that better recognize the multifaceted, varied nature of scientific evidence and reasoning. Thus, they also need to be more sophisticated about subtle process tendencies of science that can undermine legal goals. To better serve science and law, they should avoid:

- Demanding ideal evidence.
- Insisting on epidemiological studies or placing misleading restrictions on them.
- Demanding mechanistic evidence.
- Routinely excluding animal studies and structure-activity evidence.
- Routinely excluding relevant case reports (they need to learn the indicia of good reports).
- Excluding individual kinds of scientifically relevant evidence because by themselves they are insufficient for a causal conclusion.

More subtle and sensitive analyses will be needed in place of all of this.

Courts also will need to develop greater sensitivity to the process tendencies of science in order to review it more accurately and to be fair to litigants from both sides. Inter alia this includes the need to

- Recognize the difference between “no evidence of an effect” and “evidence of no effect.”


246 Benzene appears to be toxic down to 1 ppm and possibly toxic to .1 ppm (see Infante, "Benzene and Leukemia," 253–262, and Qing Lan, Luoping Zhang, Guilan Li, Roel Vermeulen, Rona S. Weinberg, Mustafa Dosemechi, Stephen M. Rappaport, Min Shen, Blanche P. Alter, Tongli Wu, William Kopp, Suramya Waidyanatha, Charles Rabkin, Weihong Guo, Stephen Chanock, Richard B. Hayes, Martha Linet, Sungkyoon Kim, Songnian Yin, Nathaniel Rothman, Martyn T. Smith, "Hematotoxicity in Workers Exposed to Low Levels of Benzene," Science 306 (2004): 1774-1776), whereas arsenic, which for years was believed to be comparatively safe at 50 parts per billion (ppb), is now judged to be toxic to levels perhaps as low as 10 ppb (or even less). See the U.S. Environmental Protection Agency, at http://www.epa.gov/safewater/arsenic.html, for discussion of the drinking water standard for arsenic.

247 Some deterrence will remain, of course, simply because the threat of suit is available to those who believe they have been wronged. However, the deterrent effect is stronger, if meritorious cases succeed.
Recognize the variety of ways that studies can be misleadingly negative, because they are too small, too short, too specific for subspecies of disease, or conducted too long after exposure has ceased to identify real adverse effects.

- Understand the "healthy worker" effect.
- Exclude evidence on the basis of scientific irrelevance only with the greatest of care.
- Recognize other ways that the process tendencies of science can unfairly impact admissibility reviews.
- Review a scientist's evidence as an integrated whole.

The various mistaken reasons some courts have given for refusing to admit expert testimony pose one serious kind of problem — judges appear to be deciding issues about the scientific relevance of evidence contrary to good scientific practices. Such decisions are even worse, however, when they are considered against the background of pragmatic barriers to the identification of causal relations that exist. The result is an unfortunate synergism between mistaken court reasons for excluding evidence and pragmatic barriers to the discovery of causal relations. This synergism in effect greatly heightens the barriers for injured parties to successfully bring their cases before a jury and have a public discussion of the human and institutional relationships that led to exposure and possibly disease. I close this chapter with some highlights of the adverse consequences for the law.

The more demanding scientific requirements courts place on the evidence that forms the basis of expert testimony, the more this heightens plaintiffs' barriers to the law simply because so little is known about the universe of chemical substances. The less subtle judges are about science, the more likely it is that the process tendencies of science or uninformed skepticism will inadvertently trump the law's mandated evenhandedness in making admissibility decisions.

In addition, ignorance about the universe of substances burdens the extent to which it is possible in a particular case to obtain minimal evidence about whether exposure to a product has contributed to a plaintiff's disease. Thus, the more courts constrain with considerable specificity the kinds of evidence that must be used in support of testimony (instead of permitting scientists to utilize the wide range of evidence they would ordinarily rely on for causation judgments), the more this restricts scientists' tools of the trade and reduces plaintiffs' access to the legal system. Given the widespread ignorance about substances, a better presumption would be that there might be little or no human epidemiological evidence and no mechanistic information about the effects of any particular substance. When either is available, this is more a result of good fortune than a routine matter.

The higher the evidentiary standard that must be met before expert testimony is admitted, the burden of production is carried, or a scientific finding is justified to the satisfaction of a jury, the easier this makes an adversary's task because the opposing party may be "more inclined to rest on the non-credibility of the proponent's proofs, and less inclined to produce affirmative evidence." Such standards will not minimize the weighted number of mistakes between plaintiffs and defendants, and will reduce incentives for opponents to produce their own affirmative evidence frustrating two aims of the tort law. The more legitimate scientific information courts have to arrive at a decision, the more accurate outcomes likely will be.

There is a further and important concern. By their rulings on the admissibility of expert testimony and its scientific foundations, courts are at least implicitly making policy for the tort law, but making it in ways that are hidden from the public and perhaps even from large numbers of the bar. These are not merely policies concerning admissibility; courts are also constructing substantive law. For example, when a court precludes a case from going to trial unless there is statistically significant epidemiological evidence showing a RR > 2, such decisions in effect are predisposing legal outcomes, especially for those whose evidence may not match the "judicially imposed standard. Admissibility rulings can have several broader policy implications:

- They can create a legal view of science that is at odds with many in the scientific community.
- They can establish a policy that precludes plaintiffs from recovery for harm suffered if they cannot provide specific kinds of evidence. Thus, a person's injuries are not legally worthy of just rectification unless there is such evidence. When this occurs, the law does not follow where science leads, but refuses to recognize a case for recovery, unless there is a special kind of evidence.
- There also can be different standards for admissibility in different courts. To the extent that there are, litigants are treated differentially depending on which circuit their case is heard. This invites "forum shopping" by litigators to find the most favorable venue for their trial.

Short of major modifications in tort law liability and consistent with the requirements of Daubert, courts can take some steps to modestly, but importantly, address the problems discussed earlier. To do this, however, they must

---

251 Berger, "Upsetting the Balance Between Adverse Interests," noting a number of policy decisions that are implicitly made.
correct some of the apparent mistakes noted in this chapter. Were they to do this, courts’ admissibility decisions would not decide the cases, but they would permit testimony to be presented and critiqued before a jury.253

The gatekeeping rules and the public values at stake, however, increase courts’ responsibilities to perform their tasks well to achieve the goals of torts within existing liability rules and available scientific evidence.254 Failures in these tasks do not seem to be options; they are too threatening to the legitimacy of the law and the possibility of justice. To fulfill this responsibility and avoid some of these tendencies from materializing, courts need a better understanding of institutional tensions between science and the law and how better to review the scientific foundation of testimony.

253 There is an important asymmetry between these two examples. By excluding cases from trial unless they have expert testimony that rests on very good human evidence, courts are in effect concluding that cases that fall short of this scientific standard as without legal merit. They decide that class of cases. By permitting experts to testify on the basis of less direct kinds of evidence, they are not necessarily deciding such cases; they are merely letting the proceedings continue further to be considered by juries and subject to later defense objections and motions and any appeals.

254 In Chapter 8, I raise the question of whether current liability rules are adequate.

7

Enhancing the Possibility of Justice under Daubert

If some courts have been using unduly constrained, idealized, or overly simple heuristics for reviewing scientific testimony on causation, how might they conduct this task differently? How can they better address complex patterns of evidence? Can their admissibility decisions better serve the aims of both law and science?

Addressing these questions is the subject of the current chapter. I briefly consider the use of court-appointed experts, and then discuss an alternative, building on suggestions made by the Supreme Court in Kumho Tire and the Third Circuit Court of Appeals. After discussing this proposal, I consider some more nuanced patterns of evidence from consensus scientific committees to illustrate some complex patterns of evidence that courts might face, should be able to recognize and review favorably. I then present some decisions in which judges have recognized the subtlety of issues they faced or in which they addressed well the shortcomings of studies or reasoning against which they needed to guard. Toward the end of the chapter, I revisit some decisions discussed in Chapter 1 to suggest more specifically some of the problems they raise.

Courts have choices in how they implement Daubert and its progeny. They can unduly restrict scientific testimony, or fail to recognize more subtle scientific mistakes that can affect litigants. By contrast, they could review admissibility decisions on expert testimony to assess whether they fall within a “zone where reasonable scientists would disagree” (adopting a guiding heuristic from the Court in Kumho Tire).1 Were they to do this, it is reasonable to expect several consequences to result.

Admissibility decisions should be better founded scientifically than at present and comport better with how scientists themselves assess evidence. Such outcomes would increase the acceptability of admissibility decisions within the scientific community and reassure respectable scientists who testify that their testimony will not be judicially condemned as inadequate.

1 Kumho Tire v. Carmichael, 526 U.S. at 153.