American Journal of PUBLIC HEALTH

EMAIL CONTENT DELIVERY e-toc

QUICK SEARCH: [advanced]
Author: Keyword(s):
Go
Year: Vol: Page:

HOME HELP FEEDBACK SUBSCRIPTIONS ARCHIVE SEARCH TABLE OF CONTENTS

Access via PAY PER ISSUE

July 2005, Vol 95, No. S1 | American Journal of Public Health S121-S128
© 2005 American Public Health Association
DOI: 10.2105/AJPH.2004.044735

This Article
- Abstract FREE
- Full Text (PDF)
- Submit a response
- Alert me when this article is cited
- Alert me when eLetters are posted
- Alert me if a correction is posted

Services
- Similar articles in this journal
- Similar articles in PubMed
- Alert me to new issues of the journal
- Download to citation manager
- Cited by other online articles

Google Scholar
- Articles by Cranor, C.
- Articles citing this Article

PubMed
- PubMed Citation
- Articles by Cranor, C.

PUBLIC HEALTH MATTERS

Scientific Inferences in the Laboratory and the Law

Carl Cranor, PhD

The author is with the University of California, Riverside.

Correspondence: *Requests for reprints should be sent to Carl Cranor, PhD, Department of Philosophy, University of California, 900 University Ave., Riverside, CA 92521 (e-mail: carl.cranor@ucr.edu).*

ABSTRACT

Following the Supreme Court's *Daubert v Merrell Dow Pharmaceuticals, Inc* decision, courts have struggled in reviewing scientific evidence and appear to have appreciated insufficiently the nature of scientific reasoning. To assist their evidentiary reviews, courts need to appreciate both scientific complexity and ignorance about human toxicity caused by the universe of chemical substances.

Conscientious, well-motivated, respectable experts can come to different conclusions about scientific evidence without making obvious mistakes and without being charlatans. So that justice is done between parties, courts need to allow for reasonable scientific disagreement to avoid excluding from trials respectable experts and all relevant scientific evidence.

The public health community can assist courts to understand ranges of scientific evidence and to recognize the reasonableness of scientific disagreements.

INTRODUCTION

The US Supreme Court in *Daubert v Merrell Dow Pharmaceuticals, Inc* and its sequelae, *General Electric Co v Joiner* and *Kumho Tire Co v Carmichael*, sought to ensure a closer relation between the law and the science in toxic tort and other legal domains.[1–3] As they implemented these decisions in toxic tort suits, the Supreme Court, lower courts, and some state courts have struggled in reviewing scientific evidence.

I briefly review here the decisions that changed the law on admitting scientific evidence for consideration at trial, and note the difficulties courts have had in excluding scientific evidence (animal studies and case studies in particular), in recognizing the structure of scientific inferences, and in demanding human evidence. (These particular issues are discussed in more detail elsewhere.[4,5]) I then consider forms of inference that are fundamental to scientific reasoning: for epidemiological studies, for more complex patterns of evidence, and for case reports that conclude that substances are toxic to humans.[5] Because very little is known about most substances registered for use in commerce, inferences about their causal properties and courts' tasks in assessing those inferences have been made more difficult. Thus, quite frequently scientists must utilize the evidence available, which may not be the best, yet when combined it may be sufficient to support causal inferences.

Well-motivated, respectable scientists can come to different conclusions about evidence without making obvious mistakes and without being charlatans. To avoid excluding respectable experts from trials, judges need to allow for these differences, letting juries hear all relevant scientific evidence. Will this be sufficient to deliver justice? I do not consider that question here, but letting in the scientific evidence is a first step.

For the public health community, admissibility of scientific evidence in tort trials is important. By authorizing compensation for those harmed by others, the law of torts assists in "regulating" risks. Secondarily, successful tort suits, to some extent, deter others from similar conduct; unsuccessful suits can send the opposite signal. Neither of these results is possible if suits are halted short of trial, as happens when judges exclude scientific evidence and testimony offered by experts that would have linked the tort or injury to a cause. If too few meritorious cases go forward, the tort law tends to become less protective, resulting in increased risks to the public's health.

When courts mishandle scientific evidence, distorting the science used in the law, they discourage scientists from participating in trials–something they are reluctant to do in the first place. If scientists avoid this important role, contrary to the Supreme Court's aim in *Daubert,* court decisions may not benefit from an accurate view of science. In addition, when courts make scientific mistakes they may perpetuate them through precedents and examples followed by other courts.

The public health community, alert to mistaken admissibility decisions and calling attention to good scientific inferences, can assist judges. This will serve justice, help to deter wrongs, and protect public health.

THE LEGAL BACKGROUND

In *Daubert*, federal trial court judges were given a heightened duty to review scientific expert testimony and its foundation prior to trial that "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."[1] *Joiner* and *Kumho*, addressing procedural issues, gave judges great latitude in determining both how decisions to admit scientific evidence into trial are to be made and in making them. A judge should determine whether an expert's testimony falls inside "the range where experts might reasonably differ" in order to be admitted.[3] These admissibility decisions can be appealed and subsequently overturned only if the judge "abused his/her discretion."[2] The abuse-of-discretion standard makes it extremely difficult for a party losing an admissibility hearing to appeal successfully, meaning less appellate court policing of district courts than there might have been. Trial courts consequently have greater responsibilities *on their own* to ensure that their decisions are reasonable and fair.

But many judges have found it hard to assess different kinds of scientific evidence and forms of scientific inferences, thus threatening the reasonableness of trial decisions.

COURT REVIEWS OF SCIENTIFIC EVIDENCE

Court-Excluded Evidence

Animal studies and human case reports are *relevant* to assessing the toxicity of substances to human beings, in science and the law. Evidence is scientifically and legally relevant to making a causal judgment, if it tends to make that judgment more or less correct.[6–11] The general causation question in toxic tort cases is whether a substance can be harmful to human beings. The specific causation question is whether defendant's substance in fact caused plaintiff's injuries. How much each piece of relevant evidence can contribute to an explanation depends on the other evidence available for making such judgments and how the evidence "fits together."[12] Just how much a scientist judges each piece of evidence to assist in explaining whether or not exposure contributes to a disease may depend on the scientist's discipline as well as on his/her individual judgment.

Animal studies. Controlled animal experiments, in which randomly selected animals are exposed to different

doses of a possible toxicant (experimental groups) and compared with a randomly selected group not so exposed (the control group), are highly regarded and routinely relied on by the scientific community for assessing the causal properties of substances. In experimental design, animal studies resemble clinical trials. Judges, however, reviewing the scientific foundation of expert testimony, have been overly critical of animal studies, failed to understand them, or dismissed them outright for what appear to be poor reasons.[13–18]

Animal studies are not mathematically certain guides to the toxicity of substances in humans (and this is certainly not required by the law of torts). They are, however, scientifically relied on for identifying carcinogens and other toxicants, being in most cases, more likely than not, reliable for general causation in the law.[4] They can pose a problem for non-scientists who may be misled by superficial physical differences between animals and humans (rodents are smaller, have fur, whiskers, funny noses, tails, a lesser brain, and behave differently), and even by some biological differences (size, lifespan, metabolic rate, and some different organs), while overlooking important and deep biological similarities.

Well-respected researchers have articulated reasons for the presumptive strength of animal studies for identifying carcinogens. The late David Rall, former head of the National Institute of Environmental Health Sciences, noted "experimental evidence to date suggests that there are more physiologic, biochemical, and metabolic similarities between laboratory animals and humans than there are differences."[19] Furthermore, fundamental "biological processes of molecular, cellular, tissue, and organ functions that control life [such as sodium and potassium transport and ion regulation, energy metabolism, and DNA replication] . . . vary little in the aggregate as one moves along the phylogenic ladder."[20]

And finally, "The more we know about the similarities of structure and function of higher organisms at the molecular level, the more we are convinced that mechanisms of chemical toxicity are, to a large extent, identical in animals and man."[20]

Patterns of carcinogenic responses in mammals suggest that a chemical causing cancer in one species will likely do so in another.[21] Independent researchers, the National Academy of Sciences, and the US Environmental Protection Agency (EPA) concur in many of these judgments.[22–26] A group of researchers from the National Institute of Environmental Health Sciences, the National Cancer Institute, the National Institute for Occupational Safety and Health, the University of North Carolina, the International Agency for Research on Cancer (IARC), and Linkoping University (Sweden) have concluded that experimental results, in particular long-term carcinogenicity tests, have proved to be valid predictors of human risk.[26]

The legal news for animal studies is not all bad. Some courts have recognized the scientific value of animal studies and have overcome a substantial record that developed from the unusual evidentiary records in the Agent Orange and Bendectin cases. The Third Circuit Court of Appeals in *In re Paoli Railroad Yard* has a discussion of some of these cases.[27]

Case studies as evidence. Courts have a mixed record on admitting case studies as relevant evidence in court

cases about toxic torts or vaccine injuries. Out of 64 ordinary federal tort cases, 14 courts accepted case studies as relevant evidence and 36 rejected them, with the remainder not making them the focus of discussion.[28] Sometimes individual court judges dismiss case studies out of hand.[29] In an additional 33 federal claims court cases under the National Vaccine Injury Compensation Program, federal magistrates all appear to regard case studies as scientifically relevant for causation, in judging causation in vaccine cases that are not automatically compensable in so-called "on table injuries." For plaintiffs to receive compensation for injuries not included in these tables, they must make a scientific case for compensation reasonably similar to those in tort cases.[28] (Most of the magistrates in the vaccine program accept *Daubert* criteria.) Thus, the record of federal magistrates in Federal Claims Court cases in recognizing case studies as evidence is better than ordinary federal judges in tort cases, but I consider only tort cases (the vaccine injury cases have some special features, but not on the scientific relevance of case reports). The vaccine injury magistrates appear to be more familiar with a wider range of scientific evidence than less specialized district court and appellate judges, which may account for differences between the judges.

Case studies are in general not mathematically certain guides to causal relationships, because some are good studies, others poor. Some merely report an association between an exposure to a drug, poison, vaccine, or other substance and the onset of disease, without analysis about whether the exposure might explain the disease. For case studies to provide evidence for causal relationships, they should satisfy analytical criteria that resemble standard scientific inferences to be discussed in the next section.[10,11]

The World Health Organization and the US Institute of Medicine both use case studies as evidence in making causal inferences (especially about reactions to drugs, vaccine, and poisons, but also about carcinogens). They have explicit criteria by which to judge when case studies constitute good evidence for human causation.[30–32] When judges decide whether studies can assist the jury in coming to its decision and thus be put in evidence, they too must consider whether the case studies provide good support of causality.

Causal Inferences

Lawyers and judges do not always concur with the scientific community's acceptance of animal and case studies. The legal community, moreover, may not appreciate the numerous ways in which generically relevant evidence can fit together to support causal inferences. Consider a dramatic example: how scientists pieced together evidence to come to a powerful conclusion about the structure of DNA, the foundation of many recent developments in biology.

> [Irwin] Chargaff's discovery that there are approximate regularities in the relative proportions of adenine and thymine, guanine and cytosine in DNA is hardly, by itself, strong evidence that DNA is a double-helical, backbone-out macromolecule with like-with-unlike base pairs; Franklin's X-ray photographs of the B form of DNA are hardly, by themselves, strong evidence that DNA is a double-helical, backbone-out macromolecule with like-with-unlike base pairs. That the tetranucleotide hypothesis is false is hardly, by itself, strong evidence that DNA is a double-helical, backbone-out macromolecule with like-with-unlike base pairs, and so on. But

> put all these pieces of evidence together, and the double-helical, backbone-out, like-with-unlike base pairs structure of DNA is very well warranted (in fact, the only entry that fits).[12]

This esoteric but fundamental discovery of the structure of DNA is likely much more difficult than inferring whether an exposure to a substance has made a causal contribution to disease. Yet it points to the necessity for scientists, whether in the courtroom or laboratory, to piece together scientifically relevant evidence in a plausible way to explain phenomena.

Courts, in ruling on the admissibility of scientific evidence, may fail to appreciate (a) the fact that causal conclusions result from inferences, (b) the form of inference used to make such judgments, (c) the greater complexity of some inferences, and (d) the difficulties in assembling evidence needed to make such inferences.

How can we describe the inference structure that leads to a causal conclusion, such as a well-warranted explanation of the structure of DNA or toxicity causation? Some philosophers of science and scientific methodologists call them "inferences to the best explanation," "diagnostic induction," "diagnostic arguments," or "differential diagnosis."[8,33–37] Regulatory agencies use the term "weight of the evidence" for the same inference structure.[22]

An inference to the best explanation involves a process of reasoning. Scientists consider different plausible explanations of the phenomena in question; take into account all the relevant and available evidence that assists explanation, as well as evidence that might help distinguish between explanations; piece the relevant evidence together in the most plausible way; and then judge which of the competing explanations has the best support. Several factors guide a scientist's reasoning process: skill in making plausibility judgments, background knowledge, understanding of possible causes, or explanations of circumstances that need explanation. The plausibility of a casual inference depends importantly on an expert's skill, understanding, and knowledge of the evidence in question; there is no set of necessary and sufficient conditions that guide causal inferences.[5,34,38]

In toxic tort cases, the diagnostic question is, "Does the available relevant evidence, taken together, better support the claim that the exposure contributed to the disease in question or does it better support some alternative explanation of the person's disease?" For the admissibility decision—whether or not to admit the expert's opinion in the trial—the question is whether the scientific expert's reasoning about the phenomena falls within the range where respectable scientists could reasonably disagree.[3]

Inferences to the best explanation have posed problems for the courts. The district court in *Joiner v General Electric* seemed to analyze and reject each piece of evidence individually as inadequate to support plaintiff's ultimate conclusion, while not assessing the body of the evidence taken as a whole.

On appeal the 11th Circuit Court of Appeals, correctly seeing that it was incumbent on courts to evaluate whether all the relevant evidence taken together provided a sufficient basis for admitting scientific experts,

overturned the *Joiner* district court's review of the evidence for admissibility. However, the Supreme Court overruled the 11th Circuit on the procedural issues, and then went beyond procedure and seemingly endorsed assessment of each piece of evidence for how well it supported the ultimate conclusion. The Supreme Court seemed to reject standard scientific inferences.[2,5] Justice Stevens, in dissent, recognized the salience of inferences to best explanation (he called them "weight of the evidence" arguments) and would have remanded the case for a hearing on whether all plaintiff's evidence in support of its explanation of Mr. Joiner's lung cancer satisfied the requirements for admissibility in that case.[2]

It is one thing to reject plaintiff's experts because no one piece of evidence by itself adequately supports the ultimate conclusion that exposure to PCBs contributed to lung cancer (something no scientist would do). It is quite another to claim that the scientific experts' reasoning and methodology, using an inference to the best explanation and considering all the relevant evidence together, supported the explanatory conclusion (something the Court did not analyze). The Court's opinion in *Joiner* risks misleading lower courts, inviting similar mistaken rejections of particular evidence or having a chilling effect on efforts to review scientific evidence in the same way that scientists do.

The structure of inference to the best explanation explains why, in a particular case before a court, there may be evidence that individually does not support the claim that exposure to a substance makes a causal contribution to disease, but when pieced together with other evidence in a scientifically plausible way does support the conclusion.[5] How well the various pieces of evidence fit together is what matters in supporting a causal inference.[5,12] In short, there is no univocal pattern that is always right or that is always required to draw an inference.[4] Both toxicological and epidemiological evidence correctly assembled may support an inference of a substance's toxic effects.

Court-Required Evidence

Contrary to scientific views of relevant evidence, some courts have demanded epidemiologic evidence; a few have required the expert to offer a mechanistic explanation for causation before admitting the expert's opinion in trial.

Mechanistic evidence. A few courts, perceiving the importance of understanding the mechanisms by which diseases occur, have suggested that mechanistic evidence is a necessary condition of evidence that provides a foundation for expert testimony.[39–41] Understanding the mechanism by which exposure to a toxic substance causes disease—e.g., asbestos (mineral fibers) deposited in the lungs causes mesothelioma; particular bacteria cause stomach ulcers—greatly assists causal inferences.[6] However, mechanistic evidence is often not available. Until as recently as 1991, mechanisms to explain beneficial and harmful effects of aspirin were unknown.[42] Moreover, mechanistic evidence is asymmetrical—when present it can greatly strengthen a causal inference, but when absent it does not necessarily undermine the inference.[38]

Epidemiologic evidence. Many courts have required experts to provide a foundation for their testimony using human epidemiologic studies. Michael Green documented this, pointing to a line of cases from the

Agent Orange and Bendectin litigation that he argued "are best understood as adopting an epidemiologic threshold, or more restrictively, a statistically significant epidemiologic threshold, for plaintiffs to establish in order to create a submissible case."[43,44] Some other courts, designating epidemiologic studies among the best kinds of evidence (but asserting they are not absolutely necessary), have, in effect, denigrated other kinds of evidence, such as animal studies and case reports. They have concluded that scientific evidence and expert opinion not based on epidemiology is defective—tantamount to requiring epidemiologic studies.[18,43–45]

The attraction of epidemiological studies is clear. They consider the effects of exposures on humans, the concern in toxic tort suits. To demonstrate a substance's human toxicity, well-designed and executed epidemiologic studies, plus clinical trials, human case studies, and molecular and toxicity data from humans, can be among the best evidence.[4]

Despite the obvious attraction of epidemiologic studies there are various problems, limitations, and shortcomings that affect their usefulness, especially in toxic tort suits. For many substances, epidemiological data simply are not available.[19] If deliberate exposure of people would be unethical, then studies must be conducted when adventitious exposures occur, making it difficult to collect accurate exposure information. And such studies are expensive. Stringent statistical significance criteria, which increase a study's scientific respectability by reducing false positives, also risk being too insensitive to detect the risks in question.[43,46]

Even when a substance contributes to adverse effects, studies too small or too short may not detect them.[47] Human studies delayed until long after exposure can also underestimate risks.[48,49] Courts must thus be especially careful in admitting insensitive or misleading studies because they may be falsely negative, an issue of equal concern in the tort law along with false positives.[4,43,46] Finally, conducting an appropriate epidemiological study is often not an option, for a variety of reasons: too few may be exposed, exposure data may be poor, the diseases may be rare, or studies may be very expensive.[4]

Epidemiologic studies are not controlled experiments. Their reputation in the scientific community may not be nearly as strong as their importance for federal judges. As one National Cancer Institute scientist notes:

> 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. . . . 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.[50]

Although difficult to document, judges may also take negative epidemiological studies at face value, a mistake to which scientists are sensitive: "no evidence of an effect" is not "evidence of no effect." Thus, a consensus scientific body such as the IARC requires that some highly specific and detailed conditions should

be met before making such an inference. Even then the results apply only to the disease endpoint of what was studied and not to unrelated diseases or tumors.[32]

PATTERNS OF EVIDENCE

Causal Inferences in Epidemiology

Epidemiological evidence does not carry conclusions "on its face," but like all scientific evidence underdetermines the explanation; it is logically insufficient to guarantee a particular conclusion in the same way that the premises of a mathematical argument or deductive logic can logically ensure a conclusion.[5,8] Thus, scientists must infer what the evidence shows: "The basic strategy of the practitioner of causal inference is to distinguish among several alternative hypotheses, all of which are underdetermined by the evidence."[51]

Such inferences provide "many opportunities within the practice of causal inference for scientists to hold different opinions about which scientific values are important to the assessment of evidence."[51] Researchers may reasonably disagree on considerations used to assess whether a study shows a causal relationship or on their application of these considerations—strength of association, coherence, consistency, or confounding.[51] One researcher might, for example, place greater emphasis on lack of consistency between one study and others. Another might have a different understanding of biological plausibility, although no one rejects biological plausibility as a consideration.

Differences in judgment can lead to different conclusions about what studies show. Even when competent, well-intentioned, conscientious scientists use identical data and identical generic considerations for interpreting it, they might reasonably and without mistake interpret the data and criteria somewhat differently. They might use somewhat different "rules of inference" for applying the criteria, and, as a result, reach different conclusions.[51] The public health community can help judges understand the origins of scientific disagreement, so that they do not exclude experts whose disagreements are within the boundaries of reasonable scientific inferences.[3]

Causal Inferences from Mixed Scientific Evidence

Scientists, piecing together more complex patterns of evidence—from human, animal, and other studies—to reach conclusions about causation, face more daunting tasks. To illustrate scientific inferences from complex and varied data, I review patterns of causal evidence from which scientists have drawn conclusions about adverse health effects. I utilize consensus judgments about probable or known human carcinogens from prominent national and international scientific bodies—perhaps not totally uncontested (because of scientific disagreement), but the result of extensive scientific assessment and discussion, and carrying widespread endorsement.

A few examples are of known human carcinogens; several are of substances classified as "probable human

carcinogens" by the IARC or substances "reasonably anticipated to be carcinogenic to humans" by the National Toxicology Program. Either assessment entails a greater degree of certainty than "more likely than not"—the standard of proof in the tort law or the less stringent admissibility requirements in torts.[1,27] The former director of the IARC Monograph Series noted that the classification "'probably carcinogenic to humans' would mean that the evidence is just a little short of certainty that the exposure in question has actually caused human cancer."[52]

There is almost no direct human evidence of human carcinogenicity for 4,4'-Methylene bis(2-chloroaniline) (MOCA), used as a "curing agent in urethane and epoxy resin systems" and in making "jet engine turbine blades, radar systems, and home appliances." Three human case studies showed elevated urinary bladder cancer among workers screened at a MOCA factory, but were insufficient for a statistically significant inference of causal association. Animal evidence is strong: MOCA causes liver, kidney, lung, and mammary gland tumors in rats, liver tumors in female mice, and urinary bladder tumors in beagle dogs. On this evidence alone—cancer in multiple sites in two sexes of one mammalian species, in multiple sites in three different species (mice, rats, dogs)—many scientists and several consensus scientific committees judge MOCA a probable human carcinogen.[53]

There is mechanistic evidence. MOCA forms adducts with DNA in rats (and DNA adducts in one human exposed to high doses). Adducts on DNA interfere with DNA functioning in both animals and humans, suggesting a likely toxicity mechanism common to mammals. IARC has found that MOCA causes DNA damage across a wide range of species, thus, making it "comprehensively genotoxic."[53] Animal evidence—combined probable mechanism of action—provided IARC with sufficient evidence to judge that MOCA was a probable human carcinogen.

Consider a second evidentiary pattern, the evidence on 1-(2-chloroethyl)-3-cyclohexyl-1-nitrosourea (CCNU), an anticancer drug. There are no human clinical trials or epidemiological studies showing that it is a carcinogen. In 1981, IARC noted a few human case reports of patients administered CCNU (along with other cytotoxic agents) who developed acute nonlymphocytic leukemia. More important for the IARC inference was that CCNU induced lung tumors in rats, also causing a slight increase in lymphomas in mice. Importantly, CCNU (and several related compounds) act directly as alkylating agents in animals,[54] meaning they attach an alkyl chemical group to the DNA and induce genetic mutations and chromosomal breakage, effects likely to lead to cancer in humans.[4,55]

Because virtually all the evidence that CCNU is a carcinogen comes from animal studies, the evidentiary pattern is noteworthy. A single study showed "increased frequency of sister chromatid exchanges" in humans (a measurement of DNA repair and an indirect measurement of DNA damage).[56] Apart from the few case studies in 1981, no direct evidence from human studies contributes to the scientific committees' consensus that it is a probable human carcinogen.

Were a similar case in court, defendants—manufacturers or those who prescribed the drug—might tell

judges that there are no statistically significant epidemiological studies demonstrating adverse effects from exposure. This is a red herring. None has been done, and they would be difficult to conduct. Populations administered CCNU are likely to be small and a much smaller number yet will live to develop tumors. If such studies could even be done, they would likely be insensitive—unable to separate out the causes of CCNU from other toxic drugs that can disrupt normal biological functioning received by these cancer patients.

Yet the CCNU evidence is persuasive because powerful biological analogies between animals and humans at the genomic, molecular, cellular, and tissue levels are sufficient to persuade scientists that CCNU is a probable human carcinogen, without specific human epidemiologic or clinical trial evidence linking cancer to CCNU.

Scientists have judged commercial benzidine-based dyes, a third example, to be probable or known human carcinogens. Animals and humans metabolize some of these dyes to benzidine, a known human carcinogen; thus, the US National Toxicology Program judged them to be known human carcinogens.[57]

Other benzidine-based dyes (a larger class) do not necessarily metabolize to benzidine, but are so similar in chemical structure that the IARC designates them probable human carcinogens. For three dyes in particular —Direct Black 38, Direct Blue 6, and Direct Brown 95—there were no specific epidemiological studies. A study of "silk dyers and painters who had had multiple exposure to benzidine-based and other dyes" demonstrated bladder cancer.[58] Nonetheless, studies provided sufficient evidence of carcinogenicity for all three in rats or mice or both. Direct Black 38, administered in drinking water, produced liver and mammary tumors in mice as well as hepatocellular carcinomas, urinary bladder tumors, and liver and colon tumors in rats.[58] Direct Blue 6 induced heptocellular carcinomas in rats within only 13 weeks after its oral administration.[58] Direct Brown 95 similarly produced hepatocellular tumors in rats after only 13 weeks, an extremely short period in a 102-week study.[58]

Other "relevant data" contributed to the inferences. Benzidine-based dyes are all structurally related to benzidine, a known human carcinogen; thus, there is a good chemical structure–biological activity link between the benzidine dyes and benzidine. The dyes are also contaminated with small amounts of benzidine. Workers exposed to some benzidine-based dyes have traces of benzidine in their urine. All the dyes show some mutagenicity or effects on mammalian DNA.

The three substances or classes of substances presented here lacked traditional human studies showing a causal association between exposure and cancer. Yet all have been judged by scientists to be probable carcinogens. They are not isolated examples: for about three quarters of 54 substances listed by the IARC as probable human carcinogens, limited or inadequate human evidence exists.[59] IARC inferences are supported instead by controlled studies in animals and by other biological evidence, usually revealing a mode of action that strengthens the inference. IARC scientists ruled out alternative explanations, for example, that the animal studies are statistical accidents; animal mode-of-action studies are not relevant to humans; and

particular chemical structure–activity relationships are not relevant to humans. They accepted animal results as powerful evidence of adverse effects in humans, in that cellular and DNA damage in animals and humans are similar. Scientists conclude that the evidence taken as a whole supports the conclusion through a process of considering plausible alternative explanations, ruling out major competitors as being less likely to be part of good scientific inferences, and concluding that the exposure contributes to the disease.

Causal Inferences from Case Studies

An anesthesiologist was repeatedly exposed to halothane during his early medical career. After his first exposure, he developed hepatitis. On at least five other occasions over a 6-year period, each time he was exposed to halothane he developed hepatitis—except a few times, when he was on prednisone. When no longer exposed to halothane, he recovered. To confirm the causal link, he requested a "rechallenge," a deliberate exposure to halothane under controlled conditions. He suffered an acute attack of hepatitis within 24 hours. The authors of the study concluded, "It is highly probable that halothane was responsible for the recurrent attacks of hepatitis in this case."[60] Moreover, this causal inference was endorsed by the Institute of Medicine and is paradigmatic of Mill's Method of Difference for drawing causal inferences.[30,61] (The Institute also identified other case studies as supporting causal judgments.[30])

In some respects, such an inference is easy because of the temporal proximity between exposure and the adverse reaction as well as the patient's consistent recovery when halothane was removed. Yet halothane has posed difficulties for courts.

A federal district court judge in California refused to admit plaintiff's experts in a case concerning halothane-caused hepatitis. The particular case concerned chronic, not acute, hepatitis, but the judge excluded the experts because "such case reports are not reliable scientific evidence of causation."[29] His ruling might have indicated he was concerned only with the evidence before him, but his use of "such" conveyed wider condemnation, and his view has been widely repeated as a reason for rejecting all case studies as evidence. Some judges continue to reject temporal proximity between exposure and disease as relevant to showing causation, despite cases such as the anesthesiologist's in which temporal proximity assists causal inference.[28] The judge may not have understood the range of case studies—descriptive case studies (which may not be good evidence for toxicity) and more analytic case studies—to rule out other explanations. The latter use all other available evidence to assess whether a case study probably or certainly shows causation.

The identification of vinyl chloride monomer, used to produce polyvinyl chloride, as a potent human carcinogen in a Goodyear polyvinyl chloride plant in Kentucky depended on a more difficult causal inference, also relying largely on case studies. On the basis of 3 cases of a rare form of liver cancer in workers (angiosarcoma), two occupational physicians inferred that exposure to the monomer was the likely cause. A follow-up search revealed 8 other cases of the disease. Animal studies of vinyl chloride, where the monomer had induced the same kind of cancers, and a few human reports from other plants supported the conclusion. The causal inference was made easier, because the disease was very rare (only 25 cases among

250 million US citizens) and had only three other known causes (excessive alcohol consumption, arsenic exposure, and thorium dioxide exposure), which scientists used to rule out other causes for all but one of the workers.[62]

Lack of Information about Potential Toxicants

Substantial lack of information about the universe of chemical substances exacerbates courts' tasks in reviewing scientific evidence. Some evidentiary heuristics adopted by judges may result from an assumption that good scientific evidence about each substance exists (or perhaps that it is easily acquired). Perhaps this accounts for an emphasis on epidemiologic studies. If judges believed, for example, that epidemiologic evidence were readily available, it is not surprising that they would reject plaintiffs' experts for not presenting it. Perhaps judges have been spoiled by robust evidence offered by the defendants in the Bendectin cases, anticipating comparable evidence in every case.[43,44]

The data about any given substance are likely far from ideal and not readily available. Even a cursory review of chemical substances reveals that scientists and the regulatory community appear to know little about the approximately 100000 substances registered for commerce and in common use in the United States. Another 800 to 1000 are added to the list each year with minimal or no testing.[63,64] Not all are problematic, as about 23% are polymers and about 33% present no or only little exposure.[65] In 1984, the National Academy of Sciences found that there were 12860 substances produced in volumes exceeding 1 million pounds per year. For 78% of them, no toxicity information was available. There were 13911 chemicals produced in volumes of less than 1 million pounds (76% with no toxicity data), 8627 food additives (46% with no toxicity data), 1815 drugs (25% with no toxicity data), 3410 cosmetics (56% with no toxicity data), and 3350 pesticides (36% with no toxicity data).[66]

The ingenuity and motivation of free-enterprise institutions produces large numbers of substances without substantial information about their effects on human health and the environment. Unfortunately, these substances and knowledge about them appear out of the control of the institutions that could protect public health and the environment. In the early 1990s there was insufficient change in the data to justify updating the National Academy Report.[67] For the 3000 substances produced in the highest volume (and perhaps representing the most serious problems), substantial toxicity data were lacking in 1998 for about 75% of them, when the EPA entered into an agreement with the producers to close the knowledge gaps.[68] However, in 1995 there were another 1000 to 12000 high-production substances for which extensive toxicological information was judged quite important but was not available.[69]

Most substances enter commerce subject only to postmarket regulatory laws. Little or no premarket testing occurs. If then the chemicals cause harm, the victims may sue, but they face extremely high knowledge and scientific barriers to bring their case to trial about their injuries. Similarly, substances screened under premarket statutes are usually tested in very limited studies that are insufficient to identify risks to the large, heterogeneous populations that will be exposed when the substances enter commerce.

To obtain relief in the tort law, plaintiffs must then find or develop information that institutions failed to require or that the company failed to pursue about its own products. Defendants defend themselves by denying there is evidence about the toxicity of substances in humans and then by disparaging the relevance and reliability of other kinds of evidence. They will get plaintiffs' expert testimony excluded if judges do not appreciate the relevance of different kinds of evidence. Judges who impose scientifically mistaken or overly stringent admissibility barriers create powerful incentives that reinforce the structural problem. There is little reason for firms to investigate their own products as it becomes easier to avoid a tort law trial by challenging a plaintiff's experts at an admissibility hearing.[4,44,46,70]

The Fact of Reasonable Scientific Disagreement

Respectable experts will disagree, for several reasons:

- the nature of scientific inferences
- disagreements about application of agreed-on causal criteria
- lack of scientific knowledge about potentially toxic substances
- complex evidentiary patterns they will be forced to assess

However, it is not simply that there will be disagreements, but these disagreements are perfectly reasonable and legitimate, given the complex inferences that must be made. (I borrow the idea of "reasonable disagreement" from Rawls.[71])

Scientific evidence underdetermines or less than fully determines explanations for a disease or adverse outcome. Not surprisingly, scientists can disagree about the interpretation of identical studies. On scientific advisory panels, members often disagree about the toxicity of a substance or exposure; usually such panels must vote in order to come to conclusions.[72,73] Legitimate disagreements arise more frequently when complex inferences must be drawn based on the toxicological evidence scientists assemble in coming to conclusions. When evidence is conflicting and complex, it is more difficult to assess and evaluate. Even when scientists fully agree about the kinds of studies and the criteria that are appropriate for interpreting them, they may reasonably disagree about the weight to be attributed to each.[51] Relevant concepts are often vague and subject to difficult cases. This results in indeterminacy, leading reasonable persons to disagree. Finally, because of how individual scientists assess different kinds of evidence, the weight they give to it, and the background knowledge they bring to it, an irreducible scientific judgment enters into assessing scientific evidence to infer that a substance does or does not cause a disease. When such conditions obtain for scientific experts, it is not just that experts disagree with one another; the disagreements are quite reasonable and legitimate.

For judges not to intrude on litigants' Seventh Amendment right to a jury trial, which may happen when they mistakenly exclude before trial particular scientific testimony, they should recognize the fact of reasonable scientific disagreement that can easily exist between respectable experts. Courts, in addition to understanding the nature of scientific inferences and a variety of scientific evidentiary patterns, must also

recognize that conscientious, respectable scientists, exercising their scientific judgment, can legitimately come to conclusions that differ with one another without being charlatans, lying, or reasoning egregiously.[4]

If the *Daubert* trilogy of decisions tried to ensure that the law better comports with the relevant science, this will happen only if courts recognize the complexity of scientific evidence, how scientists draw inferences from such evidence and the fact of reasonable disagreements between respectable scientific experts. With the help of the public health community, judges might rectify overreactions to the initial *Daubert* teaching and help ensure that the courts use in the courtroom the kinds of inferences that public health scientists use in their research. This in turn would modestly assist in increasing public health protections from toxicants. Whether such changes will ultimately be sufficient is a further and more difficult issue that must be addressed more fully at another time.[4,70]

Acknowledgments

I received grant support from the National Science Foundation (grant #99-10952), the University of California Toxic Substances Research and Teaching Program, the University of California Academic Senate Intramural funds, and the Project on Scientific Knowledge and Public Policy.

I gratefully acknowledge helpful suggestions from Eula Bingham, PhD, and research assistance from David Strauss on the case studies data.

Footnotes

Peer Reviewed

Accepted for publication September 23, 2004.

References

1. *Daubert v Merrell Dow Pharmaceuticals, Inc*, 509 US 579 (1993).

2. *General Electric Co v Joiner*, 522 US 136 (1997).

3. *Kumho Tire Co v Carmichael*, 526 US 137 (1999).

4. Cranor CF, Eastmond DA. Scientific ignorance and reliable patterns of evidence in toxic tort causation: Is there a need for liability reform? *Law Contemp Prob*. 2001;64:5–48.

5. Cranor CF. Justice, inference to the best explanation and the judicial evaluation of scientific evidence. In:

Campbell JK, O'Rourke M, Shier D, eds. *Topics in Contemporary Philosophy, Vol. III: Law and Social Justice*. Cambridge, Mass: MIT Press; 2005.

6. Thagard P. *How Scientists Explain Disease*. Princeton, NJ: Princeton University Press; 1999.

7. Federal Rules of Evidence, Rule 401. 2001.

8. Wright L. *Critical Thinking: An Introduction to Analytical Reasoning and Reading*. New York, NY: Oxford University Press; 2001.

9. Kramer MS, Lane DA. Causal propositions in clinical research and practice. *J Clin Epidemiol*. 1992;45:639–649.[CrossRef][ISI][Medline]

10. Hutchinson TA, Lane DA. Standardized methods of causality assessment for suspected adverse drug reactions. *J Chronic Dis*. 1989;39:857–890.[CrossRef]

11. Hutchinson TA, Lane DA. Assessing methods for causality assessment of suspected adverse drug reactions. *J Clin Epidemiol*. 1989;42:5–16.[CrossRef][ISI][Medline]

12. Haack S. An epistemologist in the bramblebush: at the Supreme Court with Mr. Joiner. *J Health Polit Policy Law*. 2001;26:217–237.[Free Full Text]

13. In re Agent Orange, 611 F Supp 1223 (ED, New York, 1985).

14. *Viterbo v Dow Chemical Co*, 826 F2d 420, 5th Cir 1987.

15. *Lynch v Merrell Nat. Lab.*, 830 F2d 1190, 1st Cir 1987.

16. *Brock v Merrell Dow Pharmaceuticals, Inc,* 874 F2d 307, 5th Cir 1989.

17. *Turpin v Merrell Dow Pharmaceuticals, Inc,* 959 F2d 1349, 6th Cir 1992.

18. *Allen v Pennsylvania Engineering Corp*, 102 F3d 194, 5th Cir 1996.

19. Rall DP, Hogan MD, Huff JE, Schwetz BA, Tennant RW. Alternatives to using human experience in assessing health risks. *Ann Rev Public Health*. 1987;8:355–385.[CrossRef][ISI][Medline]

20. Huff J, Rall DP. Relevance to humans of carcinogenesis results from laboratory animal toxicology studies. In: Last JM, Wallace RB, eds. *Maxcy–Rosenau–Last Public Health and Preventive Medicine*. 13th ed. Norwalk, Conn: Appleton & Lange; 1992:433–452.

21. Gray GM, Li P, Shlyakhter I, Wilson R. An empirical examination of factors influencing prediction of carcinogenic hazard across species. *Reg Toxicol Pharmacol*. 1995;22:283–291.[CrossRef][ISI][Medline]

22. US Environmental Protection Agency. Proposed guidelines for carcinogen risk assessment. *Fed Regist*. 1996;61(79)17960–18011.

23. Huff J. Chemicals and cancer in humans: First evidence in experimental animals. *Environ Health Persp*. 1993;100:201–210.[ISI][Medline]

24. Pest control: an assessment of present and alternative technologies. In: *Contemporary Pest Control Practices and Prospects: The Report of the Executive Committee*. Washington, DC: National Research Council; 1975.

25. Fung VA, Barrett CJ, Huff J. The carcinogenesis bioassay in perspective: Application in identifying human cancer hazards. *Environ Health Persp*. 1995; 103:680–683.[ISI][Medline]

26. Tomatis L, Huff J, Hertz-Picciotto I, et al. Avoided and avoidable risks of cancer. *Carcinogenesis* 1997;8:97–105.[CrossRef]

27. In re Paoli R.R. Yard PCB Litig., 35 F3d 717–799, 3rd Cir 1994, cert. denied, 513 US 1190; 1994.

28. Cranor CF, Strauss DJ. *Case Studies in Law and Science*. Forthcoming.

29. Casey v Ohio Medical Products, 877 F Supp 1380 At 1385 (ND, California 1995.

30. Stratton ER, Howe CJ, Johnston RB, eds. *Adverse Events Associated with Childhood Vaccines: Evidence Bearing on Causality*. Washington, DC: National Academy Press; 1994.

31. Collet J-P, MacDonald N, Cashman N, Pless R, and Advisory Committee on Causality Assessment. Monitoring signals for vaccine safety: the assessment of individual adverse event reports by an expert advisory committee. *Bull World Health Organ*. 2000; 78:178–185.[ISI][Medline]

32. International Agency for Research on Cancer. Preamble to the IARC Monograph Series (8. Studies of Cancer in Humans). Available at: http://www-cie.iarc.fr/monoeval/studieshumans.html. Accessed March 15, 2005.

33. Harman G. The inference to the best explanation. *Philos Rev*. 1964;LXXIV:89–90.

34. Hill AB. The environment and disease: Association or causation? *Proc R Soc Med*. 1965;58:295–300. Reprinted in: Greenland S, ed. *Evolution of Epidemiologic Ideas: Annotated Readings on Concepts and Methods*. Chestnut Hill, MA; 1987:15–19.[ISI][Medline]

35. Cullen MR, Rosenstock L, Brooks SM. Clinical approach and establishing a diagnosis of an environmental medical disorder. In: Brooks SM, Gochfeld M, Herzstein J, Jackson RJ, Schenker MB, eds. *Environmental Medicine*. St Louis, Mo: Mosby; 1995:217–231.

36. Rothman KJ, Greenland S. *Modern Epidemiology*. 2nd ed. Philadelphia, Pa: Lippincott-Raven; 1996.

37. Wright L. *Practical Reasoning*. New York, NY: Harcourt Brace Jovanovich; 1989.

38. Electric and magnetic fields risk evaluation guidelines. Sacramento: California Department of Health Services; 1999.

39. *Black v Food Lion, Inc,* 171 F3d 308, 5th Cir 1999.

40. *Westberry v Gislaved Gummi*, 178 F3d 257, 4th Cir 1999.

41. *Hollander v Sandoz Pharmaceuticals Corp,* 95 F Supp 2d 1230–1239 (WD, Oklahoma, 2000).

42. Santone KS, Powis G. Mechanism of and tests for injuries. In: Hayes WJ Jr, Laws ER Jr, eds. *Handbook of Pesticide Toxicology*. New York, NY: Academic Press; 1991:169–214.

43. Green MD. Legal Theory: Expert witnesses and sufficiency of evidence in toxic substances litigation. *NWU Law Rev*. 1992;86:643–699.

44. Green MD. Bendectin and birth defects. Philadelphia: University of Pennsylvania Press; 1996.

45. Glastetter v Novartis Pharmaceuticals Corp, 107 F Supp 2d 1015–1046 (ED, Missouri, 2000).

46. Cranor CF. *Regulating Toxic Substances*. New York, NY: Oxford University Press; 1993.

47. Cranor CF, Fischer JG, Eastmond DA. Judicial boundary-drawing and the need for context-sensitive science in toxic torts after *Daubert v Merrell-Dow Pharmaceuticals*. *Va Environ Law J*. 1996;16:1–77.

48. Rinsky RA, Hornung RW, Silver SR, Tseng CY. Benzene exposure and hematopoietic mortality: A long-term epidemiological risk assessment. *Am J Indus Med*. 2002;42:474–480.[CrossRef][ISI]

49. Silver SR, Rinsky RA, Cooper SP, Hornung RW, Lai D. Effect of follow-up time on risk estimates: Longitudinal examination of the relative risks of leukemia and multiple myeloma in a rubber hydrochloride cohort. *Am J Indus Med*. 2002;42:481–489.[CrossRef][ISI]

50. Weed DL. [National Cancer Institute.] Personal Communication; January 2003.

51. Weed DL. Underdetermination and incommensurability in contemporary epidemiology. *Kennedy Inst Ethics J*. 1997;7:107–114.[ISI][Medline]

52. Correspondence from Jerry M. Rice to Vicent DelPizzo of the California Electric and Magnetic Fields Program. Sent to the author February 1, 2003.

53. International Agency for Research on Cancer. IARC Monograph Series, Vol. 57, 4,4'-METHYL-ENEBIS(2-CHLOROANILINE) (MOCA). Available at: http://www-cie.iarc.fr/htdocs/monographs/vol57/15-moca.html. Accessed March 15, 2005.

54. International Agency for Research on Cancer, IARC Monograph Series, Supplement 7, Chloroethylnitrosoureas. Available at: http://www-cie.iarc.fr/htdocs/monographs/suppl7/chloroethylnitrosoureas.html. Accessed March 15, 2005.

55. Gregus Z, Klaassen CD. Mechanisms of toxicity. In: Klaassen CD, ed. *Casarett and Doull's Toxicology*. 6th ed. New York, NY: McGraw-Hill; 2001:35–82.

56. Preston RJ, Hoffman GR. Genetic toxicology. In: Klaassen CD, ed. *Casarett and Doull's Toxicology*. 6th ed. New York, NY: McGraw-Hill; 2001:321–350.

57. Public Health Service, National Toxicology Program. *10th Annual Report on Carcinogens*. Washington, DC: US Department of Health and Human Services; December 2002.

58. International Agency for Research on Cancer. Probably carcinogenic to humans, Benzidine-based dyes.

Available at: http://www-cie.iarc.fr/htdocs/monographs/suppl7/benzidinedyes.html. Accessed March 15, 2005

59. International Agency for Research on Cancer. Overall evaluations of carcinogenicity to humans: Group 2A: probably carcinogenic to humans. Available at: http://www-cie.iarc.fr/monoeval/crthgr02a.html. Accessed March 15, 2005.

60. Klatskin G, Kimberg DV. Recurrent hepatitis attributable to halothane sensitization in an anesthetist. *N Engl J Med*. 1969;280:515–522.[ISI][Medline]

61. Mill JS. *A System of Logic: Ratiocinative and Inductive*. London, England: Longmany, Green & Co; 1941.

62. Falk H, Creech JL Jr, Heath CW Jr, Johnson MN, Key MM. Hepatic disease among workers at a vinyl chloride polymerization plant. *JAMA*. 1974;230:59–68.[CrossRef][ISI][Medline]

63. Huff J, Melnick R. Chemical carcinogenesis bioassays: Critical for the sensible application of the precautionary principle for protecting public health. Presentation to Collegium Ramazzini, October 23–24, 2002.

64. *Identifying and Regulating Carcinogens*. Washington, DC: Office of Technology Assessment, US Congress; 1987.

65. Huff J, Hoel D. Perspective and overview of the concepts and value of hazard identification as the initial phase of risk assessment for cancer and human health. *Scand J Work Environ Health*. 1992;18:83–89.[ISI][Medline]

66. National Academy of Sciences. *Toxicity Testing: Strategies to Determine Needs and Priorities*. Washington, DC: National Academy Press; 1984.

67. Bailor JC, Bingham E. Personal communications at Collegium Ramazzini, 2002.

68. EPA, EDF, CMA agree on testing program targeting 2,800 chemicals. *Environ Health Letter*. October 1998;37:193.

69. *Screening and Testing Chemicals in Commerce*. Washington, DC: Office of Technology Assessment, US Congress; 1995.

70. Berger MA. Eliminating general causation: Notes toward a new theory of justice and toxic torts. *Colum Law Rev*. 1997;97:2117–2151.[ISI]

71. Rawls J. *Political Liberalism*. Cambridge, Mass: Harvard University Press; 1993.

72. *Assessment of Health Effect from Exposure to Power-Line Frequency Electric and Magnetic Fields*. Research Triangle Park, NC: National Institute of Environmental Health Sciences; 1998.

73. Neutra RR, DelPizzo V, Lee G. An evaluation of the possible risks from electric and magnetic fields (EMFs) from power lines, internal wiring, electrical occupations, and appliances. Final Report, Sacramento: California Department of Health Services; 2002.

This article has been cited by other articles: (Search Google Scholar for Other Citing Articles)

American Journal of PUBLIC HEALTH HOME

E. Bingham, L. Boden, R. Clapp, P. Hoppin, S. Krimsky, D. Michaels, D. Ozonoff, and A. Robbins
BINGHAM ET AL. RESPOND
Am J Public Health, February 1, 2006; 96(2): 206 - 207.
[Full Text] [PDF]

American Journal of PUBLIC HEALTH HOME

R. L. Melnick
A Daubert Motion: A Legal Strategy to Exclude Essential Scientific Evidence in Toxic Tort Litigation
Am J Public Health, July 1, 2005; 95(S1): S30 - S34.
[Abstract] [Full Text] [PDF]

American Journal of PUBLIC HEALTH HOME

D. Michaels
Scientific Evidence and Public Policy
Am J Public Health, July 1, 2005; 95(S1): S5 - S7.
[Full Text] [PDF]

This Article

- **Abstract** FREE
- **Full Text (PDF)**
- **Submit a response**
- **Alert me when this article is cited**
- **Alert me when eLetters are posted**
- **Alert me if a correction is posted**

Services

- **Similar articles in this journal**
- **Similar articles in PubMed**
- **Alert me to new issues of the journal**
- **Download to citation manager**

Google Scholar

- **Articles by Cranor, C.**
- **Articles citing this Article**

PubMed

- **PubMed Citation**

Articles by Cranor, C.

HOME HELP FEEDBACK SUBSCRIPTIONS ARCHIVE SEARCH TABLE OF CONTENTS
Copyright © 2005 by the American Public Health Association