Journal of Law and Policy

Volume 15 | Issue 1 Article 2

2007

A Framework for Assessing Scientific Arguments: Gaps, Relevance and Integrated Evidence

Carl F. Cranor

A FRAMEWORK FOR ASSESSING SCIENTIFIC ARGUMENTS: GAPS, RELEVANCE AND INTEGRATED EVIDENCE

Carl F. Cranor, Ph.D., M.S.L.[*]

INTRODUCTION

The United States Supreme Court, in deciding *Daubert v. Merrell-Dow Pharmaceuticals* and related cases,[1] overturned the *Frye* rule for assessing expert testimony[2] and heightened judges' duties to review expert testimony, including scientific testimony in toxic tort cases. This job may have turned out to be more

---

[*] Professor of Philosophy, University of California, Riverside. M.S.L., 1981 Yale Law School; Ph.D. 1971, University of California, Los Angeles; B.A. 1966, Mathematics, University of Colorado. I want to thank Margaret Berger for the opportunity to discuss these issues at the Science for Judges Program at Brooklyn Law School March 3-4, 2006.

[1] Daubert v. Merrell Dow Pharmaceuticals, 509 U.S. 579 (1993); General Elec. Co. v. Joiner, 522 U.S. 136 (1997); Kumho Tire Co. v. Carmichael, 526 U.S. 137 (1999).

[2] Frye v. U.S., 293 F. 1013 (D.C. Cir. 1923). The *Frye* test required that the generic kinds of studies, tests or techniques on which an expert might rest testimony must be "generally accepted in the pertinent field." *Id.* at 104. Specifically it noted that, "Just when a scientific principle or discovery crosses the line between the experimental and demonstrable stages is difficult to define. *Id.* 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." *Id.*

difficult than the justices envisioned because one hurdle facing experts is the nature of scientific (or nondeductive) arguments themselves. The structure of nondeductive inferences, despite common use of them in our daily lives and their pervasive use in science and other technical areas, may not be fully appreciated because of the complex relationship between premises and conclusion and the substantive expertise needed to assess the arguments.

Section I of this article discusses the amount of data known about some chemicals currently in commerce and how this profoundly affects information that experts have at their disposal, likely increasing the complexity of the arguments they can offer in support of legal claims. Section II of this article describes nondeductive (and scientific) inferences. Section III then focuses on three specific problems that pose particular challenges for courts: gaps in the arguments, scientific relevance of information, and the need to assess arguments based on all the integrated relevant evidence.[3] Scientific testimony must utilize scientific arguments. These are nondeductive inferences, widely utilized, but perhaps under-appreciated. Nondeductive arguments, even the best of them, will have gaps that courts will need to recognize but not overreact to. A reviewing body needs to permit an expert's testimony to rely on all the scientifically relevant information he or she utilizes in his or her arguments. The arguments must then be assessed by considering *all the scientifically relevant evidence taken together on which the argument is based* (not rejecting each piece of evidence); otherwise the review would not be a scientifically fair evaluation. Proper understanding of scientific arguments suggests that some courts have not fairly reviewed them as *Daubert* mandates. Lastly, Section IV will sketch how scientists would assess such arguments and then suggest a procedure courts might follow to mimic this for legal admissibility.

[3] This Article does not seek to undertake a comprehensive discussion of the numerous issues concerning such arguments.

I. IGNORANCE ABOUT THE CHEMICAL UNIVERSE

Chemical substances are quite often the objects of toxic tort suits, but regrettably little is known about them. This is a substantial barrier for those who might have been harmed by exposure to a toxic chemical. Moreover, it also greatly burdens experts who testify about such exposures because they are often forced to assemble disparate kinds of evidence that is both scientifically relevant and sufficiently good to come to conclusions about the toxicity of products. This may pose reviewing difficulties for judges and others who are not intellectually close to scientific arguments because such arguments are complex.

In 1984 there were between 65,000 and 100,000 chemicals registered for use in commerce. At that time little was understood about their toxicity properties. The National Academy of Sciences found that for 78 percent of the 3,000 top-volume chemicals in commerce, the most basic toxicity results could not be found in the public record.[4] The Academy's findings with regard to particular groups of registered chemicals were equally disturbing. The Academy found:

- 12,860 substances produced in excess of one million lbs/year, but 78 percent had no toxicity data at all;
- 13,911 chemicals produced in less than one million lbs/year (76 percent had no toxicity data);
- 21,752 substances whose production volume was unknown (82 percent had no toxicity data);
- 8,627 food additives (46 percent had no toxicity data);*
- 1,815 drugs (25 percent no toxicity data);*
- 3,410 cosmetic (56 percent had no toxicity data); and
- 3,350 pesticides (36 percent had no toxicity data).[5]

---

[4] NATIONAL RESEARCH COUNCIL, TOXICITY TESTING 84 (1984).

[5] *See* U.S. CONGRESS, OFFICE OF TECHNOLOGY ASSESSMENT, SCREENING AND TESTING CHEMICALS IN COMMERCE 1 (1995). If derivatives and metabolites are included, some experts suggest that the more appropriate number is 100,000. NATIONAL RESEARCH COUNCIL, TOXICITY TESTING 84 (1984). Categories of substances marked with an asterisk (*) are currently

Although this data is 20 years old, there is additional current evidence. In the early 1990s NRC Committee members, in response to being asked whether the report should be updated, found that there was insufficient change in the data to justify revisiting the 1984 findings.[6] As recently as 1997, 75 percent of the 3,000 chemicals produced in the highest volumes still lacked basic toxicity data in the public record.[7] This point about the scientific ignorance of substances should neither be overemphasized nor underemphasized. Although many substances have been created, some of them likely have not been pursued.[8] Others will be industrial intermediates without significant human exposures and some may be sufficiently large molecules that they are not likely to pose toxicity problems (although they may not be wholly risk free).[9] Nonetheless, there are a sufficiently large number of products in the market where there is likely to be public and workforce exposure that citizens should not rest easy. Wide anecdotal evidence suggests that the public believes that products to which they are exposed *are* legally required to be tested before they enter the market and perhaps that scientists know as much about most substances as they do about pharmaceutical products.[10] However, as the data shows, this suggestion is clearly mistaken.

Lack of scientific knowledge about the chemical universe will complicate judicial review of testimony in toxic tort cases.

---

subject to premarket testing and approval. This probably accounts for somewhat greater knowledge about them. James Huff & David Hoel, *Perspective and Overview of the Concepts and Value of Hazard Identification as the Initial Phase of Risk Assessment for Cancer and Human Health*, 18 SCAN. J. WORK ENVIRON HEALTH 83, 85 (1992) (estimating that 50,000-100,000 exist chemicals in the marketplace).

[6] John C. Bailor & Eula Bingham, personal communication, Annual Meetings of the Collegium Ramazzini, (2002).

[7] *See* DAVID ROE ET AL., ENVIRONMENTAL DEFENSE FUND, TOXIC IGNORANCE 7 (1997).

[8] *See* Huff & Hoel, *supra* note 5.

[9] *Id.*

[10] This is a common view expressed by students in my classes and by audiences to whom I give presentations.

Experts likely will be forced to rely upon less than ideal information to make inferences about whether exposure to a substance can and did cause harm, because for any randomly chosen substance there is probably poor toxicity data. In particular, there will most likely be a shortage of good human data, because, for a variety of reasons, human data will be unavailable.[11] This predicament forces scientists to infer any toxicity properties of products based on more complex evidentiary patterns such as animal evidence, chemical structure-biological activity evidence, mutagenicity studies,[12] molecular evidence and so on.[13]

II. THE STRUCTURE OF SCIENTIFIC ARGUMENTS

More complex patterns of evidence present increasing challenges to reviewing scientific testimony because the kinds of evidence and how they are combined in scientific arguments are less familiar to those not steeped in the science. Nonetheless, courts need to be prepared for the task because even complex arguments can be quite good. In order to carry out their gatekeeping duties and be fair to both sides of the bar, judges will need to be able to understand the major outlines of such arguments. In addition, scientific arguments pose their own

---

[11] In particular, human evidence is likely to be missing for many substances because often they have not been studied and human epidemiological studies tend to be insensitive, *See* Huff & Hoel, *supra* note 5 (Table 2) and *supra* Section I.

[12] Mutagency has been described as:

> [t]he induction of mutations is due primarily to chemical or physical aterations in the structure of DNA that result in inaccurate replication of that region of the genome. The process of mutagenesis consists of structural DNA alternation, cell proliferation that fixes the DNA damage, and DNA repair that either directly repairs the alkylated base or bases or results in removal of larger segments of the DNA.

Henry C. Pitot III & Yvonne P. Dragan, *Chemical Carcinogenesis, in* CASARETT AND DOULL'S TOXICOLOGY 6th ed., 86 (Curtis D. Klaassen ed., McGraw-Hill 2001) at 241-319, esp. at 256.

[13] *See* discussion *infra* at notes 148-61.

difficulties.

Providing a causal explanation of an event is a matter of inference and of argument. How does one *infer* the best explanation of an event? Providing a philosophic account of causal explanation is difficult enough on its own terms; some philosophers emphasize the idea of making inferences to the best explanation,[14] some are Bayesians,[15] and so on.[16] This Section will identify some of the major steps in inferences that are common to the different approaches by which scientists come to conclusions about the causal effects of exposures to substances. The aim is to provide a sufficiently accurate overview of nondeductive reasoning and to provide a characterization of such inferences. This is then used to highlight some difficulties to which courts must be alert in reviewing experts' reasoning for causal conclusions about toxicity. This Section focuses on causal explanations of disease because this is especially germane to courts' tasks in toxic tort cases.

A. Deductive and Nondeductive Arguments

Arguments in support of conclusions are of two kinds: deductive and nondeductive.[17] Deductive arguments are typical of mathematics and formal logic. The defining property of such

---

[14] GILBERT HARMAN, *The Inference to the Best Explanation*, PHIL. REV., 74, 89-90 (1964); LARRY WRIGHT, CRITICAL THINKING: AN INTRODUCTION TO ANALYTICAL READING AND REASONING 206, 206-17 (2001). Thagard adopts much of this view, indicating that scientists:

> can infer that the factor causes the disease if this hypothesis is part of the best explanation of the full range of evidence . . . . [and that the factor that is identified as causing] the disease must be a better explanation of the correlation between the factor and the disease than the assertion that some other cause is responsible for both the factor and the disease.

PAUL THAGARD, HOW SCIENTISTS EXPLAIN DISEASE 129 (1999).

[15] *See generally* BRIAN SKYRMS, CHOICE AND CHANCE: AN INTRODUCTION TO INDUCTIVE LOGIC (1966).

[16] This Article does not advocate on either side of the debate, nor does it choose between different fundamental accounts.

[17] LARRY WRIGHT, PRACTICAL REASONING 38-46 (1989).

arguments is that the conclusion is "guaranteed" logically or semantically by the premises: if the premises are true, the conclusion must be true.[18] In a deductively *valid* argument, if one finds the conclusion to be false, at least one of the premises must be false as well. Or, if one accepts the truth of the premises, but rejects the truth of the conclusion in a valid argument, one contradicts oneself.[19] For example, if one accepts that A = B and B = C, but denies that A = C, on the face of the argument one contradicts oneself. Logical tightness gives deductive arguments great inferential power, as the success of mathematics and formal logic shows.[20] Deductive arguments are of limited use in making empirical inferences about the world.[21]

By contrast, nondeductive arguments are simply those whose conclusions are *not guaranteed* by their premises.[22] These are typically utilized in support of empirical claims,[23] e.g., an explanation of the ocean tides or of benzene causing leukemia. Theorists have called such arguments "inferences to the best explanation," "diagnostic arguments," "diagnostic induction," "inductive arguments," and "differential diagnosis."[24] In nondeductive arguments even if the premises are true, the inferential link between premises and conclusions will have varying degrees of strength, unlike a deductive argument.[25] Thus, in a nondeductive argument the premises will provide strong, weak, or moderate support for the conclusion; one might

---

[18] *Id*.

[19] *Id*. at 40.

[20] Theorems of mathematics and formal logic are derived by deductive reasoning.

[21] WRIGHT, PRACTICAL REASONING, *supra* note 17, at 38-46.

[22] *Id*.

[23] *Id*.

[24] *Id*.; *See also* HARMAN, *supra* note 14, at 89 (noting that the term "corresponds approximately to what others have called 'abduction,' 'the method of hypothesis,' 'hypothetic inference,' 'the method of elimination,' 'eliminative induction,' and 'theoretical inference.'"), and SKYRMS, *supra* note 15. This Article largely uses the phrase "inference to the best explanation" because of the reasoning process it suggests.

[25] WRIGHT, PRACTICAL REASONING, *supra* note 17, at 38-46.

say that the argument will be strong or weak or in between, but not valid or invalid.[26]

Moreover, the given premises will provide support for different possible conclusions (or as the literature puts it, the evidence provided in the premises may be consistent with different *explanations*).[27] The task, then, in evaluating such inferences is to determine which conclusion is the most plausible (or best supported) given the premises (or which explanation best accounts for the evidence in the premises).

B. Major Steps in Using Nondeductive Reasoning to Make Inferences to the Best Explanation for Causes of Disease

This Section highlights the major steps in using nondeductive reasoning to make inferences concerning what may have caused a disease. A typical first step that would lead to an interest in disease causation is that a researcher may notice something that needs understanding or calls for an explanation.[28] This might be a correlation or association between some exposure or condition and a disease.[29] Once such a correlation has been observed, it invites an explanation, if it is sufficiently interesting.[30] For example, in a polyvinyl chloride plant in Kentucky in the mid-1970s, occupational physicians noticed three individuals with angiosarcoma of the liver (an extremely rare liver cancer).[31]

---

[26] *Id.* at 48-50.

[27] *Id.* at 107-11.

[28] The reconstruction of nondeductive arguments is largely mine, but I have learned a great deal from conversations with my colleague Larry Wright and his books, especially WRIGHT, PRACTICAL REASONING, *supra* note 17, at 99-121, WRIGHT, CRITICAL THINKING, *supra* note 14, and THAGARD, HOW SCIENTISTS EXPLAIN DISEASE, *supra* note 14.

[29] *Id.*

[30] The mere fact that there are correlations between two things often provides something to be explained**.** *See* WRIGHT, PRACTICAL REASONING, *supra* note 17, at 154.

[31] *See generally* J.L. Creech, Jr. & M.N. Johnson, *Angiosarcoma of Liver in the Manufacture of Polyvinyl Chloride,* 16 JOURNAL OF

Subsequently, they identified ten other cases at vinyl chloride plants in the U.S.[32] Questions posed about this phenomena included: Does this observed correlation have a causal explanation or not? If it does, what is it? In the law one would ask, "If there is a causal explanation for an observed association, is it one for which a responsible party should be held accountable or not?"

Second, in trying to understand casual relationships a researcher would consider a sufficiently complete list of plausible explanations to account for the evidence.[33] This important step is one of the "most basic [yet] least understood" and difficult steps in nondeductive inferences.[34] For this step, some philosophers would emphasize finding a list of plausible explanations to try to account for the phenomena, and would argue that this is based on scientists' experience, expertise, background knowledge, and other evidence of the effects.[35] More Bayesian-oriented theorists would talk of conditioned properties.[36] When scientists sought to identify the cause of

---

OCCUPATIONAL MEDICINE 150 (1974).

[32] Clark W. Heath, Jr., Henry Falk & John L. Creech, Jr., *Characteristics of Cases of Angiosarcoma of the Liver among Vinyl Chloride Workers in the United States,* 246 ANNALS NEW YORK ACADEMY OF SCIENCES 231 (1975).

[33] WRIGHT, PRACTICAL REASONING, *supra* note 17, at 99-102.

[34] SKYRMS, *supra* note 15, at 107.

[35] WRIGHT, PRACTICAL REASONING, *supra* note 17, at 99-104, esp. 103.

[36] Bayesian-oriented philosophers would, as Skyrms puts it, try to ascertain:

> [W]hat factors are likely to be relevant to the conditioned property in which we are interested [the thing to be explained]; there must be some way of setting up a list of *reasonable length* of possible conditioning properties, which probably contains the necessary or sufficient conditions being sought. The only way to do this is to apply inductive logic to previously acquired body of evidence.

SKYRMS**,** *supra* note 15, at 107. Conditioning properties on Skyrms' view are those that produce a causal effect. His account of "conditioning properties" may in fact be somewhat wider than "possible explanations" endorsed by the other view, but this is not germane to our discussion (I owe this point to

angiosarcoma at the polyvinyl chloride (PVC) plants, they had a fortuitously short list of possible causes from which to seek an explanation—excessive alcohol usage, exposure to arsenic compounds, exposure to thorium dioxide, bad luck, and a possible new cause, exposure to vinyl chloride monomers (VCM).[37] Frequently, the list of possible causes of an adverse condition could be much longer or less well-defined.

Third, scientists would then rank the list of rival explanations according to their *plausibility* or initial degree of support based on the evidence available at the time.[38] Such evidence would include both evidence collected at the time of the investigation and background knowledge about the subject being studied.[39] Plausibility rankings would refer to "the list of rival explanations [to explain what is going on] in the order of their plausibility."[40] Thus, "[w]hen we judge the [explanatory] rivals [of nondeductive arguments] to be more or less plausible, we are estimating how well or badly they explain what happened, or what is going on, given what we know about it."[41] Such plausibility judgments have degrees of gradation or degrees of strength.[42] What was the most plausible initial explanation of the liver cancer among polyvinyl chloride workers?

Fourth, the initial plausibility rankings would in turn provide clues about what *other evidence* might be available that would distinguish between the explanations.[43] Scientists would seek additional evidence to separate the more plausible from the less

---

Larry Wright).

[37] Heath, et al., *supra* note 32, at 234; D. B. CLAYSON, TOXICOLOGICAL CARCINOGENESIS 11–12 (2001).

[38] WRIGHT, PRACTICAL REASONING, *supra* note 17, at 93-96.

[39] *Id*.

[40] *Id*. at 101.

[41] *Id*. at 107.

[42] *Id*. at 47. Individuals can develop skills in ranking the different conclusions from the premises based on their *plausibility*. Such skills are quite important for scientists and the explanations they consider within their fields. Courts need to recognize the importance of the implicit scientific skill in recognizing and utilizing scientific inferences to the best explanation.

[43] WRIGHT, PRACTICAL REASONING, *supra* note 17, at 103-04.

plausible explanations.[44] What research in the form of tests, studies, or background information could help discriminate between different explanations and assist the search for a best explanation?

Often, finding such information is much easier said than done. Sometimes experiments cannot be conducted; studies available may not directly address needed issues; diseases may be sufficiently rare, or, conversely, sufficiently common that they are difficult to detect; disease processes can be too complex at present to discern causes and so on.[45] This could be especially problematic for the law where studies have not been explicitly designed to address the legal questions. In the PVC plant case, once researchers suspected that an occupational exposure might have caused the cancers, they sought evidence from other PVC plants as well as animal studies.[46] They found some earlier evidence that animals exposed to vinyl chloride contracted similar diseases as well as 1972 preliminary results of a study by Cesar Maltoni in Italy,[47] which reinforced their considerable initial concerns that had been based on case reports.

Fifth, there is widespread agreement that *all relevant* information bearing on possible explanations must be considered in drawing a conclusion about which explanation or conditioned property is most likely.[48] Scientifically "relevant information" is

---

[44] *Id*.

[45] THAGARD, *supra* note 14, at 131; CARL F. CRANOR, TOXICS TORTS; SCIENCE, LAW AND THE POSSIBILITY OF JUSTICE (2006) 170-82.

[46] GERALD MARKOWITZ & DAVID ROSNER, DECEIT AND DENIAL: THE DEADLY POLITICS OF INDUSTRIAL POLLUTION 182-83 (2002).

[47] *Id*. *See also* Cesare Maltoni, *Occupational Carcinogenesis Predicitive Value of Carcinogenesis Bioassays,* 271 ANNALS N.Y. ACAD. SCI. 431 (1976).

[48] *See*, *e.g.*, THAGARD, *supra* note 14, at 129; SKYRMS, *supra* note 15, at 107; Tom A. Hutchinson & David A. Lane, *Assessing Methods for Causality Assessment of Suspected Adverse Drug Reactions*, 42 J.CLINICAL EPIDEMIOL. 5 (1989). *See also* Jerome P. Kassirer & Joe S. Cecil, *Inconsistency in Evidentiary Standards for Medical Testimony: Disorder in the Courts*, J. AM. MED. ASS'N 1382, 1386 (Sept. 18, 2002) (noting writers from different methodological perspectives who agree on this point). Hutchinson and Lane put this point especially strongly:

information that has *any impact* on the probability of a scientist's conclusions and the plausibility of explanations (or conditioned properties).[49] Relevance judgments may not always be without controversy, but the standard for what constitutes "relevant" information is quite minimal, since typically *any* information that can affect a scientist's belief, ranking of possible explanations, probability of conditioned properties or conclusions should be included.[50] When trying to identify the cause of angiosarcoma in PVC plants no epidemiological studies were then available. However, there were good case reports from the Kentucky plant, as well as a few from some other PVC plants, revealing surprising clusters of an unusually rare disease.[51] Also, good animal studies from Maltoni's laboratory greatly assisted the inferences.[52] All of this was relevant evidence.[53]

What constitutes scientifically relevant information or data for drawing a scientific conclusion is a matter of scientific

---

> 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 incorporable by the method into the 'state of information' on which the assessment is based.

Hutchinson & Lane, *supra* at 10 (emphasis added).

[49] *See, e.g.*, WRIGHT, PRACTICAL REASONING, *supra* note 17, at 104; WRIGHT, CRITICAL THINKING, *supra* note 13, at 206-17; SKYRMS, *supra* note 14.

[50] Legal conceptions of "relevant" evidence are remarkably similar to scientific definitions. KENNETH S. BROUN ET AL., MCCORMICK ON EVIDENCE 541-542 (Edward W. Cleary ed., 3d ed. 1984); FED. R. EVID. 401 ("Evidence having any tendency to make the existence of any fact that is of consequence to the determination of the action more probable or less probable than it would be without the evidence.").

[51] Heath, et. al., *supra* note 32, at 232.

[52] MARKOWITZ & ROSNER, *supra* note 46, at 182-83 (Maltoni found that cancers were produced in rats after vinyl chloride exposure at levels lower than the existing occupational standard in the U.S.).

[53] Maltoni, *supra* note 47.

judgment.[54] This is not intended to be a tautology, but to make the point that scientists have the *substantive expertise* to assess what is relevant for making scientific inferences.[55] This is not to say that a relevance judgment is totally subjective or that it can be idiosyncratic; peers can correct other experts. Scientists *may* differ about relevance judgments, but this is likely to be unusual, simply because satisfying relevance considerations tends to be quite easy.[56] Finally, even in cases in which scientific conclusions might be controversial, what constitutes *relevant data* may be less controversial than the conclusions drawn from the data.[57]

Sixth, central to drawing scientific conclusions is a need to *integrate* all the available relevant evidence to come to a conclusion.[58] In assessing a substance's toxicity, good scientific practice dictates that scientists consider available human evidence, evidence from experimental animals, chemical structure-biological activity evidence, mechanistic evidence (which is rarely available), and so on, in order to evaluate the conclusion that follows.[59] The metaphor of fitting the pieces of a puzzle together to see what "picture" the totality of evidence provides is often used to describe this process.[60]

---

[54] WRIGHT, PRACTICAL REASONING, *supra* note 17, at 104.

[55] *Id*.

[56] WRIGHT, CRITICAL THINKING, *supra* note 14, at 206-17.

[57] From the nature of relevance judgments one can see that they tend to be much easier in science and the law than conclusions that might be drawn from them. *See* WRIGHT, PRACTICAL REASONING, *supra* note 17, at 104; WRIGHT, CRITICAL THINKING, *supra* note 14, at 206–17; MCCORMICK ON EVIDENCE, *supra,* note 50, at 541-42.

[58] INSTITUTE OF MEDICINE AND NATIONAL RESEARCH COUNCIL, COMMITTEE ON THE FRAMEWORK FOR EVALUATING THE SAFETY OF DIETARY SUPPLEMENTS, DIETARY SUPPLEMENTS: A FRAMEWORK FOR EVALUATING SAFETY (2005) at 255-56, 262.

[59] V.J. Cogliano, R.A. Baan, K. Straif, Y. Grosse, M. Secretan, F. El Ghissassi & P. Kleihues, *The Science and Practice of Carcinogen Identification and Evaluation*, 112 ENVIRONMENTAL HEALTH PERSPECTIVES 1269, 1272 (2004).

[60] Susan Haack, *Trial and Error: The Supreme Court's Philosophy of Science*, 95 AMERICAN JOURNAL OF PUBLIC HEALTH, (Suppl. 1) S66, S70

At the International Agency for Research on Cancer (IARC), for example, the scientific committees explicitly go through a stepwise process. The committee considers any evidence that a substance causes cancer in humans, and evidence that it causes cancer in animal studies.[61] These *lines of evidence* are then combined to provide a default evaluation of the substance's likelihood of causing cancer in humans.[62] The committee then considers mechanistic and other kinds of evidence to "determine whether the default evaluation should be modified."[63] The current director of that program emphasizes the role of scientific judgment in integrating evidence and coming to conclusions in scientific argument. "The final overall evaluation is a matter of *scientific judgment*, reflecting the weight of the evidence derived from studies in humans, studies in experimental animals, and mechanistic and other relevant data."[64]

This view of the IARC process makes explicit that professional judgment is central to drawing scientific inferences. Moreover, scientific judgment has a crucial role at several points, not just in drawing a final conclusion. An expert reviews data that appear to bear on causal judgments, selects the scientifically relevant data, assesses and weighs studies for their quality, weighs the importance of different kinds of data vis-à-vis one another (e.g., animal studies versus human studies versus short-term studies versus structure-activity relationships versus any case studies), and brings her background understanding of biology and toxicology, as well as her understanding of the phenomena, to the causal issues. She then evaluates different possible explanations in light of all the evidence and the particular phenomena (i.e., a disease) that she wants to explain.

An expert considers and integrates *all* scientifically relevant

---

(Suppl. 1 2005); Margaret A. Berger, "*What Has a Decade of Daubert Wrought?*" 95 AMERICAN JOURNAL OF PUBLIC HEALTH S59, S60 (Suppl. 1 2005).

[61] Cogliano et. al., *supra* note 59.

[62] *Id*.

[63] *Id*.

[64] *Id*. (emphasis added).

evidence in order to assess what it shows. Then the expert enters into an assessment of the strength of the best explanation vis-à-vis alternative explanations. As both medical and legal commentators put the point: "In the final analysis, assessment of evidence and causal inferences depend on accumulating *all* potentially relevant evidence and making a subjective judgment about the strength of the evidence."[65] When professional judgment is so central to the drawing of inferences, professionals may disagree.

Finally, the search for causal understanding takes into account all relevant information, and focuses on how much more probable or plausible an effect is with a cause than without that cause.[66] For example, was it more plausible that employees in the polyvinyl chloride plants contracted their liver cancer as a matter of coincidence or from exposure to thorium oxide, alcohol consumption, arsenic compounds, or the vinyl chloride monomer? In pursuit of the best explanation, a scientist would seek evidence that would *increase the plausibility "gap"* between the highest ranked explanation and the next highest ranked one. That is, during an investigation, the initially top-ranked explanation may gain in strength and plausibility. Alternatively it may lose strength, and, thus, the gap between it and other possible explanations would narrow (or even disappear altogether), which shows that its strength and plausibility compared with rival explanations is weakening (or the others have risen in plausibility).[67] If the evidence supports one plausible explanation so overwhelmingly that one can *reject* all other explanations, this would be more a matter of good fortune than occurs in typical nondeductive arguments.[68] If two

---

[65] Kassirer & Cecil, *supra* note 48, at 1384. *See also* Jerome P. Kassirer, "Diagnostic Reasoning," *Annals of Internal Medicine* 110, 893-900 (1989); JEROME P. KASSIRER & R. I. KOPELMAN, LEARNING CLINICAL REASONING (Williams and Wilkins 1991).

[66] Thagard, *supra* note 14, at 102; WRIGHT, PRACTICAL REASONING, *supra* note 17, at 107.

[67] WRIGHT, PRACTICAL REASONING, *supra* note 17, at 103-06.

[68] For example, HARMAN, *supra* note 14, at 89-90, seems to be thinking of easy cases of nondeductive arguments in which one explanation is so

hypotheses have approximately the same plausibility, there might be no "best" explanation, but rather two equally plausible rival explanations.[69]

The overall strategy sketched above in the search for explanations is broadly similar across many fields that utilize nondeductive arguments. It is widely endorsed by epidemiologists,[70] toxicologists, methodologists inferring causes from well-analyzed case studies,[71] governmental scientists assessing risks, investigators seeking to explain airplane or space shuttle accidents, and ordinary persons making empirical inferences.[72]

C. Weight-of-the-Evidence Arguments

The strength of scientific inferences depends both on the truth of the evidentiary claims in the premises and on the *cumulative support that all the relevant evidence* contained in the premises offers for the conclusions in question.[73] Another name for these arguments is "weight-of-the-evidence" arguments.[74]

---

superior to all others that one can properly be said to *reject* them. One explanation can be better than another without rejecting the second.

[69] This point follows from Wright's analysis at WRIGHT, PRACTICAL REASONING, *supra* note 17, at 103-06.

[70] Austin Bradford Hill, *The Environment and Disease: Association or Causation?* 58 PROC. ROYAL SOC'Y OF MED. 295, 300 (1965), *reprinted in* EVOLUTION OF EPIDEMIOLOGIC IDEAS: ANNOTATED READINGS ON CONCEPTS AND METHODS 15-20 (Sander Greenland ed. 1987). *See also* Douglas Weed, *Underdetermination and Incommensurability in Contemporary Epidemiology*, 7 KENNEDY INSTI. ETHICS J. 107, 114 (1997).

[71] *See*, *e.g.*, Hutchinson & Lane, *supra* note 48, at 12.

[72] For a regulatory use of such inferences, *see* U.S. Environmental Protection Agency's Proposed Guidelines for Carcinogen Risk Assessment, 61 *Fed. Reg.*17, 960, 17,961 (Apr. 23, 1996) (to replace 51 Fed. Reg. 33992 when finalized). One can easily become aware of the implicit nondeductive arguments in use around us by noticing the argument form.

[73] WRIGHT, PRACTICAL REASONING, *supra* note 17, at 49.

[74] This is a term often used in regulatory settings, but is not restricted to them. Scientists often speak of what conclusion the "weight of the evidence supports." Frequent conversations with David A. Eastmond, Chair,

This is a term from both scientific and regulatory contexts.[75] The metaphor "weight-of-the-evidence" is intended to convey the persuasiveness of the kind and amount of evidence in favor of different conclusions. In regulatory science, for example, researchers might be concerned whether a substance is a human carcinogen. In such circumstances, scientists consider which rival conclusions are better supported by the weight of available evidence: Is the substance a human carcinogen; is it is a probable human carcinogen; is the evidence so equivocal that one cannot decide; or is it not a human carcinogen at all? The implicit question to be addressed is whether the weight of the available scientific evidence better supports the claim that a substance causes (or contributes to), or more likely than not causes (or contributes to) cancer or to support some other claim.[76] Scientists assessing the most likely cause of angiosarcoma in PVC plants concluded that the strongest explanation, which was quite good, was exposure to vinyl chloride monomers.[77]

D. Example: Plaintiffs' Argument in Allen v. Pennsylvania Engineering

To further illustrate the points about scientific arguments, consider in schematic outline an argument offered by plaintiffs in *Allen v. Pennsylvania Engineering*.[78] This statement of their

---

Environmental Toxicology, University of California, Riverside.

[75] For example, IARC researchers note that for the conclusions of a consensus scientific committee "the final overall evaluation [of evidence that a substance is carcinogenic to humans] is a matter of scientific judgment, reflecting the *weight of the evidence* derived from studies in humans, studies in experimental animals, and mechanistic and other relevant data." Cogliano et al., *supra* note 59, at 1272 (emphasis added).

[76] For a discussion of the weight of the evidence procedure, see the U.S. Environmental Protection Agency's Proposed Guidelines for Carcinogen Risk Assessment, 61 FED. REG.17,960, 17,961 (Apr. 23, 1996) (to replace 51 Fed. Reg. 33992 when finalized).

[77] Heath et. al., *supra* note 32.

[78] 102 F.3d 194 (5th Cir. 1996). This statement of their argument is truncated and does not do it full justice, but it illustrates the points above.

argument is truncated and does not do it full justice, but it illustrates the above mentioned points.

1. Walter Allen was periodically, but sporadically, exposed to ethylene oxide (ETO), a sterilizing agent, while changing ETO cylinders on hospital sterilizing units over a 20-year period; he contracted brain cancer ("BC").
2. Rat studies, but not mice studies, show that ETO causes comparatively rare BC and that it can cross the blood-brain barrier.
3. Several small human studies with low exposures suggest that ETO is associated with BC. A large meta-analysis shows no association.
4. ETO is a small molecule, a direct-acting alkylating agent (it interferes with DNA), and a multisite mutagen.

There are clearly several possible conclusions to this argument for *general causation*:

5a. It is more likely than not that ETO can cause human brain cancer.

5b. It is not more likely than not that ETO can cause human brain cancer.

5c. One cannot draw any very definitive conclusion from the data.

In addition, there are several possible conclusions concerning *specific causation* (addressing the question of whether Mr. Allen's brain cancer was likely caused by his exposure to ethylene oxide). Clearly, one would need to have some more developed premises concerning the extent of his exposure, but the possible conclusions concerning specific causation would be the following:

6a. It is more likely than not that ETO caused or contributed to Mr. Allen's BC.

6b. It is more likely that something else caused his BC: bad luck, unlucky genes, some other exposure, and so on.

6c. One cannot draw any very definitive inference about what caused his BC.

In addition to the premises of these arguments supporting several possible conclusions, plaintiffs' experts presented several different kinds of evidence as scientifically relevant to the inferences they needed to make, e.g., human, molecular and animal evidence in particular.[79] This evidence, together with exposure information taken together, they argued, more likely than not showed that ETO can cause brain cancer and did cause brain cancer in Mr. Allen.[80]

III. WHAT PROBLEMS MIGHT NONDEDUCTIVE INFERENCES POSE FOR JUDICIAL REVIEW?

Scientific arguments are not readily accessible to those not steeped in the substantive scientific fields, nor are they easy to evaluate. They have several features that make their assessment more difficult for those who are more distant from them on a day-to-day basis. First, nondeductive arguments will always have gaps between their premises and their conclusions. This is readily apparent in the arguments from *Allen*.[81] Courts will need to recognize that scientific arguments have such gaps and not expect deductive "tightness" in them.

Second, in utilizing scientific arguments, scientists consider all the *scientifically relevant* evidence in drawing their conclusions. There can be, I believe, temptations for those not familiar with the scientific substance of such arguments to not consider some pieces of evidence because they appear not to support conclusions by themselves or to contribute too little to conclusions. For example, Mr. Allen's experts were not permitted to utilize all of the scientifically relevant evidence in their argument because the court rejected the various pieces of evidence individually.[82] Thus, courts sometimes seem to struggle

---

[79] *Id.* at 196.

[80] Expert report by Karl T. Kelsey and Anthony D. LaMontagne for plaintiffs (October 13, 1992) in Allen v. Pennsylvania Eng'g Corp., 102 F.3d 194 (5th Cir. 1996).

[81] *See supra* Section II.D.

[82] Allen v. Pennsylvania Eng'g Corp., 102 F.3d at 198.

with scientifically relevant evidence on which scientists typically rely, or to review each "piece" of evidence for whether it strongly supports plaintiffs' ultimate causal conclusion.[83] Finally, in drawing conclusions, scientists consider *all the relevant evidence as an integrated whole*.[84] Since scientists view the evidence as a whole, courts will need to recognize this and assess the evidence as an integrated whole when reviewing expert testimony for admissibility (the *Allen* court did not permit this).[85] This does not ease courts' tasks, but arguably makes them more difficult because judges must review the *scientific substance* of an expert's argument (taking account of all the relevant evidence) to ensure that it is sufficiently plausible to assist a jury.

A. Gaps in Scientific Arguments

All nondeductive arguments will have gaps because there is no logically tight relationship between premises and conclusion.[86] The gaps in the argument from *Allen v. Pennsylvania Engineering* exist not merely because I presented a brief statement of more elaborate arguments. The gaps are traceable to the structure of the arguments themselves.[87] Even the best inferences can appear to have gaps, simply because it is easy to conceive alternative conclusions to the arguments.

For example, no matter how well established Newtonian gravitational theory is to describe the attraction of physical objects on the surface of the earth and near celestial bodies, one

---

[83] *Id*.

[84] *See supra* Section II.C.

[85] Courts must conduct reviews in this manner, if they are to assess scientific arguments (testimony) as scientists do.

[86] *Supra* Section II.A.

[87] There is no logically tight relationship between premises and conclusion. Moreover, there is a more subtle point: how elaborate and detailed such arguments need to be presented very much depends upon the substantive understanding of those to whom the argument is addressed. Comparable experts will need less elaborate arguments than less well informed readers.

might easily ask whether it applies to far away bodies with which we are not familiar.[88] For another example, a well-known astronomer, Thomas Gold, contrary to well-grounded views in geology, argued until his death that oil does not come from decomposed organic matter (such as plants and dinosaurs), but from geologic processes deep within the earth's core.[89] Also, no matter how well established evolution is, any relationships between premises based on existing evidence and conclusions are not so "tight" as to deter some from easily suggesting alternative theological explanations for the observed phenomena. This does not mean they are right, only that it is comparatively easy to imagine or find some kinds of gaps in the argument.

The possibility of gaps provides openings for skepticism about nondeductive arguments and may raise judicial concerns about them.[90] If premises support alternative conclusions, this can easily invite skepticism about a particular conclusion a scientist has drawn, even if the challenge is not particularly well-founded by evidence. It is effortless to be a skeptic; one needs only to suggest a different possible conclusion to the argument or notice a gap in the argument and exploit it.[91] The possibility of skepticism toward scientific arguments may reinforce any natural skepticism judges have toward the arguments of adversarial counsel and experts. Moreover, the obvious gaps in arguments will likely increase difficulties courts will have in reviewing them because it will correctly appear that

---

[88] At one time I suggested this as a speculative comment. However, in recent years, two physicists, Mordehai Hilgrom and Jacop Bekenstein, have proposed that the appropriate force equations for certain very distant galaxies is $F=ma^2/a_0$ (where F is force, m is the mass of the object, a^2 is acceleration and a_0 is a cosmological constant) instead of $F=ma$ (force equals mass times acceleration). This view would modify early Newtonian views. Adam Frank, *Gravity's Gadfly*, DISCOVER: SCIENCE, TECHNOLOGY, AND THE FUTURE, 32-37 (August 2006).

[89] Lissa Harris, *CU's Thomas Gold, noted astronomer and 'gadfly,' is dead at 84*, CORNELL CHRON. (Jul. 1,2004), *available at* http://www.news.cornell.edu/Chronicle/04/7.1.04/Tom.Gold.obit.html.

[90] This was a concern of the Supreme Court in *Gen. Elec. Co. v. Joiner*, 522 U.S. 136, 146 (1997).

[91] WRIGHT, PRACTICAL REASONING, *supra* note 17, at 104.

every such argument has a gap. However, the mere presence of gaps alone should not disqualify a scientific argument from being presented to a jury because every nondeductive argument has a gap. The challenge for courts is to review arguments in an informed and thoughtful manner.

In *General Electric Co. v. Joiner*, a gap in plaintiffs' argument was a concern of the district court (and ultimately the Supreme Court) in reviewing Joiner's weight-of-the-evidence argument (inference to the best explanation) that he had contracted lung cancer at least in part because of exposure to polychlorinated biphenyls (PCBs).[92] The Court noticed a "gap" between plaintiff's animal evidence and the experts' opinion.[93] It urged that there could be too great a "gap" between data on which experts rely and a scientist's opinion testimony.[94] But if there is always some gap, how should judges review such arguments? Moreover, can they review expert testimony without comparing plaintiff's conclusions with defendant's conclusions, as the original *Daubert* decision cautioned against?

First, not every nondeductive argument admits of *legitimate* critique simply because it has a gap. Some will be quite solid and strong—oil comes from decomposed organic matter—or some even implausible—Martians caused the space shuttle Challenger disaster. Second, in reviewing expert arguments courts will need to consider a scientist's conclusion in relation to the data and information on which it is based, but they should only consider a conclusion in relation to the evidence in the argument in order to assess arguments as scientists would.

Good argument evaluation would not license the conclusion to be considered in isolation from all the scientifically relevant evidentiary premises in support of it.[95] Moreover, the law on the admissibility of evidence does not permit a comparison between plaintiff's conclusion (or argument) and defendant's scientific

---

[92] Gen. Elec. Co. v. Joiner, 522 U.S. 136, 141, 146 (1997).

[93] *Id.*

[94] *Id.* The Court's reaction to this evidence was likely an overreaction. *See* CRANOR, *supra* note 45.

[95] WRIGHT, PRACTICAL REASONING, *supra* note 17, at 114.

conclusion (or argument), unless there is simply no issue of material fact to be resolved between parties.[96] Thus, for admissibility a judge would need to make some minimal assessment of the strength of a scientist's nondeductive argument to see whether it is sufficiently minimally plausible for legal purposes.[97] Following this step, some courts, beginning with the *Daubert* litigation and subsequent to it, appear to have engaged in a comparative weighing of plaintiff's evidence with defendant's evidence prior to trial, but that is controversial, simply because it appears to intrude on the jury's authority.[98]

There can be arguments, of course, in which all the data on which experts rely so poorly support the inferences drawn that one might say they have no support at all. When this occurs it is probably better to say not that there is "too great a gap between data and the conclusions drawn from them," but, rather, that the conclusion is simply too "speculative" given the evidence in the premises.[99]

---

[96] FED. R. CIV. P., 56 C.

[97] *See also supra* Section IV.

[98] *See, e.g.,* Logerquist v. McVey, 1 P.3d 113, 131 (Ariz. 2000) ("The *Daubert/Joiner/Kumho* trilogy of cases . . . puts the judge in the position of passing on the weight or credibility of the expert's testimony, something we believe crosses the line between the legal task of ruling on the foundation and relevance of evidence and the jury's function of whom to believe and why, whose testimony to accept, and on what basis."); Howerton v. Arai Helmet, Ltd., 597 S.E.2d 674, 692 (N.C. 2004) ("[T]rial courts asserting. . .authority under *Daubert* may unnecessarily encroach upon the constitutionally-mandated function of the jury to decide issues of fact and to assess the weight of the evidence."); Bunting v. Jamieson, 984 P.2d 467, 472 (Wyo. 1999) (adopting *Daubert*, but nonetheless expressing concern that applying the Daubert approach to exclude evidence has been criticized as a "misappropriation of the jury's responsibilities," and that "'it is imperative that the jury retain its fact-finding function'" (*quoting* Springfield v. State, 860 P.2d 435 (Wyo.1993)); Brasher v. Sandoz Pharm. Corp., 160 F. Supp. 2d 1291, 1295 (N.D. Ala. 2001) (applying *Daubert*, but noting that the jury's right to decide the facts of the case is usurped when a trial court "overreach[es] in the gatekeeping function" and "determine[s] whether the opinion evidence is correct or worthy of credence.").

[99] Michael J. Saks, *The Aftermath of* Daubert*: An Evolving Jurisprudence of Expert Evidence*, 40 JURIMETRICS J. 229, 236 (2000).

B. Potential Problems with Scientific Relevance

A second set of potential problems with scientific inferences concerns the scientific relevance of studies, data, and information on which an expert bases her testimony. Courts have had some difficulties with animal studies, case reports and molecular evidence. A generic difficulty appears to be that courts have engaged in assessing *individual* pieces of evidence as a means of reviewing an experts' *overall testimony* (or argument).[100] Such a method for analyzing scientific arguments confuses an assessment of the overall argument with an assessment of one piece of the evidence (or one premise) in the argument.[101] It conflates the analysis of the argument *as a whole* with analysis of *the evidence in one premise of the argument*. Such an approach is fraught with numerous difficulties.

In order to see these difficulties, several distinct questions should be introduced. First, is a particular piece of evidence *relevant* to assist a scientific inference about toxicity? Second, if it is scientifically relevant, how much *weight* or (in legal terms) *probative value* does, and should, it have in an overall inference about toxicity?[102] Third, does one piece of evidence *taken by itself* provide sufficient support for a scientist's *toxicity conclusion*?[103] Finally, does all the relevant evidence considered provide a sufficiently plausible argument for the expert's conclusion, so he or she may be admitted to testify to assist the jury? Failure to distinguish at least these four questions can contribute to confusion about the issues and frustrate scientifically accurate reviews of expert testimony.

One issue is whether a piece of scientific evidence is scientifically *relevant* to an inference of toxicity or to a

---

[100] As explained below, this clearly occurred in *Allen v. Pa. Eng'g Corp.*, 102 F.3d 194, 197 (5th Cir. 1996), and seemed to be endorsed by the Supreme Court in *General Elec. Co. v. Joiner*, 522 U.S. 136, 142-46 (1997), which led Justice Stevens to register his dissent at 150-55.

[101] Confusing court language in the original *Daubert* decision as well as subsequent commentary and court decisions has contributed to this.

[102] WRIGHT, PRACTICAL REASONING, *supra* note 17, at 123.

[103] *Id.*

judgment of the relative plausibility of different explanations of the evidence. As noted above, it should be easy for evidence to satisfy the relevance criteria—if it would have *any impact,* positive or negative, on the plausibility of different explanations, it is relevant.

Before introducing the other points, reconsider the argument from *Allen v. Pennsylvania Engineering*.[104] The Fifth Circuit Court of Appeals reviewed a judge's exclusion of plaintiffs' experts.[105] The court of appeals held that epidemiological studies did not show ETO caused or contributed to brain cancer in humans.[106] Next, referring to *Brock v. Merrell-Dow Pharmaceuticals, Inc*,[107] the Court noted that animal studies "must be carefully qualified in order to have explanatory potential for human beings."[108] Then, quoting defense experts, the court argued that the rat studies could not be reliably extrapolated to humans, holding that studies showing that rats contracted brain cancer, "furnishes at best speculative support for appellants' causation theory."[109] Finally, the Court argued that while the cell biology shows ETO to have mutagenic and genotoxic properties in living systems, this does not necessarily show it can cause brain cancer in humans or did cause brain cancer in Mr. Allen.[110] The court concluded: "[N]one of the

---

[104] *See supra* Section II.D.

[105] Plaintiffs used a more elaborate version of the argument sketched above in *supra* Section II.D.

[106] Allen v. Pa. Eng'g Corp., 102 F.3d 194, 197 (5th Cir. 1996) ("First, although occupational exposure to ETO has been studied for many years, not a single scientific study has revealed a link between human brain cancer and ETO exposure.").

[107] Brock v. Merrill-Dow Pharm., Inc., 874 F.2d 307, 313 (5th Cir. 1989), *modified,* 884 F.2d 166 (5th Cir. 1989).

[108] Allen v. Pa. Eng'g Corp., 102 F.3d at 197.

[109] *Id*. ("[T]he lack of capacity for the F-344 rat to predict how even the mouse model responds necessarily undercuts confidence that the rat will predict accurately how other species including humans will respond [to EtO exposure].").

[110] *Id*. at 198. ("Third, the cell biology data show only that ETO has mutagenic and genotoxic capabilities in living organisms, not that it necessarily causes brain cancer in humans or in Allen's particular case. That

scientific data on which appellants' experts rely furnishes a scientifically valid basis for the conclusion they would draw. The paucity of epidemiological evidence, the unreliability of animal studies, and the inconclusiveness of cell biology combine to undercut the expert testimony."[111]

We see in this decision some court language that is fairly common. Instead of assessing the scientists' argument (their testimony) based on all the relevant evidence, the court addressed each piece of evidence by itself. One reason for this approach may be that the judges were concerned about the scientific relevance of the individual pieces of evidence to the expert's arguments, and whether plaintiffs' expert could use them to draw a conclusion. They did not use this language, however.

Moreover, given the idea of scientific relevance, judges should exercise considerable care in assessing whether individual pieces of evidence are scientifically relevant, simply because they are so intellectually distant from the relevant fields and because criteria of both scientific and legal relevance are easy to satisfy.[112] Of course, they could query experts on such issues and inquire about how the evidence might affect (however slightly) their argument.[113] That is, they could ask, "How does this piece of evidence contribute (however slightly) to your argument?"

Could the judges in *Allen* have been concerned that each piece of evidence was so poor that it could not even have been scientifically relevant for the toxicity judgment? Their assertions about animal evidence could have been so construed: they held that the fact that ethylene oxide caused brain tumors in rats could not be evidence for the claim that ETO could cause brain tumors in humans, simply because ETO did not correspondingly

---

ETO may have these effects on living cells or genes is the beginning, not the end of the scientific inquiry and proves nothing about causation without other scientific evidence.").

[111] *Id.*

[112] See *supra* notes 49-50.

[113] They would in effect have a discussion on how such evidence was *relevant* to the conclusion for which they argued.

cause brain tumors in phylogenetically similar mice.[114] That is, they might have believed the rat evidence was simply scientifically irrelevant to humans. Had they explicitly used such an argument that would have been clearer. However, if this was their concern, there is independent evidence that they were mistaken on scientific grounds.[115] This is a risk of non-scientists reviewing the details of scientific arguments.

The court might not have understood (because it was too distant from the science), or plaintiffs might not have adequately

---

[114] Allen v. Pa. Eng'g Corp., 102 F.3d 194, 197 (5th Cir. 1996).

[115] As part of an NSF-funded research project at UC Riverside, David Eastmond and I sought scientific peer reviews of expert reports in a small number of legal cases. *Allen* was one. We sent the plaintiffs' and defendants' expert reports without names or affiliations attached to two extramural referees who were experts in the toxicology of ETO and one epidemiologist. Following the lead of some federal judges and the language from *Kumho Tire* we asked them whether the experts' opinions fell within a range where "reasonable experts would disagree." Both an industry oriented scientist and a non-industry scientist agreed that plaintiffs presented good scientific arguments that ethylene more likely than not could cause brain cancer. Consider just one expert's view.

> The evidence presented by these experts clearly establishes ETO as a carcinogen with a high likelihood of human risk. They appropriately cited literature showing that ETO is a direct-acting DNA alkylating agent, is mutagenic in multiple in vivo and in vitro studies including human cells, and consistently showed induction of chromosomal damage in peripheral lymphocytes of exposed workers (chromosomal aberrations, sister chromatid exchanges, and micronuclei). ETO also induces heritable translocations in rodents (not mentioned by the plaintiff's experts). The plaintiff's experts also cited studies showing tumor induction at multiple sites, including brain, in male and female rats exposed to EtO, and they cited studies demonstrating that EtO forms DNA adducts in the brain. The latter piece of information is important because it demonstrates that EtO can cross the blood brain barrier. (Peer review of plaintiffs' expert's report by anonymous reviewer, Jan. 12, 2004) (any emphases in the original).

Moreover, one expert noted (and then went on to illustrate the view) that "The defendant's experts made several unjustified assumptions and misstatements to support their contention that the plaintiff's brain tumor was not due to exposure to ETO." (Peer review of defendants' expert's report by anonymous reviewer, Jan. 12, 2004). This issue is discussed more fully in CRANOR, *supra* note 45 at 324-28.

explained, the significance of the rat studies even in the absence of similar results from mouse studies. It can be argued that in this case rats would be a better model for predicting effects in humans simply because rats generally have a slower metabolism and breathing rate than mice, thus retaining ETO in their bodies more like humans.[116] The rat studies also show that the small molecule of ethylene oxide can cross the blood-brain barrier, something that it is typically difficult for chemicals to do.[117]

If ETO can have this biological effect in rats, which have a metabolism that is more similar to humans than mice, it is plausible that ETO can have this biological effect in humans, which is what plaintiffs had argued.[118] Toxicologists would explain ethylene oxide's inability to cross the mouse blood-brain barrier as based on special features of mice that make them different from rats and humans.[119] By contrast, rats' slower metabolism and respiratory rates would result in them retaining ETO longer, giving that small molecule time for absorption into various tissues, including the brain.[120] The *Allen* court seemed to *assume* that there was something special about rats, which was not applicable to mice and to humans, when the opposite appears to be true: the mice have special features that distinguish them

---

[116] Interview with David A. Eastmond, (June 2003).

[117] The blood-brain barrier can be characterized as one that:

> is not an absolute barrier to the passage of toxic agents into the CNS [Central Nervous System]. Instead, it represents a site that is less permeable than are most other areas of the body. Nevertheless, many poisons do not enter the brain in appreciable quantities because of this barrier.

Karl K. Rozman & Curtis D. Klasassen, *Absorption, Distribution, and Excretion of Toxicants, in* CASARETT AND DOULL'S TOXICOLOGY 6th ed., 86 (Curtis D. Klaassen ed., McGraw-Hill 2001) 107-32, 122. The blood-brain barrier is a physiological barrier that seems to have evolved to provide protections to the brain. *Id*. at 122-23.

[118] Karl T. Kelsey & Anthony D. LaMontagne, Plaintiff's Expert Opinion Aff.; Allen v. Pennsylvania Eng'g Corp., 102 F.3d 194 (5th Cir. 1996).

[119] Interview with David A. Eastmond, Chair, Environmental Toxicology, University of California, Riverside, May 2003.

[120] *Id*.

from rats and make their responses less applicable to humans.[121]

However, the court did not use the language of relevance, but instead sometimes asked whether *each piece of evidence* was reliable.[122] This suggests its concern might have been a) whether one piece of evidence is or can be sufficiently reliable by itself to support an expert's overall conclusion, perhaps b) whether by itself it can provide major support for a conclusion or maybe c) whether it unerringly contributes support to a conclusion.[123] Despite some passages in the Supreme Court's *Daubert* opinion, the discussion of animal evidence in *Joiner* and some other courts' opinions, this is puzzling. No single piece of evidence about toxicity is ever likely to support a conclusion—such an expectation is both scientifically mistaken and an unreasonable interpretation of nondeductive arguments.[124] Moreover, if a scientist or a court can challenge each piece of evidence individually as insufficiently reliable to support an expert's conclusion, this would almost certainly undermine *every* scientific argument because they are typically based on multiple premises (pieces of evidence) in support of conclusions

For example, Watson and Crick's well-known paper on the structure of DNA rested on several considerations that individually seemed like comparatively weak evidence.[125] However, taken together, Susan Haack argues, the evidence supports an inference concerning the "structure of DNA [that] is very well warranted (in fact, the only entry that fits)."[126] This

---

[121] CRANOR, *supra* note 45, at 325.

[122] Allen v. Pennsylvania Eng'g Corp., 102 F.3d 194, 198 (5th Cir. 1996).

[123] *See id*. at 197-98 (discussing possibilities of interpretation).

[124] In a typical argument a single piece of evidence might appear to be very important (and in fact can be very important), but it typically has this significance because of the presence of other pieces of evidence utilized in the inference. *See* Carl F. Cranor & David A. Eastmond, *Scientific Ignorance and Reliable Patterns of Evidence in Toxic Tort Causation: Is There a Need for Liability Reform?*, LAW & CONTEMP. PROBS. 5, 34-41 (2001). *See also* CRANOR, *supra* note 45, at 313-14 for discussion of some examples.

[125] *See supra* notes 185-86 and accompanying text.

[126] Susan Haack, *An Epistemologist in the Bramble-bush: At the Supreme*

argument, as we now know, has revolutionized biology. Had the review procedures of some judges been applied to this ground-breaking paper, it might have died a premature and mistaken death.

Finally, no *kind* of evidence—epidemiological, animal, or molecular studies—unerringly supports a conclusion concerning toxicity. Whether a particular piece of evidence does *contribute* to a conclusion depends upon what it shows, as well as the other evidence utilized in the argument. Sometimes molecular evidence can be quite strong as it is for the toxicity of ethylene oxide[127] and sometimes not. Sometimes case reports can be especially strong evidence as they were in identifying the toxicity of vinyl chloride and sometimes not.[128]

There is a cluster of deeper and more disturbing issues.

Court with Mr. Joiner, 26 J. HEALTH POL. POL'Y & L 217, 237 (2001). *See also supra* note 102.

[127] *See* discussion and accompanying *supra infra* notes 77-79.

[128] The idea of reliable evidence is puzzling and troubling as it has come to be utilized in legal decisions. The Court in *Daubert* might have had in mind the reliability of underlying tests in support of a legal case as defendants argued the blood pressure test was reliable in *Frye*. This was a test or technique that directly addressed one of the key legal issues of that case—was the defendant telling the truth? Moreover, it is possible that such a test could be quite accurate, e.g., 90 percent accurate, in providing evidence about who was telling the truth and who was lying, but of course it was not. However, using reviews of such techniques as an analogy for reviewing a variety of pieces of evidence that would assist scientists in coming to judgments about the toxicity of a substance is likely to be misleading. Sometimes individual structure-activity or mechanistic evidence can quite helpful in assessing the toxicity of a product when combined with other information and sometimes not. CRANOR, *supra* note 45, at 111-15, 245-48. Sometimes an epidemiological study will be very helpful when it shows a positive association between exposure and disease and sometimes not. Usually negative or "no effect" epidemiological studies are quite misleading if they are taken to provide evidence that there is no adverse effect from exposure. *Id*. at 27, 243-45, 264, 277. Individual pieces of toxicity evidence do not lend themselves readily to being judged "reliable" or not for the ultimate toxicity questions; *their contribution to a toxicity judgment is properly assessed in the presence of all the other evidence in support of a toxicity claim and how well all the evidence taken together supports the toxicity conclusion*.

Several courts appear to have rejected whole categories or kinds of evidence as insufficiently reliable to support a scientist's conclusions. Some have rejected animal studies, and molecular evidence, and many have rejected human case reports.[129] Each of these can easily be scientifically relevant evidence, well-endorsed by scientists, and, depending upon the other evidence available in a particular case, each can be especially powerful.

1. Animal Studies

A variety of considerations point to the scientific relevance and probative value of animal studies in making toxicity judgments. National and international consensus scientific committees routinely rely upon animal evidence for judging the toxicity of substances. Consider carcinogens as an example, which is an especially well developed area. The International Agency for Research on Cancer lists about sixty-six substances or groups of substances, excluding mixtures and exposure conditions, as probable human carcinogens.[130] For more than

---

[129] On animal studies, for example, consider two cases: Plaintiff's expert . . . relied on a study of the effect of picloram on rats that showed that when exposed to large amounts of the chemical, the rats developed cancerous tumors and died. He admitted that the effects of chemicals differ between humans and rats . . . . We then are left to conclude that the study, at most, is only evidence that picloram may produce some unidentified effect on humans.

Viterbo v. Dow Chem. Co., 826 F.2d 420, 424 (5th Cir. 1987). "The animal studies are not helpful in the instant case because they involve different biological species. They are of so little probative force and are so potentially misleading as to be inadmissible." In re Agent Orange Prod. Liab. Litig., 611 F. Supp. 1223, 1241 (E.D.N.Y. 1985). For some discussion of case reports see CRANOR, *supra* note 45, at 256-57 and Section B.3 *infra,* at notes 162-80. Other courts have admitted testimony based on animal studies. Some of these are summarized in *In re Paoli Railroad Yard PCB Litig.*, 35 F.3d 717, 780 (3d Cir. 1994).

[130] IARC, *Overall Evaluation of Carcinogenicity to Humans, Group 2A: Probably Carcinogenic to Humans, Group 2A: Probably Carcinogenic to Humans*, MONOGRAPH SERIES. REV. (July 7, 2004); *available at*

forty of these substances (about 60 percent), evidence of carcinogenicity in humans is inadequate or limited.[131] Nonetheless, the overall classification is based on sufficient evidence in animal studies plus "other data relevant to the evaluation of carcinogenicity and its mechanisms."[132] The U.S. National Toxicology Program lists about 185 substances as "reasonably anticipated to be a human carcinogen." Of these, a large percentage has been identified on the basis of good animal studies.[133] This has been confirmed by agency personnel.[134]

The Environmental Protection Agency ("EPA") classifies substances as probable human carcinogens based upon animal studies, even if the evidence that a substance is carcinogenic to humans is inadequate.[135] These are substances "likely to produce cancer in humans due to the production or anticipated production of tumors by modes of action that are relevant or assumed to be relevant to human carcinogenicity."[136]

Scientific principles underscore the importance of animal studies and point to the foundation for the use of animal studies described above. Huff and Rall summarize considerable science:

> From data available so far, therefore, it appears that chemicals that are carcinogenic in laboratory animals are likely to be carcinogenic in human populations and that, if appropriate studies can be performed,

---

httyp://www-cie.iarc.fr/moneval/crthgr02a.htm (listing 66 substances that are known human carcinogens).

[131] That is, studies would be of insufficient quality to permit an inference concerning human carcinogenicity, or the association is credible, but alternative explanations of the positive results cannot be ruled out with sufficient confidence to justify a causal inference. *See id.*

[132] *See id.*

[133] Criteria were first listed and published on September 26, 1996, and are listed at the NTP Web site, http://ntp.niehs.nih.gov/ntpweb/index.cfm?objectid=03C9CE38-E5CD-EE56-D21B9.

[134] Interview with Ronald Melnick, National Institute of Environmental Health Science, in Brooklyn N.Y. (March 3, 2006).

[135] Environmental Protection Agency Proposed Guidelines for Carcinogen Risk Assessment, 61 Fed. Reg. 17, 17960, 17985 (Apr. 23, 1996).

[136] *Id.*

> there is qualitative predictability. Also, there is evidence that there can be a quantitative relationship between the amount of a chemical that is carcinogenic in laboratory animals and that which is carcinogenic in human populations.[137]

This conclusion is supported by more specific relationships identified by scientists. Although there are readily apparent differences between laboratory animals and humans (such as external physical characteristics, lifespan, metabolic rate, and heterogeneity) that too often receive greater attention than similarities, "experimental evidence to date certainly suggests that there are more physiologic, biochemical and metabolic similarities between laboratory animals and humans than there are differences."[138] The "biological processes of molecular, cellular, tissue, and organ functions that control life are strikingly similar from one mammalian species to another."[139] In addition, based upon current information, there is great similarity in the carcinogenic processes between animals and humans.[140] Furthermore, "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."[141] The toxicology authors in the Federal Judicial Center Manual on Scientific Evidence and the EPA agree.[142] A 2005

---

[137] James Huff & David P. Rall, *Relevance to Humans of Carcinogenesis Results from Laboratory Animal Toxicology Studies, in* Maxcy-Rosenau Last Public Health & Preventive Medicine 433, 437 (John M. Last & Robert B. Wallace eds., 13th ed. 1992).

[138] David P. Rall et al., *Alternatives to Using Human Experience in Assessing Health Risks,* 8 ANN. REV. PUBL. HEALTH 355, 356 (1987).

[139] *Id.* at 434.

[140] Some researchers make even stronger claims. For example, see James Huff, *Chemicals and Cancer in Humans: First Evidence in Experimental Animals*, 100 ENVTL. HEALTH PERSP. 201, 204 (1993) (stating that the array and multiplicity of carcinogenic processes are virtually common among mammals, for instance between laboratory rodents and humans).

[141] *Id.* at 204.

[142] See Bernard D. Goldstein & Mary S. Henifin, *Reference Guide on*

Institute of Medicine and National Research Council report gives animal studies a strong endorsement, echoing numerous earlier reports. According to this report, animal studies are:

> powerful because controlled studies can be conducted to predict effects that might not be detected from customary use by humans until they result in overt harmful effects. Animal studies are especially useful in detecting effects of chronic exposures and effects on reproductive and developmental processes because epidemiological methods of studying humans are especially problematic in these areas . . . .[143]

Recall also how animal studies were especially strong evidence in identifying vinyl chloride monomer as a human carcinogen.

Despite these findings, courts may nonetheless be concerned about inferences from animals to humans. However, such concerns can be misplaced. Consider a hypothetical example similar to one used by the Federal Judicial Center and the Judicial Council of California to assist in educating judges about scientific evidence:

> Suppose a hypothetical plaintiff Mr. Jones was exposed to XXBC in drinking water. As evidence for the toxicity of XXBC, suppose in one study mice exposed to 5 milligrams of substance XXBC per kilogram of body weight have approximately six times the rate of liver cancer as unexposed mice. The

---

Toxicology, *in* REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 421 (Federal Judicial Ctr. ed., 2d ed. 2000) and Environmental Protection Agency Proposed Guidelines for Carcinogen Risk Assessment, *supra* note 72, at 17977 ("[T]here 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 [i.e., must be in the same tissue in rodents and humans].").

[143] INSTITUTE OF MEDICINE AND NATIONAL RESEARCH COUNCIL, COMMITTEE ON THE FRAMEWORK FOR EVALUATING THE SAFETY OF DIETARY SUPPLEMENTS, DIETARY SUPPLEMENTS: A FRAMEWORK FOR EVALUATING SAFETY 157 (National Academies Press 2005) [hereinafter DIETARY SUPPLEMENTS].

> disease rate in the exposed mice was .20 (18 of 90 mice in this group had tumors) compared with the control group (3 of 90 had tumors, a disease rate of .03). Thus, the disease rate in mice attributable to XXBC is 17/100 or 17 percent. The dose to which the experimental mice were exposed was 170 times that to which Mr. Jones was exposed. Suppose there were similar results in rat studies. Suppose also that the evidence from the two rodent studies suggests that XXBC was a genotoxic carcinogen and that there was a linear relationship between dose and response.[144]

In addition, although were no statistically significant human studies exist, Mr. Jones and a few others who were sampled in the area had DNA damage consistent with DNA damage seen in animal studies, but these studies were not statistically significant.[145] Major alternative explanations of Mr. Jones' liver cancer (hepatitis B and exposure to aflatoxins) could be ruled out with reasonable confidence.[146]

Consider only the exposure of mice compared with exposures of humans. Many people may believe because rodents receive higher doses of a substance than those to which humans are exposed, that such studies are irrelevant to humans. However, this may result from a misunderstanding of the studies, a considerable public relations campaign against them, and, in any case, is often a red herring. Given the above information, scientists can calculate the likely disease rate for humans who had lower environmental exposures to XXBC. It is simply the disease rate in the animals divided by the much higher dosage they received because the toxicant acts by means of a linear mechanism. Thus, in the exposed animals the disease rate was 17/100 or 17 percent. If there were a particular cancer in 17 percent of the population, it would be among the very

---

[144] University of California/Judicial Council Summit, CASE STUDY VIGNETTE: Toxic Tort, Oct. 3, 2006.

[145] *Id.*

[146] *Id.*

highest cancer rates![147] Discounting the disease rate in humans by their lesser dose (170 times smaller) would yield a disease rate in an exposed human population (at the lower dosage) of 1/1000. A cancer rate of 1/1000 in an exposed human population would have to be compared with the rate in the general unexposed human population to see whether the relative risk was sufficiently high to merit legal compensation, and scientists would have to determine whether the particular individual had other possible exposures or conditions that would have increased his liver disease rate.

However, to make things simple, if liver cancer rate in the general population were 1/2000, those exposed to XXBC (at the assumed level) would have double the risk of liver cancer compared with the general population (or a relative risk of 2:1). If the disease rate in the general population were lower, e.g., 1/5000, the relative risk for those exposed to XXBC would be even greater (5:1). This would be quite a high relative risk.

To conclude, animal studies are relevant evidence for addressing the toxicity of a substance for general causation in the law because scientists routinely rely upon them for the reasons given above. Moreover, particular animal studies should not be disqualified as evidence for causation simply because animals are exposed to higher doses of a toxicant than humans. In particular, one would not tolerate a disease rate in humans as high as 17 percent of the exposed population as existed among rodents in the hypothetical example above. Finally, animal evidence can be quite powerful for general causation and possibly for specific causation in the law, as the above example shows, when appropriate extrapolations can be made. Of course, not all animal studies will necessarily be as valuable as the hypothetical, but animal evidence is often relevant and powerful, especially for carcinogens and reproductive toxicants.

---

[147] For comparison, a woman's chances of contracting breast cancer between birth and 70 years of age is 7.13% and her chances of contracting it between birth and death is 13.22%. A man's chances of contracting prostate cancer between birth and 70 years of age is 14.51% and of contracting it between birth and death is 17.93%. American Cancer Society, Surveillance Research, http://www.cancer.org/downloads/stt/CAFF06Prob.pdf.

2. Molecular Evidence

A second kind of toxicological evidence that can be quite important, but that fares badly in courts, is evidence about the molecular structure and its attendant biological activity in mammalian systems. For example, the *Allen* court struggled with the relevance of molecular studies, noting that they were "the beginning, not the end of the scientific inquiry and proves nothing about causation without other scientific evidence."[148] The court excluded the molecular data and subsequently the expert as well.[149]

A standard toxicology textbook notes the importance of molecular evidence: "An agent's structure, solubility, stability, pH sensitivity, electrophilicity, volatility and chemical reactivity can be important information for hazard identification [that is, for identifying hazards caused by substances]."[150] "Historically, certain key molecular structures have provided regulators with some of the most readily available information on the basis of which to assess hazard potential."[151] These include information about some carcinogens, structural alerts for "aromatic amine groups," and certain dyes as potential carcinogens. Some provide important information about developmental toxicants.[152]

The Institute of Medicine and National Research Council, addressing the toxicity of dietary supplements, highlights the underlying scientific rationale for the significance of structure-activity data for identifying adverse effects of toxicants:

> The physical-chemical properties and biological effects of a substance are derived from its chemical

---

[148] Allen v. Pa. Eng'g Corp., 102 F.3d 194, 198 (5th Cir. 1996).

[149] This court might have believed that the molecular evidence was scientifically relevant, but inadequate in the absence of what it saw as other needed evidence to support the causation claims (some of its arguments suggest this point). *Id* at 197-98.

[150] ELAINE M. FAUSTMAN & GILBERT S. OMENN, *Risk Assessment*, *in* CASARETT AND DOULL'S TOXICOLOGY 6th ed., 86 (Curtis D. Klaassen ed., McGraw-Hill 2001).

[151] *Id*.

[152] *Id*.

> structure. If the chemical structure of a dietary supplement is known, but additional insight into the biological activity is needed, then it is scientifically appropriate to consider the information about the biological activity of structurally related substances. It is assumed that the biological effects of chemicals, including toxic effects, are implicit in their molecular structures (referred to as toxicophores when they are associated with toxic effects). This concept is most clearly illustrated with the example of ephedra, which is considered by some scientists to have similar physiological actions, although less potent, to the chemically related substance amphetamine, as well as the recently banned pharmaceutical agent phenylpropanolamine.[153]

Scientists also recognize that certain classes of structure-activity relationships have been quite important in identifying chemical groups that are known to interact with mammalian DNA or proteins. Such relationships provide *strong*, but not quite infallible, reasons for thinking that substances with chemical similarities have similar biological activity.[154]

Courts can be quite dismissive of molecular or chemical structure data.[155] One possible reason is that similar chemical structures are not mathematically certain guides to similar toxicity effects, but mathematical certainty is not required in tort

---

[153] DIETARY SUPPLEMENTS, *supra* note 143, at 205-06, *citing* Food and Drug Administration, 69 Fed. Reg. 6787, 6787-854 (2005); I. Furuya & S. Watanabe, *Discriminative Stimulus Properties of Ephedra Herb (Ephedra Sinica) in Rats*, 13 YAKUBUTSU SEISHIN KODO 33, 33-38 (1993); C. R. Lake & R. S. Quirk, *CNS stimulants and the Look-Alike Drugs*, 7 PSYCHIATRY CIN. NORTH AMERICAN 689, 689-701 (1984).

[154] FAUSTMAN & OMENN, *supra* note 150, at 83-104; J. ASHBY & 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*, 204 MUTATION RESEARCH 17, 17-115 (1988).

[155] As occurred in Allen v. Penn. Eng'g Corp., 102 F.3d 194, 198 (5th Cir. 1996).

cases, given the tort law standard of proof. Chemical structure-biological activity data is ordinarily scientifically relevant evidence that can assist toxicity judgments.[156] Whether it is or not will depend upon the particular evidence in question—it should not be dismissed because it is a particular kind or category of evidence. In addition, sometimes such evidence can contribute substantial evidence of causation for some classes of substances, as it could have in *Allen v. Pennsylvania Engineering*.[157] In any case, such data should be part of scientifically reasonable patterns of evidence of causation, if it is relevant. How much scientific weight, probative value, or evidentiary strength molecular data has for a scientific argument in individual cases will depend on the substance, its properties, adverse effects, and other evidence that is available. Following the *Daubert* mandate, if the courts are to make the law better comport with the pertinent science in a case, they must recognize the *scientific relevance* and sometimes *quite strong evidentiary weight* of structure-activity relationships and other molecular properties.

The structure-activity evidence for ethylene oxide (ETO) in *Allen v. Pennsylvania Engineering* was especially powerful evidence of its particular toxicity. ETO is a multisite mutagen (that is, it causes DNA mutations in many tissues), a quite significant biological feature of a substance.[158] Moreover, because it is a small molecule that requires no metabolic

---

[156] CRANOR, *supra* note 45, at 111-12; IARC, WORLD HEALTH ORGANIZATION, *Preamble* to MONOGRAPH SERIES, Section 4 (a description is provided of any structure-activity relationships that may be relevant to an evaluation of the carcinogenicity of an agent, the toxicological implications of the physical and chemical properties, and any other data relevant to the evaluation that are not included elsewhere.).

[157] *See* discussion and accompanying text *infra* notes 200-01.

[158] "Mutagenicity testing, combined with an evaluation of chemical structure, has been found to identify a large proportion of trans-species, multiple-site carcinogens," R. JULIAN PRESTON & GEORGE R. HOFFMANN, *Genetic Toxicology*, *in* CASARETT AND DOULL'S TOXICOLOGY 342 (Curtis D. Klassen et al. eds., 6th ed. 2001).

transformation to produce toxic effects (it is "direct-acting"),[159] it could reach nearly any target site in the body. The rat studies utilized in *Allen* noted above showed it could cross the blood-brain barrier and reach the brain.[160] Contrary to the view of the court, the ETO molecular evidence provided particularly powerful evidence of ETO's toxicity.[161]

3. Case Reports

Case studies or case reports typically "arise from a suspicion, based on clinical experience, that the concurrence of two events—that is, a particular exposure and occurrence of a cancer—has happened rather more frequently than would be expected by chance."[162] Case reports for vaccines and drugs are part of what health professionals call "passive reporting schemes that rely on the vigilance of health care providers to detect events that are felt to be due to the administration of a drug product."[163] They also are one of five major kinds of evidence

---

[159] "Direct-acting carcinogens are typically carcinogenic at multiple sites and in all species examined. A number of the direct-acting alkylating agents, including some used in chemotherapy, are carcinogenic for humans." (Vainio et al., 1991). HENRY C. PITOT III AND YVONENE P. DRAGAN, *Chemical Carcinogenesis*, *in* CASARETT AND DOULL'S TOXICOLOGY 681, 686 (Curtis D. Klassen et al. eds., 6th ed. 2001).

[160] This was plaintiffs' experts' argument, reported in CRANOR, *supra* note 45, at 324-28.

[161] This issue is described in greater detail in CRANOR, *supra* note 45, at 324-28. Moreover, as part of NSF Grant #99-10952, David Eastmond and I had the plaintiffs' and defendants' experts reports in *Allen v. Pennsylvania Engineering* reviewed by two extramural experts, who judged that plaintiffs' arguments that ETO could cause brain cancer in humans (based largely on the animal studies and information about the molecular properties of ETO) were within a range where scientists could reasonably disagree (and in fact offered the kinds of arguments that consensus scientific committees typically hear about carcinogens).

[162] IARC, WORLD HEALTH ORGANIZATION, *Preamble* to MONOGRAPH SERIES, Section 8.

[163] J. P. Collet et al., *Monitoring Signals for Vaccine Safety: The Assessment of Individual Adverse Event Reports by an Expert Advisory Committee,* 78 pt. 2 BULL. OF THE WORLD HEALTH ORG 178 (2000).

utilized in occupational settings to identify toxicants.[164]

Case studies function best to reveal adverse causal reactions to vaccines, drugs, poisons, some anesthetics, and even dietary supplements.[165] Their evidentiary value tends to increase when there is a fairly short interval between exposure and reaction and where adverse reactions are reasonably easily identified.[166] However, they have also been used to identify adverse reactions from occupational exposures (recall the discussion of vinyl chloride above).[167]

Moreover, the Institute of Medicine Committee on the adverse effects from vaccines concluded that

> [I]n the absence of epidemiologic studies favoring acceptance of a causal relation, individual case reports and case series were relied upon, provided that the nature and timing of the adverse event following vaccine administration and the absence of likely alternative etiologic candidates were such that a *reasonable certainty of causality* could be inferred . . . from one or more case reports. The presence or absence of demonstrated biologic plausibility was also considered in weighing the overall balance of evidence for and against a causal relation.[168]

To see how a case report can be especially good evidence, consider the following. A 42-year-old man developed Guillain-Barré Syndrome (GBS), an acute inflammatory demyelinating polyneuritis following three independent tetanus shots over a 13-year period. This disease "is characterized by the rapid onset of

---

[164] Peter S. Thorne, *Occupational Toxicology*, *in* CASARETT AND DOULL'S TOXICOLOGY 1131-32 (Curtis D. Klassen et al. eds., McGraw-Hill Med. Pub. Division 6th ed. 2001).

[165] DIETARY SUPPLEMENTS, *supra* note 143, at 131-32.

[166] *Id.*

[167] The presentation on case studies is developed more fully in CRANOR, *supra* note 45, at Chapter 4.

[168] INSTITUTE OF MEDICINE, CHILDHOOD VACCINES 31 (emphasis added).

flaccid motor weakness with depression of tendon reflexes and elevation of protein levels in CSF without pleocytosis. The annual incidence of GBS is about 1 per 100,000 for adults and approximately the same for children."[169] The Institute of Medicine judged that "because [this] case by Pollard and Selby . . . demonstrates that tetanus toxoid *did* cause GBS, in the [IOM] committee's judgment tetanus toxoid *can* cause GBS."[170] It then added, "[t]he relation between tetanus toxoid and GBS is convincing at least for that one individual, even though this man [subsequent to his last episode of GBS caused by tetanus toxoid] experienced multiple recurrences of demyelinating polyneuropathy, most following acute viral illness . . . . [Two other cases] are recorded in enough detail to be accepted as GBS."[171]

The IOM regarded this single case report as sufficiently powerful on its own to show the likelihood of causation resulting from exposure to the tetanus toxoid (tetanus toxoid was capable of causing GBS). Moreover, the IOM judged the likelihood of *general causation* from the specific causation in this case.[172] Finally, they had little other data regarding adverse effects of tetanus toxid other than some background knowledge that foreign proteins introduced by vaccines can cause various adverse immune system reactions and the specific facts of the case. In particular, they had no epidemiological studies and no animal studies that reinforced this inference.[173]

The Article I Special Masters in the Vaccine Injury Compensation Program (VICP), who assess whether persons have suffered compensable injuries from vaccines, "have debated the utility of case reports" for causation inferences.[174]

---

[169] *Id.* at 86.

[170] *Id.* at 89 (emphasis in original).

[171] *Id.*

[172] *Id.* ("[B]ecause [this] case by Pollard and Selby (1978) demonstrates that tetanus toxoid did cause GBS, in the [IOM] committee's judgment tetanus toxoid *can* cause GBS.") (emphasis in original).

[173] Other case reports are discussed in CRANOR, *supra* note 45, at Chapter 4.

[174] Stevens v. Sec'y of Dept. of Health & Human Servs., No. 99-594V,

Some had initially opposed them, but, in the end, even a judge who was at first skeptical about 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 cased 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."[175]

The case report that the judge found persuasive is the tetanus toxoid case study cited by the Institute of Medicine in the preceding paragraph.[176]

Not all case reports will be such powerful evidence as the one mentioned or as strong as the case reports that assisted in the identification of the vinyl chloride monomer as a human carcinogen.[177] Some may be poor reports that should not be considered scientifically relevant. However, courts should recognize that many case reports can be scientifically relevant and can serve as part of the evidentiary basis of a good scientific argument.[178] They would need to assess each one for scientific relevance and not dismiss them because they are a *kind* of evidence that does not unerringly point to toxicity.

2001 WL 387418, at *14 (Fed. Cl. Mar. 30, 2001) (*quoting* O'Leary v. Sec'y of Dept. of Health & Human Servs., No. 90–1729V, 1997 WL 254217, at *3 (Apr. 4, 1997)).

[175] *Id.*

[176] *Id.* at *15 n. 31.

[177] See *supra* notes at 12-13 and accompanying text for further discussion

[178] Some of the evidentiary value of case reports are discussed in more detail in CRANOR, *supra* note 45, at Chapters 4 and 6.

C. Scientists Assess All the Relevant Scientific Evidence for a Conclusion

Central to nondeductive (scientific) arguments is a need to integrate all the relevant and available evidence to come to a conclusion about the most likely explanation of what is occurring. In assessing a substance's toxicity, scientists typically utilize human evidence, if it is available, experimental animal studies, if they are available, any chemical structure-biological activity evidence, and any mechanistic evidence in order to evaluate what conclusion follows.[179] A widely shared view in setting out and assessing scientific arguments is, "never throw [relevant] evidence away."[180]

Recall from above that the International Agency for Research on Cancer scientific committees explicitly go through a stepwise process to integrate human and animal evidence together with mechanistic and other evidence to assess whether a substance is a carcinogen.[181] Recommendations for integrating evidence are not unusual; quite the contrary, they are routine.[182] For a more theoretical example of how all relevant evidence must be considered, consider Susan Haack's description of Watson's and Crick's evidence for the double-helix structure of DNA:

> 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

---

[179] This is shown by many examples from the International Agency for Research on Cancer or similar examples from the national Toxicology Program. *See* CRANOR, *supra* note 45, at 302-19.

[180] Hutchinson & Lane, *supra* note 48, at 10.

[181] *See supra* note 59 for a further discussion concerning IARC.

[182] *See*, *e.g.*, Cogliano, et. al, *supra* note 59 for a description of the process at IARC. Examples from the National Toxicology Program indicate a similar integration of evidence occurs in their deliberations.

> evidence that DNA is a double-helical, backbone-out macro-molecule with like-with-unlike base pair. 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).[183]

This esoteric theoretical discovery, based on the integration of several different kinds of evidence, 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 assimilate in a *scientifically plausible* way *all* the *relevant evidence* to explain the claims in question. As noted above, eliminating each piece of evidence one by one would undermine this major scientific argument.[184]

Moreover, different patterns or kinds of evidence may play a greater or lesser role in supporting a toxicity judgment, depending upon what other evidence may be available in a particular case.[185] Sometimes one kind of evidence may be more important, sometimes another. Animal, in vitro, and various forms of mechanistic evidence, including structure-activity relationships, can be particularly important, depending upon the presence of other evidence.[186] Other kinds of evidence have become increasingly important for assessing the toxicity of substances when there is poor or no human evidence. This has led some distinguished cancer researchers to point out that "there should be no [hierarchy of state-of-the-art approaches for making toxicity decisions]. Epidemiology, animal, tissue culture

---

[183] *See* Haack, *supra* note 126, at 237 (2001).

[184] *See supra* discussion at notes 122-26.

[185] For examples see CRANOR, *supra* note 45, at 302-19.

[186] For some examples of this point see Cranor & Eastmond, *supra* note 124, at 36-41; CRANOR, *supra* note 45, at 302-19.

and molecular pathology should be seen as integrating evidences in the determination of human carcinogenicity."[187] The Institute of Medicine and the National Research Council echo this point.[188]

The mere fact that one piece of evidence does not strongly support a conclusion does not imply that all the evidence taken together fails to support it. In other words, a particular piece of evidence might fall short of ideal evidence of its kind, or even of ideal evidence for supporting the conclusion at issue, but the total evidence may still support the conclusion as more likely than not correct.[189]

IV. JUDICIAL REVIEW OF SCIENTIFIC ARGUMENTS

How should courts approach scientific arguments, given the framework discussed above? The nature of scientific arguments (inferences to the best explanation) and their features suggest the following procedure for reviewing expert testimony independent

---

[187] Michele Carbone, George Klein, Jack Gruber & May Wong, *Modern Criteria to Establish Human Cancer Etiology*, 64 CANCER RESEARCH 5518, 5522 (2004).

[188] DIETARY SUPPLEMENTS, *supra* note 143, at 254 ("It is also not appropriate to develop a hierarchical approach to considering the different types of data—human data, animal data, *in vitro* data, or information about related substances—for various reasons. In part, such an approach is not feasible because of limitations in the quality of the data and what different types of studies can reveal, but these limitations can be overcome with other types of data. Although a hierarchical approach is not practical, it is possible to weigh the various types of data available to make conclusions regarding risk to human health.").

[189] DIETARY SUPPLEMENTS, *supra* note 143, at 255 ("Individual pieces of information from any one of the categories of information (human, *in vitro*, animal, or related substances data) may sometimes be sufficiently compelling to both exceed a threshold of concern and to justify focused evaluation or action. In many circumstances, however, data will need to be collated within the same category or across several categories to determine the appropriate level of concern. That is, even if concern raised by one category of data—for example, human data—does not meet a threshold for action, the body of evidence available across several categories may raise the level of concern.").

of legal language with which courts will be more familiar.

Courts could inquire into the *scientific relevance* of individual pieces of evidence, but only if the evidence appears obviously irrelevant. Perhaps better, if courts are concerned about the scientific relevance of individual pieces of evidence, they might query how particular pieces could, however slightly, affect a particular scientist's conclusion. Such a *relevance review* would not ordinarily result in rejection of evidence, since relevance criteria are comparatively easy to satisfy. An intellectual presumption might be that evidence is scientifically relevant, unless a judge finds a scientist cannot explain the relevance of particular evidence.

Moreover, given the paucity of understanding of the toxicity of substances, courts need to recognize the obvious potential relevance of molecular evidence, animal studies, and good case reports. These are all kinds of studies with which some have struggled when they assumed individual pieces of evidence had to support by themselves, or perhaps provide major support for, an expert's conclusions.

Once courts are satisfied that no individual pieces of evidence are irrelevant, or they have excluded any that are obviously irrelevant, they could then review proposed testimony to determine whether the testimony or the argument, *based on all the scientifically relevant and integrated evidence* that the expert had used, is minimally legally plausible for the claims in question. They should ask: how plausible (or well supported) is the argument, given all the integrated evidence utilized by the expert? This is how one would think about nondeductive arguments outside the context of the law. Finally, in reviewing all the evidence considered together, courts might well find that experts can legitimately disagree about the weight or strength that each piece of evidence has for the overall argument. Courts must recognize the fact of scientific disagreement and allow for it in reviewing expert testimony.[190]

What do these recommendations mean in terms of legal guidance with which courts would be familiar? According to the

---

[190] CRANOR, *supra* note 45, at Chapter 7 (esp. 289-96).

Amended Federal Rule of Evidence (FRE) 702 the test for reviewing expert testimony is whether:

> (1) the testimony [which I take to be scientific argument] is based upon sufficient facts or data,
>
> (2) the testimony [scientific argument] 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.[191]

The logic of Rule 702 can be read as compatible with the ways in which scientists would assess other scientists' arguments. Consider only the first two clauses. A plausible reading of clause (1) appears to suggest that courts address the following question: considering all the scientifically relevant evidence a scientist has taken into account, is his or her testimony (which can be the expert's nondeductive argument for general or specific causation) supported by *sufficient evidence for the conclusion he or she draws* in order for a jury to consider it? That is, this section in order to make it compatible with assessing nondeductive arguments could be read to suggest that courts should review whether an expert's scientifically relevant evidence taken as a whole is sufficiently strong for the conclusion for which an expert argues and sufficiently strong to pass a reliability review to assist a jury. There is no suggestion that individual items of evidence be rejected piecemeal as insufficiently supportive of the overall scientific conclusion. Perhaps this could be one interpretation of the clause, but it would be incompatible with scientific approaches to argument evaluation.

Clause (2) requires that an expert's testimony must be based upon "reliable principles and methods."[192] An inference to the best explanation is a standard inferential methodology that scientists and many other experts utilize to infer causation and other empirical claims. The methodology is widely endorsed and

---

[191] Advisory Committee on Evidence Rules, "Proposed Amendment: Rule 702" (December 2000) (emphasis in original).

[192] *Id*.

accepted.[193] Particular scientific arguments, or uses of the method, are as "good" or "reliable" as an expert constructs them. How strong is the relation between all the relevant evidence in the premises and the conclusion drawn? In the terms of this article, the question would be, "Is the inference based on all the relevant evidence sufficiently plausible to assist a jury?" This suggests the following question for the courts: "Is a litigant's expert's argument, taking into account all the scientifically relevant evidence, sufficiently minimally plausible (reliable) for the expert to be admitted to testify?" Alternatively is the relationship between the premises and the conclusion so speculative, given all the scientifically relevant evidence that the expert should not be permitted to testify?

Courts are likely to see mixed patterns of evidence with some possibly good case reports, likely some quite good animal data, molecular data, but at least sometimes with no definitive epidemiological studies. Because of scientific ignorance about the universe of chemical substances, poor testing of products by firms[194] and perhaps reduced incentives for others to test them, toxicity data is likely to be limited,[195] forcing experts to utilize arguments based upon a wide variety of scientifically relevant evidence. Thus, judges will need to recognize that scientists often must draw on more complex evidentiary patterns out of necessity, recognize that they can be quite good, and be prepared to admit testimony based on them even when there is poor or no epidemiological evidence.

The Supreme Court in *Kumho Tire v. Carmichael* provided two valuable heuristics for reviewing scientific testimony for admissibility.[196] The first is that the court should ensure "that an

---

[193] *See supra* Section II.

[194] *See* Margaret A. Berger, *Eliminating General Causation: Notes Towards a New Theory of Justice and Toxic Torts*, 97 COLUM. L. REV. 2117, 2135 (1997) (*citing* studies of Agent Orange, asbestos, Bendectin, breast implants, the Dalkon Shield, thalidomide, tobacco, MER/29 (a cholesterol-reducing drug that caused cataracts), alachlor, atrazine, formaldehyde, and perchloroethylene); CRANOR, *supra* note 45, at 166-70, 350-53, 364.

[195] *See* discussion and accompanying text *supra* notes 1-3.

[196] Kumho Tire v. Carmichael, 526 U.S. 137, 152 (1999).

expert . . ., employs in the courtroom the same level of intellectual rigor that characterizes the practice of an expert in the relevant field."[197] This heuristic appears to be amplified on the following page where it is applied to the facts of the case. It added that a court may exclude evidence if it finds that an expert's testimony 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.'"[198] Both heuristics are something of sociological guidelines for judges: How does this expert's testimony on the issue in question compare to the standards of the profession (the first) or to other experts in the field (the second).

The second heuristic is especially attractive and some judges have found it quite practical. Judge Lee utilized it in a Parlodel case,[199] and Judge Pointer used it in the *Silicone Breast Implant Litigation*.[200] Given the nature of scientific arguments, this can be quite a helpful guide, as I have argued elsewhere.[201]

It follows from the second heuristic that if an expert's

---

[197] *Id.* at 152-53.

[198] *Id.* at 153.

[199] Order of the Court, April 6, 2001, in Soldo v. Sandoz Pharmaceuticals Corp., 244 F. Supp. 2d 434 (W.D. Pa. 2003). Judge Lee requested that if the appointed experts, disagreed with plaintiffs' experts that Parlodel can and did cause stroke, address whether "opinions [should] be considered subject to sufficient genuine dispute as would permit other persons, generally qualified in your field of expertise, to express opinions that, though contrary to yours, would likely be viewed by others in the field as presenting legitimate and responsible disagreement within your professions?" Court's Instructions to Experts appointed Pursuant to Federal Rule of Evidence 706.

[200] Judge Pointer utilized this heuristic to guide the deliberations of a Rule 706 National Academy of Sciences Panel, which was instructed to "review and critique the scientific literature pertaining to the possibility of a causal association between silicone breast implants and connective tissue diseases, related signs and symptoms, and immune system dysfunction." Judge Pointer was the coordinating judge for the federal breast implant multi-district litigation. The Panel published a report, dated November 17, 1998, which is entitled *Silicone Breast Implants in Relation to Connective Tissue Diseases and Immunologic Dysfunction*.

[201] CRANOR, *supra* note 45, at Chapter 7.

testimony is *within the range of opinion where experts might reasonably differ on a scientific issue before a court even though the evidence is shaky*, then he or she should be admitted to testify. Testimony that is within such a *zone of reasonable scientific disagreement* seems precisely the kind of testimony courts should permit juries to hear on disputed issues.[202]

In toxic tort cases, there are usually at least two, and possibly more conclusions that litigants may claim are supported by the totality of the evidence. Plaintiffs typically claim that exposure to defendants' substance caused or contributed to plaintiffs' injuries, whereas defendants will claim that *something else*, such as other exposures, bad luck, genetic predisposition, unknown antecedents, etc., caused plaintiffs' injuries. Thus, one large factual issue for a jury is the comparison of two or more explanations: Is the plaintiff's explanation more probable than the defendant's? A judge's task would seem to be to determine whether each side's scientific testimony is sufficiently plausible, or "within the range of opinion where experts might reasonable differ," given the relevant evidence on the issue, for a jury to hear the testimony.[203] The jury must then assess the weight and credibility of the scientific evidence together with other evidence to decide the case.[204]

CONCLUSION

A central virtue of the law is fairness between litigants at the bar. It should be especially committed to fairness in admissibility decisions because such decisions can be so critical and outcome determinative. With the change in principles guiding admissibility from *Frye* to *Daubert* courts can no longer rely on generic admissibility reviews that were more typical under *Frye*. Instead, they will need to review *individual*

---

[202] Of course such testimony should reliably apply to the facts of the case, as *Daubert* and the Amended Rule 702 require.

[203] Kumho Tire v. Carmichael, 526 U.S. 137, 152 (1999).

[204] A more extensive discussion of these issues is in CRANOR, *supra* note 45, at Chapter 7.

scientific arguments by experts and some of the individual evidence on which they rest for relevance. This is a difficult task, much more burdensome than under *Frye*, and something that challenges those distant from the science. If expert testimony is within the range of opinion where experts might reasonably differ on a scientific issue, the expert should be admitted to testify before a jury. If not, the expert should be excluded. Admitting testimony that is with a zone where comparable experts might reasonably differ would be part of fair admissibility reviews.

The upshot of this proposal is that any assessments of plausibility or reliability (or whether the expert's argument has the same intellectual rigor that characterizes the practice of other experts in the relevant field or is within a zone of reasonable disagreement) should be applied to *overall scientific arguments*, *not typically to individual pieces of evidence*. This is not likely to make courts' tasks easier; quite the contrary, such judgments will be more complicated than assessments some courts have made about individual pieces of evidence (but such reviews frequently were mistaken ways to evaluate scientific arguments, unless they focused merely on issues of relevance). Moreover, given how little is often known about potential toxicants, courts may face complex patterns of evidence that increase their challenges. Even though this is a much more daunting procedure to implement, it is the way scientists would review analogous arguments in their field. Fairness to an expert's testimony would seem to require that courts review it as scientists would. If courts are going to review particular scientific arguments (testimony) as *Daubert* mandated, they will need to face up to the challenges of the task or perhaps return to a less intrusive, more generic guide for reviewing expert testimony.[205]

---

[205] For an alternative to *Daubert* for reviewing expert testimony, see Donaldson v. Central Illinois Public Service, 199 Ill.2d 63 (2002). Other state courts have not followed the *Daubert* approach. *See supra* note 98. *See also* Raphael Metzger, *The Demise Of Daubert In State Courts,* Commentary for Lexis Nexis MEALEY'S Emerging Toxic Torts 14 (5) (June 3, 2005); available at http://www.mealeys.com.