A Daubert Motion: A Legal Strategy to Exclude Essential Scientific Evidence in Toxic Tort Litigation

In the US Supreme Court’s Daubert v Merrell Dow Pharmaceuticals, Inc decision, federal judges were directed to examine the scientific method underlying expert evidence and admit that which is scientifically reliable and relevant.

However, if a judge does not have adequate training or experience in dealing with scientific uncertainty, understand the full value or limit of currently used methodologies, or recognize hidden assumptions, misrepresentations of scientific data, or the strengths of scientific inferences, he or she may reach an incorrect decision on the reliability and relevance of evidence linking environmental factors to human disease.

This could lead to the unfair exclusion of valid scientific evidence, particularly that which is essential to a plaintiff’s case in toxic tort litigation. (Am J Public Health. 2005;95:S30–S34. doi: 10.2105/AJPH.2004.046250)

THE US SUPREME COURT ruling on Daubert v Merrell Dow Pharmaceuticals, Inc directed federal judges to act as gatekeepers by deciding whether to allow expert evidence to be presented to a jury. Judges are expected to examine the scientific method underlying expert evidence and to admit that which is both scientifically reliable and relevant to the issue at hand. The decision may seem well intentioned, because it could eliminate scientifically ungrounded opinions (e.g., all chemicals are carcinogens, or no environmental chemicals cause human cancer, or animal findings are not relevant to human risk). However, the issues surrounding environmental health effects are not always intuitively clear, because most scientific conclusions related to human health risks are based on interpretations of several sources of data, and absolute certainty may not be achieved for individual causality.

Thus, a judge who does not have expertise in dealing with scientific uncertainty, agree with a particular interpretation, understand the full value or limit of currently used methodologies, or recognize hidden assumptions, biases, or the strengths of scientific inferences, may reach an incorrect decision on the reliability and relevance of credible evidence linking environmental factors to human disease.3,3

The case of Daubert v Merrell Dow Pharmaceuticals, Inc concerned whether or not Bendectin can cause birth defects in humans. The district court maintained that expert testimony based on in vitro and live animal studies, pharmacological similarities between Bendectin and other substances known to cause birth defects, and unpublished reanalyses of negative epidemiological studies on Bendectin were inadmissible evidence of causality.

What is reliable and therefore admissible scientific evidence? According to the Daubert opinion it is the following: (1) evidence based on a testable theory or technique; (2) the theory or technique has been peer reviewed; (3) the technique has a known error rate; and (4) there is general acceptance of the underlying science. Because there are no clear guidelines on how to objectively determine scientific validity, judges may make decisions based on their own values and preconceived notions.4 The criteria for admissible evidence indicated in the Daubert decision can be met without achieving scientific validity and, conversely, validity may exist without meeting these criteria.5

The burden on the judge is considerable because failure to fully understand the scientific issues or to distinguish reliable from unreliable testimony could result in a decision whereby juries would not hear expert witnesses present relevant, reliable, and legitimate evidence. The decision not to admit expert testimony by judges, who Chief Justice Rehnquist labeled “amateur scientists,”6 could lead to the exclusion of scientific evidence essential to a plaintiff’s claims in toxic tort litigation. Thus, a Daubert motion provides a special opportunity for defendants to exclude incriminating evidence from a court proceeding.

Scientific data relevant to human health effects come in many different forms (e.g., clinical trials, epidemiological studies in humans, controlled studies in experimental animals, or laboratory studies in vitro systems), which have strengths and limitations. Understanding the relevance and reliability of the diverse experimental approaches and findings generally depends on how the study was designed, how the data were collected, analyzed, and evaluated, and the different perspectives put forward by experts in multiple disciplines (e.g., epidemiology, toxicology, pathology, medical cellular/molecular biology, chemistry, statistics, biomathematical modeling, etc.). It is unrealistic to expect a judge untrained in these areas to understand all of the underlying issues that might impact the validity and relevance of data from each of these disciplines with respect to determining human health risks.

Weak experimental designs and methods bias data used to render interpretations of human health risk toward not finding a risk even if a risk exists. For example, an insensitive analytical method that does not detect an
agent in the environment or an insensitive or inadequately designed study that does not achieve statistical significance for an adverse effect would not necessarily mean that the agent was not present in the environment or that risk does not exist. A laboratory study with animal group sizes of 20–50 usually cannot detect a significant risk of 10% or less; a study in 10 healthy adult male volunteers is not only insensitive but also provides little information about risks to children, the elderly, women, susceptible subgroups, or those exposed simultaneously to other toxic agents; an underpowered cancer epidemiological study cannot rule out the possibility of increased cancer risk in exposed populations. For these reasons and more, health agencies develop guidelines for judging the adequacy of experimental data for evaluating human health risks. For example, to reach the conclusion that human evidence suggests the lack of carcinogenicity, the World Health Organization’s International Agency for Research on Cancer requires data from multiple, mutually consistent, adequately powered epidemiological studies covering the full range of human exposures that exclude with reasonable certainty bias, confounding, and chance as well as provide individual and pooled estimates of risk near unity with narrow confidence intervals. In addition, the International Agency for Research on Cancer cautions that “latent periods substantially shorter than 30 years cannot provide evidence for lack of carcinogenicity.”

Unless a judge has had specific training in the multiple scientific and engineering fields relevant to the expert testimony being presented, it is unlikely that he or she alone would be able to recognize all the biases or hidden assumptions that could render evidence or counterarguments suspect or invalid. If the issues were truly that clear, there would not be disagreement and debate among health scientists concerning their opinions on the relative influence of complex interactions of environmental, genetic, medical, and lifestyle factors on the health of individuals and the public. Although not unexpected, it is generally scientists representing the interests of industry who overstate matters of controversy by downplaying the value of scientific evidence that was not obtained from human exposure studies and by exaggerating the possible role of confounders in epidemiological studies that show positive associations. In order to render a valid decision, it is essential that the judge be able to recognize exaggerations or misrepresentations of scientific controversy. By disallowing testimony of the plaintiff’s witnesses who offer opinions within the boundaries of normal scientific debate, the judge has essentially interjected his or her bias into complex environmental health issues that may not be resolved in the scientific community.

A judge who does not fully understand critical aspects of scientific methodology and interpretation of data may dismiss the evidence of expert witnesses who provide opinions based on methods well established and commonly accepted in the scientific and health communities. Whereas some judges may have claimed that results from animal studies cannot be extrapolated to humans, this opinion is contrary to the positions of all public health agencies, both national and international. For example, the preamble to all International Agency for Research on Cancer Monographs on the Evaluation of Carcinogenic Risks to Humans states that “in the absence of adequate data on humans, it is biologically plausible and prudent to regard agents and mixtures for which there is sufficient evidence of carcinogenicity in experimental animals as if they presented a carcinogenic risk to humans.”

The latter view is based on the fact that all known human carcinogens that have been adequately tested have produced significant carcinogenic effects in animal models. Rodents have been widely used as models for humans in toxicity and drug safety tests because of the trans-species similarities in physiological and biochemical processes. The scientific community considers animal data to be relevant and reliable for human risk assessments, even though most animal studies are performed at exposures higher than those to which humans are exposed in the environment. Concerning human exposures in the workplace, many bioassays have been conducted at similar exposure levels.

Low-dose extrapolations from experimental studies in animals to human exposure levels are necessary because the lower limit of detection of additional risk in animals, which is approximately 10% above the background rate, is an unacceptable level of risk for humans. Dose-response analyses of experimental data are critical for estimating human risk at environmental or occupational exposures. Without adequate familiarity with experimental designs, data analyses, and methods for evaluating human risk, a judge’s pretrial Daubert decision could well lead to a jury being denied hearing “reliable and relevant” evidence from knowledgeable experts; such a decision usurps the jury’s role of assessing the validity of scientific testimony and determining whether opinions are plausible.

Because the defendant would certainly hire scientists from multiple disciplines with biases favoring their positions, the cross-examination process and presentation of relevant contrary evidence or opinions by the defendant’s experts to the jury is a fairer way of revealing testimony that is reliable and credible. The dismissal of expert testimony prior to a trial based on the possible mistaken perceptions of a trial judge is inconsistent with our national principle of equal and impartial justice for all citizens. Consider the situation in which a Daubert motion is made and by which a judge decides that the plaintiff’s expert witnesses in a toxic tort case are relevant and reliable. Shouldn’t the judge automatically then determine that the opinions of the defendant’s expert witnesses (those who would claim that the defendant’s products or agents released into the environment are not harmful to human health) are also reliable and relevant? If the latter opinions are not found to be reliable, then presumably the judge would exclude their testimony as evidence to be presented before a jury. Such a decision is theoretically possible, but would likely be challenged for not allowing the defendants the means of obtaining witnesses in their favor. Likewise, not allowing the plaintiff the right to a trial by jury is unfair to the party in a toxic tort case that is
seeking compensation for its injuries. Why shouldn’t a jury be allowed to hear all of the relevant scientific evidence and opinions regarding adverse health effects that may result when a company pollutes the environment or workplaces of US citizens? Because the plaintiffs carry the burden of proof in toxic tort litigation, dismissal of expert testimony affects plaintiffs more than it does defendants. Thus, the application of Daubert in jury trials tips the scales of justice strongly in favor of defendants, who may have adversely affected the health of others through negligent or irresponsible emission or manufacture of harmful agents.

INCOMPLETE SCIENCE DOES NOT JUSTIFY EXCLUDING EVIDENCE

Because knowledge on environmental diseases is often incomplete, it is not uncommon for individual scientists to come to different conclusions when interpreting the same data sets and assessing their implications for human health. For example, the finding of hemoglobin adducts in humans or animals exposed to a particular agent indicates that exposure occurred and that that agent or one of its metabolites was reactive with proteins. If further study shows that DNA adducts were also formed, then the level of concern might be raised because DNA maintains the code for faithful replication of the cell in which that adduct was present. If the DNA adducts detected were similar to those of a known human carcinogen, then some might feel that we should be concerned about the potential carcinogenicity of that agent to humans. If we also found that that agent was metabolized by a similar pathway as a known human carcinogen, catalyzed by the same enzymes present in animals and humans, and that animal carcinogenicity studies showed similar types of tumors for both agents, then most but not all scientists would conclude that such data provides indisputable evidence of human cancer risk despite a lack of epidemiological results specific to that agent. This is the type of evidence that is available for vinyl fluoride and vinyl bromide in comparison to the known human carcinogen vinyl chloride. If a judge still maintains his or her bias and require epidemiological evidence of carcinogenicity in humans before allowing such compelling evidence to be presented before a jury?

National and international agencies that provide evaluations of human health risks do not rely solely on associations observed in epidemiological studies. Most often, no adequate studies have been performed, especially on newly introduced chemicals. Rather, expert multidisciplinary panels use all of the available and relevant scientific evidence in reaching their overall conclusions. Interestingly, regarding dioxin and ethylene oxide, the International Agency for Research on Cancer (IARC) and the National Toxicology Program (NTP) both concluded that the presence of these chemicals is known human carcinogens (i.e., sufficient evidence exists that there is a causal relationship between exposure to the agent and human cancer), although there was less than compelling evidence of carcinogenicity from studies in humans. Evaluating each piece of evidence separately, as might occur in a Daubert decision, could often lead to incorrect judgments of causality of human disease.

For most toxic agents, reliable epidemiological evidence is not available. Protection of public health is based on primary prevention and acting on warning signals from all relevant sources of information. By reducing or eliminating exposure to cancer suspect agents, we may thankfully never see enough cancer patients to confirm their carcinogenicity. The alternative of waiting for dead bodies to appear before taking any preventative action has been referred to as the “body in the morgue approach.”

Most scientific interpretations related to health risks are based on a variety of assumptions; some are explicit, whereas others are frequently based on the different ways in which individuals evaluate available evidence and consider alternative explanations. It is virtually impossible to state with absolute certainty that an individual’s disease condition was due solely to a specific exposure; likewise, it is impossible to state with absolute certainty that a past exposure to a particular toxic agent did not contribute in some way to that disease. If a judge requires nearly absolute certainty of causation, then he or she has raised the standard of proof for plaintiffs in such toxic tort cases to a nearly unachievable level. Recognizing the difficulty in drawing conclusions from epidemiological studies, Sir Bradford Hill developed a series of criteria for determining causality in cancer epidemiology. On the issue of making health-based decisions with incomplete evidence, Hill noted “all scientific work is incomplete—whether it is observational or experimental. All scientific work is liable to be upset or modified by advancing knowledge. That does not confer upon us a freedom to ignore the knowledge we already have, or to postpone the action that it appears to demand at a given time.”

Because of uncertainties or lack of complete information on disease processes and how intrinsic and extrinsic factors may be involved, it is not possible to estimate precisely the level of human health risk from experimental toxicity data. Although many industrial chemicals have been studied for toxic effects in animals, no toxicological information is available for the majority of chemicals. Also, although new mechanistic insights on disease processes are advancing daily, there is still much to be learned about how environmental factors, human variability (e.g., genetics, gender, age, exposure to other agents), and lifestyle factors (e.g., diet, exercise, alcohol consumption) interact to influence the likelihood of disease outcome. Exposure issues such as timing, duration, frequency, and intensity, as well as exposures to other agents and latency for clinical manifestation of the disease (e.g., cancer latency may be as long as 20–40 years) also impact on evaluations of disease-exposure relationships, yet precise information on these factors is not always available for exposed populations. Because of uncertainties, scientists may come to different conclusions on the relevance of specific findings to disease causation. With appropriate hypothesis testing, knowledge gained can reduce uncertainty. However, even with additional study, it is unlikely that we will know completely the mechanisms of...
disease causation by environmental agents and thus prove that a specific exposure was the sole cause of an individual’s diseased condition. The fact that uncertainty exists does not mean that valid evidence cannot establish realistic links between exposure and disease causality.

HEALTH-PROTECTIVE DECISIONS PREVENT NEEDLESS SUFFERING

Based on sound scientific evidence, it is possible to characterize the likelihood of human risk from exposure to specific environmental agents. This principle has been adapted by most national and international health agencies that assess the health effects of environmental agents. In the face of uncertainty, these agencies consider it prudent to act on the warning signals that arise from experimental studies and make decisions that are protective of public health. Although most rodent carcinogens have not been adequately evaluated in human studies, too often carcinogenic effects that were detected in animal studies were later confirmed in human studies. In some instances, such as that of diethy stilbestrol, animal warnings were ignored and, as a result, many people suffered the consequences of exposure to an agent that causes genital and reproductive abnormalities and cancer in humans. For 1,3-butanediol, the permissible occupational exposure limit promulgated by the Occupational Safety and Health Administration was lowered from 1000 ppm to 1 ppm, but not until more than 10 years after this chemical was shown to be a potent, multiple-organ carcinogen in laboratory animals at exposures considerably lower than the Occupational Safety and Health Administration standard. Subsequent to the publication of the original animal carcinogenicity data, epidemiological studies confirmed the carcinogenicity of butadiene in humans and follow-up studies in laboratory animals demonstrated carcinogenic effects at 6.25 ppm.

In some instances, judges have excluded epidemiological evidence that shows a statistically significant increase in risk when those studies did not demonstrate increased risks greater than a doubling (relative risk of 2.0) in exposed populations. The reasoning behind this legal threshold is that if the relative risk in an exposed population is greater than 2.0, then for an exposed individual, disease causality is more likely than not to have been due to that exposure, that is, it exceeds 50% for exposed individuals. However, this judgment fails to recognize that risk probabilities are underestimated for exposures that accelerate the time of disease occurrence, that is, the time until cancers are detected is reduced in the exposed population. Several additional flaws in such rulings are also noticeable. First, for agents that are prevalent in the environment, human exposure may occur at multiple locations or sources (e.g., environmental tobacco smoke, drinking water disinfection byproducts, benzene); consequently, there is no truly unexposed reference population. Thus, the contribution of that agent to the disease rate in the reference population will result in an underestimation of relative risk in the exposed population. Second, if the relative risk estimates were obtained from occupational exposure studies, then the “healthy worker effect” may lead to underestimates of risk when disease rates are compared to the general (less healthy) population. Because of the healthy worker effect, risk for all causes of death is less than 1.0. Consequently, meaningful estimates of relative risk in workers compared to the general population need to be adjusted upwards, or relative risk estimates must be based on incidence in exposed workers versus unexposed workers; but, even then, “unexposed” workers often means “less exposed.”

Third, because risks are not uniformly distributed in exposed populations, a risk much greater than 2 may exist for various susceptible subgroups, even though the overall risk is 2.0 or less. For example, interindividual variability in the probability of disease causation may occur because of differences in the magnitude, frequency, and duration of exposure; genetic differences that influence how individuals metabolize the agent, produce or eliminate reactive metabolites, repair genetic damage, or predispose an individual to a disease; exposure to other agents (e.g., pharmaceuticals or occupational or consumer chemicals) that affect the behavior of the agent of concern in individuals; differences in health status (e.g., pre-existing disease, immune-system deficiency); and other age- and gender-related differences. Because of the complex nature and multiple interactions among risk factors, an individual’s risk cannot be estimated from epidemiological data alone. Focusing on a value of 2.0 as a measure of the relative risk in an exposed population rather than analyzing all of the data that contribute to the risk estimate is an irrational and inappropriate way to judge causality in an exposed individual. Requiring a relative risk greater than 2.0 is not a valid reason for dismissing pertinent evidence relevant to an individual’s claim of lost years of healthy life.

THE DATA QUALITY ACT SUPPRESSES SCIENTIFIC EVIDENCE

Similar to the Daubert decision, the Data Quality Act of 2000 provides another means for special interest groups to challenge the value of scientific information used by federal agencies for making regulatory decisions. For example, peer-reviewed studies by Hayes et al. were published in highly respected scientific journals showed endocrine-disrupting effects of the herbicide atrazine in frogs. These studies were challenged by the Center for Regulatory Effectiveness, which claimed that the US Environmental Protection Agency had not yet validated test protocols for demonstrating endocrine disruption. However, endocrine disruption by environmental agents has been studied and reported for over 25 years, and the following definition of an endocrine disruptor has been established: an “endocrine disruptor is an exogenous agent (synthetic or natural) that interferes with the production, release, transport, metabolism, binding action, or elimination of natural hormones in the body responsible for maintaining homeostasis and regulation of developmental processes.” Thus, the published findings that atrazine produced sexual deformities, including hermaphroditism in male frogs, as well as other studies showing delayed puberty and direct inhibition of...
Leidyg cell testosterone production in male rats clearly demonstrate the endocrine-disrupting effects of this agent. Furthermore, because hormones and hormone receptor systems are phylogenetically similar, the effects observed in one mammalian species raise concern about the potential effects in other mammalian species, including humans. It is interesting to note that the European Union has withdrawn approval of the use of atrazine because of health and environmental concerns. Challenges, such as the one by the Center for Regulatory Effectiveness under the Data Quality Act are simply attempts to exclude or delay the use of reliable scientific evidence for regulatory decisions in the United States. Because the Data Quality Act applies to research conducted by federal scientists and federal grantees but not to industry-sponsored research, an inherent bias exists for claims made under this act.

CONCLUSIONS

Evaluations of the health effects of environmental agents require thorough examination of all available and relevant scientific information by experts trained in the multiple scientific disciplines applicable to the issue. The dismissal of reliable evidence under a Daubert motion or through challenges made under the Data Quality Act results in unreasonable barriers for juries and regulatory agencies, respectively, to make appropriate decisions on the health effects of toxic agents in exposed individuals or populations. The Daubert decision and the Data Quality Act need to be reviewed for their biased impact on health-based decisions in the United States. ■

About the Author
The author is with the Environmental Toxicology Program, National Institute of Environmental Health Sciences, National Institutes of Health, Bethesda, Md.

References