

# Ten Years of Judicial Gatekeeping Under *Daubert*

Joe S. Cecil, PhD, JD

In the ten years since *Daubert v Merrell Dow Pharmaceuticals, Inc*, the standards for admissibility at trial of expert testimony in general and scientific evidence in particular have become more demanding. Reviews of recent cases and empirical studies of federal judges' and attorneys' practices indicate that judges are more likely to consider the admissibility of expert evidence prior to trial, to inquire more deeply into the reasoning and methodology that supports the expert opinions, and to limit or exclude such evidence from presentation at trial. Studies of published cases confirm this finding.

Recent cases consider more difficult questions arising from the differing methodologies used in various areas of science. The current legal framework that assesses admissibility in terms of professional practice outside the courtroom is poorly suited to cases that require expertise across a wide range of specialties and force judges to choose from among different scientific methodologies. Future research should focus on the pretrial screening of expert testimony and interactions between the attorneys and experts in shaping that testimony. (*Am J Public Health*. 2005;95:S74–S80. doi:10.2105/AJPH.2004.044776)

Many federal judges were uncertain how the 1993 Supreme Court decision in *Daubert v Merrell Dow Pharmaceuticals, Inc*,<sup>1</sup> would affect their work. But Judge Alex Kozinski, author of the Ninth Circuit appellate court decision that was vacated and remanded for further consideration by the Supreme Court, was worried. In reconsidering the case in light of the standards expressed in *Daubert*, Judge Kozinski wryly noted the following<sup>2</sup>:

Our responsibility, then, unless we badly misread the Supreme Court's opinion, is to resolve disputes among respected, well-credentialed scientists about matters squarely within their expertise, in areas where there is no scientific consensus as to what is and what is not "good science," and occasionally to reject such expert testimony because it was not "derived by the scientific method." Mindful of our position in the hierarchy of the federal judiciary, we take a deep breath and proceed with this heady task.

So how have the courts done? Law professors and other scholars have filled the law library shelves with articles analyzing published cases following *Daubert*. But published cases represent only a portion of the litigation; many more cases are resolved without a published order or decision.<sup>3</sup> Are these published cases indicative of broad shifts in the orientation of the courts that extend throughout litigation? Or are the published cases like

the tips of the trees that sway in the winds while the forest floor remains unstirred?

I review here the limited amount of empirical research that has assessed the effect of the *Daubert* decision on civil litigation as well as an emerging series of research studies intended to strengthen our understanding of the role of expert testimony in courts. I also examine a recent case involving conflicting evidence from court-appointed scientists, indicating the difficulties that arise in applying a standard based on professional practice in cases that require multiple areas of expertise.

## THE LANDSCAPE OF EXPERT TESTIMONY

Discussions of admissibility of expert testimony typically focus on some of the most demanding areas of science. Although judges often consider evidence similar to that used in assessment of public health issues, judges must consider such evidence in the context of deciding an individual case rather than the establishment of a broad social policy. The *Daubert* case involved conflicting testimony regarding epidemiology, toxicology, and pharmacology, characterized by Judge Kozinski as concerning "matters at the very cutting edge of scientific research, where fact meets theory

and certainty dissolves into probability."<sup>2</sup> Legal doctrines developed to guide judges in considering such difficult issues of science also are likely to be complex. However, it is worth noting as a preliminary matter that much of the testimony offered by testifying experts does not rise to this level of complexity. Such expert testimony requires special expertise, but much of this specialized testimony appears rather routine when measured against the evidence presented in *Daubert* or other cases involving the effects of exposures to allegedly toxic substances.

In one of the earliest empirical studies of expert testimony, Gross analyzed reports from a California jury verdict reporting service to determine the type and frequency of experts testifying in state civil trials that ended in a jury verdict.<sup>4</sup> Drawing on this data set, Gross and Syverud reported that more than half of the testifying experts were physicians, with only 3% identified as scientists.<sup>5</sup> Champagne, Schuman, and Whitaker surveyed attorneys, judges, jurors, and experts in civil cases in Texas and confirmed that about half of the experts were physicians, with scattered representation of other professions.<sup>6</sup>

More recently, my colleagues and I at the Federal Judicial Center surveyed federal judges regarding expert testimony in their most recent civil trial.<sup>7</sup> We then surveyed attorneys who served those same trials. In both the judge and attorney surveys, we asked respondents to describe the types of experts who testified in terms of the occupational categories used by Gross. Again, medical and mental health specialists were the most frequently presented category of experts, accounting for more than 40% of the experts presented overall. More specifically, medical doctors (from all specialties) accounted for approximately one third of all experts presented. Specialists from business, law, and finance accounted for 25% of the experts presented, making them the second most frequent category of experts. Economists were by far the most common type of expert

within this category, as well as the most common type of expert overall, representing almost 12% of all experts. (Economists were included in the financial category because they most often present estimates of lost profits or wages and other financial projections rather than apply broad economic theories to the facts of a case.) Engineers and other safety or process specialists registered close behind the business/law/financial sector, accounting for about 22% of all experts. Such experts respond to a wide variety of issues that are difficult to specify.

Scientists composed a small portion of the testifying experts. Specialists from scientific fields such as chemistry, ballistics, toxicology, and metallurgy accounted for only 8% of the experts that testified at trials. In this group, chemists were the most frequent type, representing 1.6% of the total number of experts. Epidemiologists and toxicologists, who were the focus of the *Daubert* decision, were rare, together constituting just over 1% of the testifying experts. In comparing these results with a similar 1991 survey, we concluded that the distribution of types of expert testimony in civil trial has changed very little following the *Daubert* decision.

Judges reported that the most frequent issues addressed by experts at trial were the existence, nature, or extent of injury or damage (68%) and the cause of injury or damage (64%), which finding is consistent with the fact that tort cases represented almost half of all cases reported. Testimony as to the amount of recovery to which plaintiff was entitled was offered by experts in 44% of trials; this type of testimony is consistent with the large number of economists who were reported as having testified. Other issues addressed by expert testimony were the reasonableness of a party's actions (in 34% of trials), industry standards/"state of the art" (30%), standard of care owed by a professional (25%), design or testing of a product (25%), and knowledge or intent of a party (16%).

## DAUBERT AS A GATEKEEPING DEVICE

At the time of the *Daubert* decision, there was considerable uncertainty about its effect on admissibility of scientific evidence. It soon

became clear that the interpretation of the decision was resulting in a more restrictive approach to admissibility of scientific testimony. A recent analysis by the Rand Corporation of a sample of 399 published and unpublished federal district court decisions appearing in the Westlaw database over a 20-year period indicates the extent to which courts have shifted toward excluding proffered scientific and technical evidence.<sup>8</sup> That analysis indicates that challenges to case "elements" involving expert evidence rose in the 3 years following the *Daubert* decision.

This interpretation is consistent with much of the commentary that followed *Daubert*. Many cases decided soon after *Daubert* involved clear instances of evidence that was not well supported by acceptable scientific methodology and permitted easy application of the *Daubert* standards. For example, the courts rejected the use of an industrial rather than medical methodology in assessing the presence of asbestos fibers in lung tissue,<sup>9</sup> the substitution of visual inspection for appropriate medical testing to conclude that cataracts were caused by radiation exposure,<sup>10</sup> and subjective opinions that failed to consider other plausible explanations for harmful effects.<sup>11</sup> Exclusion of expert testimony in such cases was generally straightforward and increased the confidence of judges in reviewing the basis of scientific testimony.

The Federal Judicial Center survey of federal judges and attorneys also confirmed a shift toward more demanding standards for admissibility of evidence. In 1998, both judges and attorneys indicated that judges were more likely to scrutinize expert testimony before trial and to limit or exclude proffered testimony compared with pre-*Daubert* litigation practice in 1991.<sup>7</sup> The survey revealed that motions filed early in litigation have become a favored pretrial device for challenging the admissibility of expert testimony and that judges are focusing more attention in pretrial proceedings on admissibility issues.<sup>7</sup>

The view from the state courts is less clear. A 1998 survey of state court judges by Gatowski et al. found judges split in their assessment of the effect of the *Daubert* decision.<sup>12</sup> Only about half of the judges from states that follow the Federal Rules of Evi-

dence (and, therefore, are likely to follow *Daubert*) felt that their gatekeeping role had changed as a result of *Daubert*, and only a third of the judges believed the intent of *Daubert* was to raise the threshold for admissibility. The study also found that state court judges did not have a good understanding of the meaning of the scientific criteria suggested by *Daubert* and raised questions about the ability of the state trial courts to apply such standards in a reasoned and thoughtful manner. Few judges, for example, were able to define the concept of "falsifiability," which was one of the factors mentioned in *Daubert*. Other studies have demonstrated that judges may be insensitive to methodological problems in social science research such as experimenter bias or the absence of a control group.<sup>13</sup>

The Rand study also identifies a notable change in the pattern of federal litigation. Beginning in 1997, the proportion of cases that were challenged for unreliable evidence declined, as did the success rate for such challenges.<sup>8</sup> The authors interpret this as evidence that parties responded to the post-*Daubert* standards by strengthening the quality of the expert testimony or abandoning those cases that were unlikely to meet the higher standards. Changes in publication practice once the case law interpreting *Daubert* became established may also have contributed to the appearance of such a change.

As the federal courts evolved to consider more difficult issues of expert testimony, the distinction between methodology and conclusions became blurred, and application of the *Daubert* standards became more difficult. The Supreme Court decision in *General Electric Co v Joiner*<sup>14</sup> illustrates the increasing difficulty in considering issues of scientific evidence during this period. In such cases, often the issue is not the absence of scientific evidence, but rather whether the existing scientific evidence can be generalized to address the specific causal relationships alleged in the case. In *Joiner*, for example, the plaintiff contended that exposure to polychlorinated biphenyls had promoted the development of his small-cell lung cancer. After establishing an abuse-of-discretion standard for appellate review, the Supreme Court then examined the record in the case and demonstrated the limited

extent to which courts should permit experts to generalize beyond established scientific findings.

The plaintiff in *Joiner* had presented a series of epidemiology studies that offered equivocal findings when taken separately, but according to the plaintiff, demonstrated a causal relationship when considered together. The plaintiff also presented studies with infant mice showing that injections of large doses of polychlorinated biphenyls led to cancer. The Supreme Court rejected such evidence after focusing on the breadth of generalization implied by the plaintiff's expert testimony and noting the following<sup>14</sup>:

Trained experts commonly extrapolate from existing data. But nothing in either *Daubert* or the Federal Rules of Evidence requires a district court to admit opinion evidence which is connected to existing data only by the *ipse dixit* of the expert. A court may conclude that there is simply too great an analytical gap between the data and the opinion proffered.

Many cases since 1997 have focused on this analytical gap between research data and expert opinion regarding the issues presented in specific litigation. Because such cases often examine opinions based in part on research methodologies that are employed in the normal course of scientific inquiry, they cannot be dismissed out of hand in the manner of those early cases decided soon after *Daubert*. Instead of dismissing the methodology used by the experts, judges who exclude such testimony typically focus on the reasoning process of the experts, questioning whether the methods and findings relied on by the experts can reasonably be extended to the facts of the case at hand. As part of such an analysis, recent courts have dismissed experts' reliance on animal studies<sup>15</sup> cautioned against extrapolation of dosage levels<sup>16</sup> and objected to generalization across similar substances.<sup>17</sup>

In questioning the propriety of generalizing research findings to the facts of specific litigation, courts often imply that the experts have abandoned an objective and impartial role appropriate for scientists and become advocates for the parties.<sup>18</sup> Judges have indicated persistent concern regarding the possible lack of objectivity on the part of testifying experts. The Federal Judicial Center survey states that both before and after

*Daubert* (i.e., in 1991 and 1998), the most frequent problem that federal judges encounter is "experts who abandon objectivity and become advocates for the side that hired them." This perceived lack of objectivity may have a number of different sources. Experts are selected by the parties based on the extent to which their testimony will advance the parties' claims, a practice that may favor the selection of extreme viewpoints.<sup>19</sup> Moreover, preparing an expert witness to offer testimony involves a socialization process that is likely to encourage the expert to identify with the interest of the party.<sup>4</sup> It is reasonable that judges, who likely were exposed to such practices prior to their arrival on the bench, would be skeptical of testimony offered by expert witnesses who had undergone such selection and coaching.

The gatekeeping role has evolved beyond a device for reviewing only scientific evidence to include all types of expert testimony. In *Kumho Tire Co v Carmichael*,<sup>20</sup> which involved an allegedly defective automobile tire, the Supreme Court extended the gatekeeping approach to all types of expert testimony. Prior to *Kumho*, the courts were divided on whether expert testimony based on experience and clinical medical testimony in particular should be subject to the *Daubert* screening process. In extending the trial court's gatekeeping obligation to all expert testimony, the Supreme Court noted that "no clear line" can be drawn between the different kinds of knowledge, and "no one denies that an expert might draw a conclusion from a set of observations based on extensive and specialized experience." Although the specific factors mentioned in *Daubert* may not be relevant to nonscientific expert testimony, other factors may provide a suitable standard for assessing such testimony. The Supreme Court indicated that all expert witnesses should employ "in the courtroom the same level of intellectual rigor that characterizes the practice of an expert in the relevant field." In effect, this decision tethered the standard for admissibility of expert testimony to standards of professional practice. This reliability requirement has also been added as a recent amendment to Federal Rule of Evidence 702, strengthening the role of the court in assessing the foundation of all expert testimony proffered for litigation.

## THE REVEALING CASE OF *SOLDO v SANDOZ PHARMACEUTICALS CORP*

Under these *Daubert/Joiner/Kumho* standards, expert testimony such as clinical medical testimony that employs a mix of traditional scientific methodologies and less rigorous case study and observational methodologies is especially problematic. In effect, the *Kumho* decision tethered the standard for admissibility of testimony by physicians to the professional standards of medical practice. But many courts have been reluctant to accept opinions based on these less rigorous methodologies, even though they are widely used in clinical medicine.<sup>21</sup> More is going on here than a straightforward distrust of witnesses. Even when the experts are clearly objective, that is, appointed by the court with no previous association or contact with the parties, a judge may question the extent to which the appointed expert's testimony can properly inform consideration of the issues in the case.

The recent case of *Soldo v Sandoz Pharmaceuticals Corp*<sup>22</sup> is revealing of the manner in which courts take account of different areas of expert testimony. Although this is a single case rather than a research study, it incorporates many aspects that would be desirable in a research study. Reputable scientists from different disciplines with no ties to the parties examined complex medical testimony and prepared independent reports regarding the scientific validity of the proffered testimony, which was then reconciled by the judge in resolving a motion to exclude the testimony. Viewed in this light, it might be properly considered a "case study" incorporating critical elements that no formal research study has yet achieved and is therefore appropriate to include as part of this review.

In *Soldo*, a young mother sustained an intracranial hemorrhage and resulting stroke soon after giving birth. She claimed that the stroke was a result of her ingestion of Parlodel, a drug manufactured and marketed by the defendant to prevent lactation. Because a heightened risk of stroke occurs in the postpartum phase of pregnancy, among the tasks facing the court was to determine if the risk of stroke among women who have taken Parlodel following pregnancy exceeds the heightened base rate of stroke among

women in the postpartum phase of pregnancy. No meaningful clinical trials or epidemiology studies of Parlodel exist, so the plaintiff's experts considered a series of animal studies, case studies, and other clinical reports in concluding that Parlodel caused the injury.

In assessing the scientific reliability of the testimony by the plaintiff's experts, the court sought the assistance of three court-appointed experts—a clinical pharmacologist, a neurologist, and an epidemiologist. The three court-appointed experts were identified with the assistance of the Registry of Independent Scientific and Technical Advisors, administered by the Private Adjudication Center of Duke University School of Law (<http://www.law.duke.edu/pac/registry/index.html>). Unfortunately, this service has been suspended. A similar service for identification of court appointed experts for federal litigation is available through the American Association for the Advancement of Science demonstration project, Court Appointed Science Expert (<http://www.aaas.org/spp/case/case.htm>).

All were well-regarded scholars with academic appointments at distinguished universities and medical schools. The experts had no previous association with any of the parties and had not previously considered the role of Parlodel in causing stroke.

In a lengthy order, the court sought the advice of the appointed experts on “whether the methodology or technique employed by plaintiff's medical witnesses . . . in formulating their opinions is scientifically reliable and whether the methodology or technique can be properly applied to the facts of this case.” The court recited an expanded version of the *Daubert* factors and noted that published studies need not be required as a basis for testimony, that “differential diagnosis and temporal analysis, properly performed, would generally meet the factors,” and that there must be a proper “fit” between the expert's opinion and the facts in the case. The court further instructed that if any of the appointed experts should find that the testimony of the party's expert is not “scientifically reliable,” then the appointed expert should also indicate if opinion by the party's expert might represent a “legitimate and reasonable” minority view within the profession.

### The Expert Reports

The three court-appointed experts, working independently, developed three different interpretations of the scientific reliability of the testimony by the plaintiff's experts, thereby demonstrating the diverse viewpoints employed by various scientific disciplines in assessing a similar body of evidence. All of the appointed experts acknowledged that the plaintiff's experts had made the most of the data that exist, but they differed in the extent to which they regarded such an effect as meeting the standards of “scientific knowledge.” The epidemiologist and the neurologist found that the testimony did not meet the standards of scientific knowledge, but for very different reasons. The clinical pharmacologist found that the testimony, with one exception, met the standards of scientific medical testimony. None of the experts used the same methodology or reasoning process.

The most demanding standard for finding information to be sufficiently reliable to be deemed “scientific knowledge” was offered by the court-appointed epidemiologist. He focused on the “analytical gap” in applying the research to the facts of the case, noting the following:

[B]ecause the information is so indirectly applicable and hypothetical in nature, the application of it to form an opinion is not a “scientifically reliable” process. The linkage between those shreds of potentially relevant information and the opinion that results is so murky that it is very difficult to see how the evidence leads to the opinions that are offered. Applying any reasonable standards of scientific evidence as the basis for drawing a conclusion leads to the judgment that we do not know enough to offer an opinion on this matter that is reasonably well grounded in science.

The court-appointed epidemiologist was especially critical of the absence of human studies and was unwilling to extrapolate findings across species, noting that, “Some form of epidemiologic or clinical evidence, even if flawed and incomplete, is needed for drawing inference about general causation in making a judgment about Parlodel and intracerebral hemorrhage.” He acknowledged that causal attributions may be made without such studies if the causal pathway is clear, such as injuries that arise when a tornado hits a mobile home park, but continued, “In the absence of clinical or epide-

miologic research, it would require a tremendous amount of indirect evidence to reach the point that even in the absence of research, the linkage is ‘obvious’ in the way that the tornado leading to injury is obvious.”

Regarding the possibility of legitimate and reasonable disagreement with his views, the epidemiologist acknowledged that views vary across sciences:

[T]he vast majority of scientists who routinely consider these sorts of evidence (epidemiologists, researchers in clinical medicine) would agree with my general conclusions. If forced to guess the proportion, I would estimate 80% of my peers would concur. Those who study basic mechanisms of disease causation (physiologists, pharmacologists, toxicologists) might well dispute my views in that the plausibility based on those lines of evidence is more supportive of the potential for causality. The counterargument to my view is that the diverse threads based on mechanism of actions for Parlodel, analogy to other agents in the same broad category of drug, and temporal linkage of the medication and the illness can be integrated scientifically into a scientifically reliable conclusion. However, essentially all scientists recognize that when the issue is the causation of clinical disease in humans, there is a sizable gap between what is plausible based on indirect evidence and what is proven based on clinical and epidemiologic studies. In the chain of reasoning, most scientists would likely share the view that leaping across the huge gulf of critical data moved the person making the inference beyond the scope of science.

The least demanding standard for assessing a causal relationship was offered by the clinical pharmacologist. He acknowledged that epidemiologic and clinical trial data would be helpful in assessing the causal relationship, but found them to be unnecessary:

[M]any important and well-recognized adverse drug reactions are not documented by well conducted epidemiologic studies that have been specifically conducted to detect them and that we still believe that the weight of the scientific evidence is sufficient to implicate their involvement to such an extent that we would remove a drug from the market. . . . [M]ost clinical practice is not guided by data from [prospective, randomized, placebo-controlled clinical trials], because they are so difficult and expensive to conduct. In addition, aggregate data from such trials often do not apply to the specifics of individual cases. To assert that any medical practice has no scientific basis because a randomized, placebo-controlled trial to answer the pertinent question was not conducted, would be to obviate the vast majority of clinical practice. It follows that other tools must

usually be used to provide sufficient evidence to guide our practice.

The court-appointed clinical pharmacologist found the plaintiff's experts' assertion of a causal link to be scientifically reliable (with one minor exception). In reaching this conclusion, the clinical pharmacologist used a "totality of the evidence" approach, bringing together assessments of animal studies, case studies, and other available evidence, "none of which taken separately may be determinant, but which when viewed as a whole may be considered convincing." Although such an approach is common in public health assessments, judges tend to consider separately the validity of each piece of evidence. Although the epidemiologist dismissed case studies as inherently unreliable, the clinical pharmacologist relied on a series of case studies, including one case study reported in a peer-reviewed journal that recorded vasoconstriction changes in carotid arteries with the introduction and removal of a drug closely related to Parlodel.

In finding a causal link, the clinical pharmacologist reasoned that evidence of vasoconstriction in peripheral arteries of dogs and humans caused by Parlodel and related substances would permit the inference that Parlodel would cause similar vasoconstriction in cerebral arteries, which may then lead to intracranial hemorrhage and stroke. Such reasoning far exceeds the modest stretch across the "analytical gap" that would have been allowed by the court-appointed epidemiologist.

The third court-appointed expert, a neurologist, was willing to consider animal studies as appropriate evidence, but he required more direct evidence of a vasoconstrictive effect of Parlodel on cerebral arteries. When the neurologist applied this standard to the proffered testimony of the plaintiff's experts, he too found it lacked necessary scientific rigor. The neurologist acknowledged that clinical or epidemiologic data are not necessary to declare a causal relationship, but he objected to reasoning by analogy that evidence of vasoconstriction in peripheral and carotid arteries is indicative of a similar relationship in cranial arteries. He pointed out that cerebral arteries respond differently to drugs like Parlodel than do peripheral arter-

ies. In the absence of specific human or animal evidence that Parlodel causes vasoconstriction in cerebral arteries, he found the assertion of such an association to be without scientific foundation.

Having found that the testimony by the plaintiff's experts was not scientifically reliable, the court-appointed neurologist then assessed the extent to which others in his field might express "a legitimate and responsible" contrary opinion. He acknowledged that others outside his specialty of experimental pharmacology of the cerebral vascular system might disagree with his conclusion because they are unfamiliar with the published scientific evidence regarding the difference in response of cerebral vasculature and peripheral vasculature to such drugs; he also stated that such knowledge is not typically part of the training of specialists in the broader fields of neurology and cerebrovascular disease. Moreover, he noted that contrary views may be found in medical textbooks and peer-reviewed articles and, in the absence of a clear cause for the hemorrhage, some measure of subjective judgment is required in assessing the evidence. In conclusion, he noted that other persons generally qualified in this field of expertise might legitimately disagree with his conclusions.

### The Court's Decision

The court ended up with three different opinions from three distinguished scholars from three different disciplines. One can imagine the disappointment, perhaps even regret, felt by the judge when he reviewed these three reports that shared little in common. It is likely that in calling for independent assessments from three different specialties, he had hoped that the opinions would converge, thereby strengthening assessment of the scientific methodology underlying the experts' opinions. Instead, he found a dramatic illustration of the variation in acceptable methodologies across well-established areas of science.

At first glance, one might assume that three different opinions from highly qualified court-appointed experts would in itself be evidence of "a legitimate and responsible disagreement" regarding the disputed issues and require that a jury resolve the conflict. Al-

though most courts have excluded such testimony regarding Parlodel,<sup>23,24</sup> at least one other federal court decided that such differences should be resolved by a jury.<sup>25</sup>

The *Soldo* court took no comfort in the diversity of views presented by the experts. Faced with such conflicting opinions, the court attempted to reconcile disputes over scientific validity that the scientific community itself had not resolved. In so doing, the court unavoidably stepped beyond the bounds of *Kumho* in assessing whether the methodology and techniques met "the same level of intellectual rigor that characterizes the practice of an expert in the relevant field" and moved on to establish legal policy regarding how such disputes among scientists were to be reconciled under federal law.

In an opinion that extends more than 100 published pages, the court dismisses the views of the court-appointed clinical pharmacologist and the plaintiff's experts and concludes that the opinions expressed by the plaintiff's experts "failed to use a reliable scientific methodology" to demonstrate general causation and specific causation. The court does not explicitly address the use by the court-appointed clinical pharmacologist of the "totality of the evidence" test, but it notes that while expert opinions may make "appropriate use of all of the available information, . . . in the absence of some minimum amount or level of scientific evidence, the opinions cannot be scientifically derived because there is too little science from which to derive them. Although it is sometimes necessary in a clinical, regulatory, or business practice to make decisions based on less than sufficient and/or reliable scientific evidence due to practical demands that require immediate decision-making, such guesses, although perhaps reasonable hypotheses based on the best available evidence, do not constitute a scientifically reliable approach when used to assess causality via the scientific method."

In conclusion, the court agreed with the court-appointed epidemiologist and neurologist:

The body of scientific evidence relating to Parlodel and stroke is simply insufficient to support a scientifically reliable application of plaintiff's expert methodology. . . . Without sufficient evidence of general causation, plaintiff's experts could not reliably apply a differential diagnosis that comports with the scientific method, notwithstanding the fact

that physicians in clinical practice may be required to proceed with a differential diagnosis on the basis of guesses or hypotheses due to the exigency of the need to treat their patients.

Having excluded the testimony of the plaintiff's experts, the court then granted summary judgment in favor of the defendant.

By finding that the report by the court-appointed clinical pharmacologist did not meet a sufficient standard of scientific reliability, the court made clear what many previous decisions have obscured; that is, conclusion of a casual relationship by a distinguished scholar unaffiliated with the parties and using generally accepted methods of clinical inference may not be sufficient to allow a court to submit such differences of opinion to a jury. Such a view would appear to conflict with the assurance of the Supreme Court in *Kumho* that expert opinions are admissible if they employ "in the courtroom the same level of intellectual rigor that characterizes the practice of an expert in the relevant field." For this court and an undetermined number of others, admissible evidence requires more than meeting appropriate standards of professional practice; it also requires demonstration of a relationship through methodologies that are not an essential part of clinical practice.

The court in *Soldo* indicated as much when it noted that even if the plaintiff's expert opinions were admissible under *Daubert*, "such evidence provides but a scintilla of support for plaintiff's position and would not be sufficient to allow a reasonable jury to find that plaintiff's [intracerebral hemorrhage] had been caused by Parlodel." One might argue that this is a more appropriate basis for making such a decision rather than for striking the testimony as inadmissible because of some perceived flaw in methodology and reasoning. Even if the experts use methods and reasoning appropriate to their profession, the courts may set, as a matter of law, a minimum threshold for evidence that is sufficient to justify submission to a jury. Of course, if such a legal standard is established in an explicit manner as being insufficient as a matter of law, then that decision would be subject to appellate review on a de novo basis and more vulnerable to reversal on appeal. Specifying such a legal standard, however, would seem preferable to declaring that

broadly approved professional practices are scientifically unsound and, therefore, do not meet the standards for admissibility.

## THE EMERGING RESEARCH AGENDA CONCERNING EXPERT TESTIMONY

The skepticism of courts toward expert testimony in general and scientific testimony in particular seems rooted in a view of science and litigation that has caused the courts to be extremely cautious in acknowledging the value of some scientific methodologies in providing an informed assessment of causal relationships. Consequently, courts often declare common methods of professional assessment based on animal research or clinical inference to be so lacking in scientific rigor that they fail to meet a suitable standard for consideration by the jury.

Several research projects currently under development are likely to shed new light on the use of scientific and clinical medical testimony in litigation. Carl Cranor and David Eastmond have received a grant from the National Science Foundation to develop reviews by independent scientists of synopses of expert testimony used in a particularly difficult toxic tort case. The assessments of the scientists can then be compared with the court's interpretation of what constitutes reliable scientific evidence. Such a study is likely to sharpen the comparison of the interpretations of scientific methodology inside and outside of the court.

The Science, Technology & Law Panel of the National Academies has developed a proposal to examine the characteristics of "litigation science," or science that has been developed in the context of the litigation process. Judge Kozinski recognized that science developed in the context of litigation may be subject to pressures that may distort the findings when he cautioned the following<sup>2</sup>:

One very significant fact to be considered is whether the experts are proposing to testify about matters growing naturally and directly out of research they have conducted independent of the litigation, or whether they have developed their opinions expressly for purposes of testifying. . . . [I]n determining whether proposed expert testimony amounts to good science, we may not ignore the fact that a scientist's normal workplace is the lab or the field, not the courtroom or the lawyer's office.

Through a series of case studies, the Science, Technology & Law Panel plans to examine a number of distinctive features of such litigation, including the litigation circumstances that brought about the research, the extent to which the participants in litigation participate in the design, analysis, and interpretation of the research, and the extent to which the research finds an audience beyond the participants in the litigation. This study is intended to identify criteria that courts can use to assess the quality of such scientific studies.

The Science, Technology & Law Panel of the National Academies also has developed a proposal to examine the scientific foundation of forensic science testimony submitted in criminal cases. This topic has been generally neglected by the broad scientific community, but it has grown in importance after the *Kumho* decision with the extension of evidentiary standards of reliability to all areas of expert testimony. An empirical analysis of appellate decisions in criminal cases by Groscup et al. indicates that the *Daubert* factors have been rarely used outside of forensic areas that are clearly scientific.<sup>26</sup> A more explicit role in screening expert testimony in criminal cases has emerged following *Kumho* and the amendments to Federal Rule of Evidence 702. The panel will convene a committee to formulate a research agenda for forensic science disciplines related to identification (e.g., fingerprints, tool marks, footprints, tire treads, question documents, and handwriting) to strengthen the scientific methodology underlying these areas of forensic science and to promote academic research in forensic sciences.

More research is needed in the manner in which attorneys identify, recruit, and prepare experts for testimony. Most difficult will be identifying the manner in which consulting experts who are not designated to testify at trial are used to shape the claims and defenses, because such activities are protected by attorney work-product privilege. Finding a similar opportunity to explore these issues with those who have served as experts will also be difficult.

Finally, more research is needed on the extent to which courts conduct pretrial inquiries into the reliability of expert testimony. The Federal Judicial Center examined such activities in cases that went to trial, but that study missed those cases in which such an

examination resulted in the case terminating prior to trial. We do not know the extent to which judges engage in the screening of expert testimony as part of a routine pretrial process and how this screening process varies across areas of expert evidence. ■

### About the Author

The author is with the Federal Judicial Center's Program on Scientific and Technical Evidence, Washington, DC.

Requests for reprints should be sent to Joe S. Cecil, Federal Judicial Center, One Columbus Circle, NE, Washington, DC 20002 (e-mail: jcecil@fjc.gov).

This article was accepted August 9, 2004.

**Note.** The views expressed in this article are those of the author and do not necessarily represent the views of the Federal Judicial Center.

### Acknowledgments

Portions of this paper were presented at the Coronado Conference on Scientific Evidence and Public Policy, Coronado, Calif, March 2003. Portions of the discussion of the *Soldo v Sandoz* case also appear in "Construing science in the quest for 'ipse dixit': a comment on Sanders and Cohen."<sup>27</sup>

### References

1. *Daubert v Merrell Dow Pharmaceuticals, Inc*, 509 US 579 (1993).
2. *Daubert v Merrell Dow Pharmaceuticals, Inc* (Daubert II), 43 F3D 1311, 1316 (9th Cir 1995).
3. Songer DR. Nonpublication in the United States district courts: official criteria versus inferences from appellate review. *J Politics*. 1988;50:206–215.
4. Gross SR. Expert evidence. *Wisc Law Rev*. 1991; 1113–1184.
5. Gross SR, Syverud KD. Don't try: civil jury verdicts in a system geared to settlement. *UCLA Law Rev*. 1996;44:1–64.
6. Champaign A, Schuman D, Whitaker E. Expert witnesses in the courts: an empirical examination. *Judicature*. 1992;76:5–10.
7. Kafka CL, Dunn MA, Johnson MT, Cecil JS, Miletich D. A survey of judges' and attorneys' experiences, practices, and concerns regarding expert testimony in federal civil trials. *Psychol Public Policy Law*. 2002;8:309–332.
8. Dixon L, Gill B. Changes in the standards for admitting expert evidence in federal civil cases since the *Daubert* decision. *Psychol Public Policy Law*. 2002;8: 251–308.
9. *Braun v Lorillard Inc*, 84 F3d 230, 234 (7th Cir 1996).
10. *O'Conner v Commonwealth Edison Co*, 13 F3d 1090, 1106-7 (7th Cir 1994).
11. *Claar v Burlington Northern R. Co*, 29 F3d 499, 502 (9th Cir 1994).
12. Gatowski SI, Dobbin SA, Richardson JT, Ginsburg GP, Merlino ML, Dahir V. Asking the gatekeepers: a national survey of judges on judging expert evidence in a post-*Daubert* world. *Law Hum Behav*. 2001;25: 433–458.
13. Kovera M, McAuliff B. The effects of peer review and evidence quality on judge evaluations of psychological science: are judges effective gatekeepers? *J Appl Psychol*. 2000;85:574–586.
14. *General Electric Co v Joiner*, 522 US 136 (1997).
15. *Newman v Motorola, Inc*, 218 FSupp2d 769, 780-1 (D Md 2002).
16. *Amorgianos v National Railroad Passenger Corp*, 137 FSupp2d 147, 189 (ED NY 2001).
17. *Mitchell v Gencorp Inc*, 165 F3d 778, 782 (10th Cir 1999).
18. *Cacciola v Selco Balers, Inc*, 127 FSupp2d 175, 184 (ED NY 2001).
19. Elliott ED. Toward incentive-based procedure: three approaches for regulating scientific evidence. *Boston Univ Law Rev*. 1989;69:487–508.
20. *Kumho Tire Co, Ltd. v Carmichael*, 526 US 137 (1999).
21. Kassirer JR, Cecil JS. Inconsistency in evidentiary standards for medical testimony: disorder in the courts. *JAMA*. 2002;288:1382–1387.
22. *Soldo v Sandoz Pharmaceuticals Corp*, 244 FSupp2d 434 (WD Pa 2003).
23. *Glastetter v Novartis Pharm Corp*, 252 F3d 986 (8th Cir 2001).
24. *Rider v Sandoz Pharmaceuticals Corp*, 295 F3d 1194 (11th Cir 2002).
25. *Globetti v Sandoz Pharmaceuticals Corp*, 111 FSupp2d 1174 (ND Ala 2000).
26. Groscup JL, Penrod SD, Studebaker CA, Huss MT, O'Neil KM. The effects of *Daubert* on the admissibility of expert testimony in state and federal criminal cases. *Psychol Public Policy Law*. 2002;8:339–372.
27. Cecil JS. Construing science in the quest for "ipse dixit": a comment on Sanders and Cohen. *Seton Hall Law Rev*. 2003;23:967–986.