Commentary

Demystifying the Law/Science Disconnect

Michelle M. Mello and
Troyen A. Brennan
Harvard School of Public Health

As the symposium in this issue of the *Journal of Health Politics, Policy and Law* makes clear, there remains a great deal of controversy and uncertainty surrounding the treatment of scientific evidence by the courts. Many of the issues are not new: they began to be identified in the legal literature in the mid-1980s (Brennan and Carter 1985). The fifteen years of struggle by courts and legal commentators have not, however, resolved the problems that occur when judges are faced with complex questions of scientific fact. To use a medical metaphor, we have been struggling to make a diagnosis—to clearly articulate the nature of these problems, with a view to proposing therapeutic reforms.

A key assumption underlying these diagnostic efforts seems to be that the gulf between law and science as disciplines is so vast that it can never fully be bridged. This is a central argument of the crisp article by John M. Eisenberg, the administrator of the Agency for Healthcare Research and Quality (AHRQ). Eisenberg identifies a series of differences between law and medicine and concludes that given these differences in institutional views, judges' problems with medical and scientific issues are not only understandable but inevitable.

One could, we think, respectfully take issue with some of the dichotomies Eisenberg sets forth, and we do so below. At the same time, however, we recognize that these concerns about the institutional competence of courts to make medical and technical judgments are, to judges, very real. Permeating the legal decision making in matters related to health

Journal of Health Politics, Policy and Law, Vol. 26, No. 2, April 2001.

care and science is a palpable self-consciousness and reluctance to venture into waters uncharted in the course of a judge's training. Thus, while the perceived disparities between law and science that Eisenberg highlights may be somewhat overstated, they do help explain why judges decide health care cases as they do. In particular, they shed light on judges' deep ambivalence about trying to use techniques of evidence-based medicine, including cost-effectiveness analysis, to adjudicate health care disputes.

Eisenberg usefully describes four ways in which law and science are commonly believed to diverge. First, while physicians engage in probabilistic reasoning in making a diagnosis, courts must decide actual causality in individual cases. Second, physicians diagnose problems on a prospective basis, but courts make decisions "after the fact." Third, the work of scientists is subjected to peer review, while that of judges and litigants is not. Finally, scientific advancement is a process of synthesizing the findings of myriad independent researchers working on the ground level, while changes in the law are brought about by watershed decisions handed down by high courts. Each of these points offers insight into the unique features of scientific and legal reasoning, but, we believe, each invites a response.

Eisenberg's first assertion, regarding probabilistic versus case-specific reasoning, is true in a strict sense. However, it overlooks the extent to which judicial determinations of actual causality do involve probabilistic judgments. In a medical negligence case, for example, the question for the trier of fact is, "Did the plaintiff prove that it is more likely than not that the defendant's negligence caused the plaintiff's injury?" This probabilistic determination is not much different from the questions physicians face: "Are these symptoms more likely a manifestation of illness X than illness Y?" and "How likely is it that treatment Z will cure the problem?"

Moreover, in addition to making probabilistic judgments, courts increasingly are using probabilistic evidence as a basis for decision. While twenty years ago it seemed that courts had a very poor comprehension of statistical evidence, and indeed the entire field of epidemiology (Brennan 1988), the situation has changed a great deal. Both judges' level of understanding and their willingness to use statistical evidence in consideration of complex scientific disputes have increased. With the burgeoning of the "toxic tort" and product liability litigation industry, and the explosion of new drugs and medical technologies giving rise to such litigation, judges are called upon to adjudicate scientific disputes with

increasing frequency. Acquiring a basic facility with statistical and epidemiological methods has become a necessity.

Eisenberg's second point is that the perspectives of law and medicine are different, with the law making decisions retrospectively and medicine taking a prospective view. But again, the nature of causality determinations in tort litigation casts this characterization into doubt. The fundamental basis for a finding of proximate causation in a negligence case is foreseeability: Could the defendant reasonably have anticipated that his or her action would cause the particular harm that the patient suffered? In a strict product liability case, foreseeability plays an even greater role, also delineating which injured parties have standing to bring a claim in strict liability against a manufacturer. The "foreseeable users, foreseeable uses" rule opens the courtroom doors only to persons whose use of the product the manufacturer reasonably could have foreseen (American Law Institute 1998: Sec. 2). The trier of fact makes these foreseeability determinations from the perspective of what the tortfeasor knew at the time of his or her decision. This is essentially the same prospective vantage point that Eisenberg identifies as characterizing medical decision making.

There is, however, a key insight in Eisenberg's proposition about the difference between pre hoc and post hoc reasoning. Courts must decide at the time the dispute is presented. Scientific conclusions can be tentative, and final conclusions reached over time (Kaye 1992). This difference accounts for one of the most peculiar aspects of mass tort litigation: the fact that the emergence of new epidemiological evidence leads to markedly different judicial resolutions of similar cases over time (Sanders 1998). Nonetheless, the key perspective for the court, no matter when the decision is made in the cycle of a mass tort, is what the defendant knew at the time the critical decisions were made. This is certainly a prospective vantage point.

Eisenberg goes on to make the point that unlike the work of scientists, the work of judges and litigants is not subjected to peer review, though perhaps it should be. He contrasts courts to medical journals, arguing that courts do not take advantage of expert opinion in the same way that scientists turn to learned peers for review of scientific evidence. But two responses might be made to this. First, the expert witnesses called in litigation provide a form of peer review for the litigants. They report on the strength of the current scientific evidence supporting the plaintiff's claim. Of course, as Eisenberg points out, this is not the dispassionate, impartial expert analysis that is the hallmark of scientific peer review.

The experts in an adversarial process are swept up into the nature of the proceeding. Indeed, this point has led many commentators to advocate the use of court-appointed experts, an approach that is gaining some momentum as a result of the breast implant litigation (Walker and Monahan 2000). But the point remains that, however imperfect, expert testimony serves as a form of peer review for the positions asserted by the litigants in a dispute involving complex scientific issues.

The work of judges, too, is subjected to peer review via the appeals process. Lower court judges are acutely aware that their written opinion and the way in which they conduct judicial proceedings are subject to scrutiny and reversal by their brethren. In addition to the vertical appeals process common to state and federal court systems, the federal courts of appeals provide a horizontal mechanism for review: the petition for a rehearing *en banc* of a case decided by a three-judge panel. Moreover, horizontal review also occurs when a lower court is presented with a question of first impression in its jurisdiction and must decide whether to follow the decisions of lower courts in other jurisdictions that have considered the issue.

There is much evidence to suggest that inferior court judges are very apprehensive of being reversed on appeal because their professional audience may question their judgment and abilities (Caminker 1994: 827). As one state court judge has succinctly put it, "There are no eight letters more intimidating to a trial judge than R-E-V-E-R-S-E-D" (Breeden and Bryan 2000). Thus, when reviewing courts decline to follow a judge's decision, the sting is much the same as that felt by a scientist whose paper is rejected by a journal due to methodological flaws. It is likely that the two forms of peer review exert similar quality-control effects.

The final law/science dichotomy that Eisenberg cogently describes is that while sea changes in the law are brought about by an important single decision by a high court judge, scientific advancement appears to be an aggregation of a series of findings by broadly disparate research groups. We disagree with this "top-down" and "bottom-up" description of law and science. Eisenberg's discussion usefully brings to mind the fact that while there is a Supreme Court to step in where lower courts have split in their decisions on a particular legal question, there is no such arbiter to resolve fissures in the scientific community. However, Eisenberg underestimates the extent to which the law is bottom-up. Most judicial decision making is done in the "laboratories" of the lower courts. Like scientists who base their protocols on methodologies developed by

their predecessors in the field and innovate on them in order to address new scientific questions, judges consult legal precedent to ascertain how others on the bench have handled similar cases and, in general, depart from precedent only where necessary to resolve a novel factual situation or issue of law. In this sense, the development of legal principles is very similar to the coalescence that occurs in science.

This is not to say that this evolution-through-aggregation is always enlightening to students of scientific evidence in the law, or to judges themselves. Even when legal edicts emanate from the Supreme Court, they may not provide clear guidance, particularly when read in combination. For example, Professors Daniel W. Shuman and Susan Haack in their articles in this issue provide excellent descriptions of the confusion created by the *Daubert-Joiner-Kumho* trilogy of cases. *Daubert* can be understood as an effort to reconcile the long-standing conflict between the liberal Federal Rules of Evidence and the restrictive *Frye* rule. However, in many respects, *Daubert* increased rather than ameliorated the confusion over what constitutes "reliable" scientific evidence. *Daubert* appeared to set forth some relatively strict criteria for judges to assess scientific evidence, including a focus on methodology rather than conclusions and reliability. The opinion was amenable to a multitude of different interpretations, however, with the result that the lower courts issued a series of conflicting and confusing opinions attempting to flesh out *Daubert*'s mandate (Graham 2000).

The Supreme Court intervened in *Joiner* and *Kumho* in order to resolve some of the most significant ambiguities in *Daubert*. *Joiner*, however, served only to cast doubt upon the criteria for the evaluation of scientific evidence that the Court set forth in *Daubert*. In holding that a judge could exclude expert testimony even though it is based on studies that are methodologically sound if there is "too great an analytical gap between the data and the opinion proffered," *Joiner* somewhat undermined *Daubert*'s pronouncement that "the focus, of course, must be solely on principles and methodology, not on the conclusions that they generate" (Saks 2000). The Court's subsequent opinion in *Kumho* clarified that the *Daubert* criteria applied to all types of expert testimony, not just "scientific" testimony, but also confused matters by holding that trial courts could decide on an ad hoc basis whether or not the specific factors described in *Daubert* are applicable to a given case.

Trying to render the *Daubert-Joiner-Kumho* trilogy comprehensible is impossible unless one acknowledges the historical and political context for the decisions. By the early 1990s, the courts were faced with a grow-

ing public perception, produced by mass tort litigation, that judges were incapable of discriminating real science from so-called junk science, particularly in the realm of clinical ecology. *Daubert* was a reaction to this, empowering judges to eliminate evidence that they did not find to be "reliable." *Daubert* hearings became a hallmark of mass tort litigation, used by relatively well-heeled defendants to increase the costs of litigation for plaintiffs. Plaintiffs in those cases that involved excellent epidemiological literature (in particular, asbestos and tobacco litigation) had the resources to prevail in these costly hearings, but the *Daubert* decision did help curb the growth of mass tort litigation based on science that did not rise to the level of good epidemiology. *Joiner* and *Kumho* can be understood as tinkering with *Daubert*'s heavy-handed effort to eliminate certain kinds of scientific evidence. But even if one comprehends the general goals of this trilogy of cases, the specific details of how lower courts are to carry out their mandate remain obscure. Legal reasoning about the role and admissibility of scientific evidence, as in other areas of law, is serpentine and slow to coalesce.

This point crystallizes our main departure from Eisenberg's view. Courts' difficulties with complex scientific issues are not a matter of disconnect of two cultures but rather reflect a historical evolution influenced by political issues and complicated by the maturation of the science of epidemiology. Ours is an optimistic view: Cultural disconnect is less likely to be resolvable than is a lack of consistency predicated on evolutionary trends.

This brings us to Eisenberg's final proposition, that rules of evidence are not well-established either for clinical medicine or for scientific issues in the courts. While we might find reasons to disagree with all of his preceding propositions, this one seems to us to be quite correct and points the way to a reasonable diagnosis. Neither federal jurisprudence nor state decisions have really enlightened the average judge as to how to evaluate scientific or technical evidence presented in a courtroom. According to Cynthia Mulrow and Kathleen Lohr, the same problem is experienced by the physician who must rely on a variety of different types of studies, with little training in how to weigh that evidence and apply it to the clinical setting. In such a situation, physicians and judges have two options: fall back on traditional rules of practice, or make a rather uneducated guess as to how to proceed in a new direction. The first course halts the evolution of clinical practice and legal doctrine; the second inevitably leads to inconsistency in decisions. It is an awkward dilemma.

One area in which this dilemma has presented itself in the courts relates to the use of cost-effectiveness analysis (CEA) in the adjudication of health care cases. Courts have exhibited varying degrees of willingness to use CEA, depending on the nature of the claim before them. Judges' uncertainty about how to evaluate scientific evidence, and their belief that legal culture is fundamentally different from the culture of science, has caused them to adhere to familiar forms of analysis rather than embark on a broad project of judicial CEA.

This juridical anxiety about CEA has led to quite different outcomes in seemingly similar cases. As Peter D. Jacobson and Matthew L. Kanna observe in their article, courts have had no problem explicitly incorporating CEA into ordinary negligence cases. They determine the standard of care by applying the Hand formula of *United States v. Carroll Towing* (159 F.2d 169 [2d Cir. 1947]). However, in medical negligence cases, they are much less likely to use CEA, with *Helling v. Carey* (519 P.2d 981 [1974]), constituting a notable exception. Instead, they defer to the medical profession and accept medical custom as the standard of care.

The same sort of dichotomy can be seen in product defect cases. CEA is perfectly acceptable for a broad variety of product liability cases, as Jacobson and Kanna note. For most products, courts will deem a product defective in design if, in their judgment, the technical evidence shows that at the time of manufacture there existed an alternative design that was safer and economically feasible (American Law Institute 1998: sec. 2). This explicit weighing of risks, benefits, and costs closely parallels the techniques of evidence-based medicine. However, the nuance that is absent from Jacobson and Kanna's account is that the *Restatement (Third) of Torts* suggests that certain medical products (drugs, vaccines, and medical devices) should not be subject to this rule. Manufacturers of these products are given heightened protection against defect claims: they need only show that the benefits of their product outweigh the risks for at least one class of users (ibid.: sec. 6[c]). Economic considerations are absent from this determination.

It is very difficult to identify a principled basis for the *Restatement*'s differential treatment of medical and other products. The regime is typically justified by reference to the need to encourage rapid development of new medical technologies by insulating manufacturers from legal liability (Conk 2000). However, it seems to us just as likely that it is driven by judges' sense of inadequacy to make determinations about the relative merits of different designs for these complex products.

A third area in which ambivalence about CEA can be seen is the liti-

gation in managed care cases. *Pegram v. Herdrich* (120 S. Ct. 2143 [2000]), clearly states that judges are not capable of engaging in medical level decision making about spending or rationing. These issues are best left to the legislatures. On the other hand, when asked to evaluate whether a treatment is experimental and thus excluded from coverage by an insurance policy, the courts have been quite willing to evaluate the evidence using various forms of clinical effectiveness or cost-effectiveness analysis. The litigation over coverage of autologous bone marrow transplantation plus high-dose chemotherapy for the treatment of breast cancer provides a large number of examples in which courts delve into just the sorts of issues that *Pegram* eschews. In characterizing benefit denial cases as narrow matters of contract interpretation, Jacobson and Kanna understate the extent to which courts engage in technical evaluations of the effectiveness of the treatments at issue. Reviewing an insurer's decision to classify a treatment as experimental involves scrutinizing evidence from clinical trials, analyzing study methodologies, and addressing deep questions about the level of scientific certainty required in order to deem a new therapy sound medical practice.

The courts' somewhat schizophrenic treatment of CEA may reflect judicial ambivalence about the entire project of CEA, an ambivalence we also see in managed care organizations and other players in the health care realm. Or it could simply be a manifestation of judges' anxieties about the lack of firm rules for evaluating scientific evidence and their own institutional competence to do CEA. Faced with a lack of confidence about the science, judges fall back, as physicians do, on accepted wisdom.

What can be done? Eisenberg points out that some answers are clear. Better training for judges about scientific evidence would help, although we must agree with Chief Justice William H. Rehnquist's opinion in *Daubert* that high hopes cannot be pinned on such training. While judges may gain some proficiency in techniques of evidence-based medicine, as they did in epidemiology, they are never going to become scientists.

The use of court-appointed experts to assist them in making technical decisions, encouraged by the American Law Institute project on tort reform in the late 1980s (Brennan 1991: 110–135), appears to be gathering some steam after the breast implant litigation. Nonetheless, it would only be feasible to use these court-appointed experts in the highest-stakes cases. Another possibility would be to remove some types of scientific and technical decisions from judges' purview altogether. For example, expert panels convened by administrative agencies could make determinations about what the scientific community considers to be "experimen-

tal" treatments. This would seem to have considerable potential for easing judges' burdens in insurance coverage disputes; however, significant study would have to be undertaken before moving to such a radical reform.

Jacobson and Kanna's suggestion that managed care organizations, rather than courts, should take the lead in expanding the use of CEA is a helpful one. Although health plans make decisions about coverage in the face of financial conflicts of interest, their institutional expertise in the area of CEA far exceeds that of judges. Courts could usefully serve as watchdogs over their decision making, looking for instances in which these conflicts of interest corrupted the CEA process. This is essentially the role they play now in regards to coverage decisions by ERISA plans.

In terms of a research agenda, perhaps the most critical need today is to better understand the characteristics of cases that lead judges to decisions that seem inconsistent. This will require investigation on a much broader footing than the research that has been done to date on the use of scientific evidence by courts. First, there is a need to closely review those areas in which repetitive decisions have been made by courts, in order to understand the characteristics of cases that lead judges to fall back on traditional analyses rather than probe the scientific evidence. Only controversies in which there are numerous opinions about a specific issue will allow us to apply econometric and statistical methods to identify these characteristics.

Second, we must go "upstream" from published decisions to analyze how courts are handling the broader universe of cases, including those that are eventually settled or for other reasons do not result in a written opinion. This would involve either a RAND-style study of jury verdicts or collaboration with chronic defendants, such as insurers, to assess the way in which repetitive claims have been adjudicated.

Survey methods could also be much more helpful than they have been in the past. Comprehensive surveys of federal or state judiciaries have been undertaken by the Federal Judicial Center and others, but sophisticated probing based on specific sets of facts have not been part of those surveys. Nor have plaintiffs and defense attorneys been surveyed about similar subjects to identify differences in their understanding of the science and their view of the strategies that lead to success in litigation. As a result, we lack basic data on the sorts of problems that courts face with scientific evidence. Given this lack of knowledge, it seems too early to jump to specific reforms, such as use of bifurcation, court-appointed experts, or science panels.

AHRQ and the Institute of Medicine recognize that the use of scien-

tific evidence by courts will be an enduring problem, and they are committed to fostering the basic research necessary to begin to design appropriate therapeutics. Based on what we know today, the outlines of this research agenda are beginning to become clear. The IOM and AHRQ have set forth the challenge; it is up to social scientists, legal academics, scientists, and jurists to take the next steps to bridge the divide between law and science.

References

American Law Institute. 1998. *Restatement (Third) of Torts: Products Liability*. Philadelphia: American Law Institute.

Breeden, J. L., and W. A. Bryan. 2000. Improving Jury Deliberations: Perspectives from the Circuit Court Bench. *South Carolina Lawyer* (September/October): 18–24.

Brennan, T. A. 1988. Causal Chains and Statistical Links: The Role of Scientific Uncertainty in Hazardous-Substance Litigation. *Cornell Law Review* 73:469–533.

———. 1991. Gaps in Health Insurance. In *Final Report. American Law Institute, Compensation and Liability for Product and Process Injury*, ed. Paul C. Weiler. Philadelphia: American Law Institute.

Brennan, T. A., and R. F. Carter. 1985. Legal and Scientific Probability of Causation of Cancer and Other Environmental Disease in Individuals. *Journal of Health Politics, Policy and Law* 10(1):33–80.

Caminker, E. H. 1994. Why Must Inferior Courts Obey Superior Court Precedents? *Stanford Law Review* 46:817–873.

Conk, G. W. 2000. Is There a Design Defect in the *Restatement (Third) of Torts: Products Liability*? *Yale Law Journal* 109:1087–1133.

Graham, M. H. 2000. The Expert Witness Predicament: Determining "Reliable" under the Gatekeeping Test of *Daubert, Kumho*, and Proposed Amended Rule 702 of the Federal Rules of Evidence. *University of Miami Law Review* 54:317–357.

Kaye, D. H. 1992. Proof in Law and Science. *Jurimetrics Journal* 32:313–359.

Saks, M. J. 2000. The Aftermath of *Daubert*: An Evolving Jurisprudence of Expert Evidence. *Jurimetrics Journal* 40:229–241.

Sanders, J. 1998. *Bendectin on Trial: A Study of Mass Tort Litigation*. Ann Arbor: University of Michigan Press.

Walker, L., and J. Monahan. 2000. Scientific Authority: The Breast Implant Litigation and Beyond. *Virginia Law Review* 86:801–833.