

Evidence: Its Meanings in Health Care and in Law

(Summary of the 10 April 2000 IOM and AHRQ
Workshop, "Evidence": Its Meanings and Uses in Law,
Medicine, and Health Care)

Clark C. Havighurst
Duke University School of Law

Peter Barton Hutt
Covington & Burling

Barbara J. McNeil
Harvard Medical School

Wilhelmine Miller
Institute of Medicine

In April 2000 the Institute of Medicine (IOM) and the Agency for Health-care Research and Quality (AHRQ) jointly hosted a one-day workshop to explore an intriguing and important intersection of medicine and law: the courtroom presentation of science-based medical evidence and expertise. This workshop was inspired by a concern that legal uses and interpretations of science-based medical evidence, particularly population studies and the findings of controlled clinical trials, may diverge substantially from the uses and interpretation of that evidence by the medical and health care researchers who produce it and of the practitioners and health plans that use it in making clinical decisions and policies.

Recognizing that a preliminary discussion among professions was needed even to describe the nature of their differences, the IOM and AHRQ, at the instigation of John M. Eisenberg, director of AHRQ, convened about twenty clinicians, epidemiologists, health services researchers, health plan executives, practicing and academic lawyers, jurists, and social scientists in the field of legal medicine (see appendix for participants). Participants and presenters were asked to formulate empirical research questions concerning both evidence-based medicine (EBM) and judicial practices that might increase familiarity with, and therefore promote greater reliance on, the use of science-based medical evidence by the courts. Workshop participants were further asked to identify policy

issues relating to the application of evidence-based medical findings that were emerging in the context of congressional consideration of patient protection legislation and reform of health plan liability law.

The four background papers commissioned for this workshop provided the participants with a common frame of reference for the issues to be addressed during the day. These papers were the first drafts of the authors' articles in this special issue. The authors were variously asked to address the following questions:

- What do physicians take to be evidence that justifies their practices and treatment decisions, and how has this understanding changed over time?
- To what extent has EBM affected the practice of medicine?
- What kinds of questions do rules of evidence allow medical experts to address in the courts?
- How do judges and juries understand and weigh scientific claims about the outcomes and efficacy of particular medical practices?
- What impact have recent Supreme Court decisions regarding the role of the judge in qualifying expert witnesses and screening scientific and technical evidence for presentation to juries had on malpractice cases and health plan coverage disputes?
- How are courts likely to deal with science-based medical evidence in cases involving health plan coverage disputes and medical necessity determinations under proposed liability reforms?
- What is the place of cost-effectiveness and cost-benefit analysis in health plan coverage policies, and how will courts consider coverage choices based on these kinds of analyses?
- How can those involved in developing the evidence base for medical practice most effectively present this information in legal settings?

In addition to these commissioned papers, Susan Haack, professor of philosophy at the University of Miami, provides a broad overview of judicial rulings on and interpretations of scientific evidence and expert testimony of the past century. In her article, "An Epistemologist in the Bramble-Bush: At the Supreme Court with Mr. Joiner," Haack argues that inquiry in the natural sciences is, in practice, similar to empirical inquiry of other sorts and cannot be distinguished by particular methodologies, contrary to what recent Supreme Court decisions have presumed. Substantive scientific knowledge, and not simply ascertaining that the proper scientific techniques or methods are followed in producing the evidence in question, is needed to determine the degree of war-

rant of a particular scientific claim or theory. In some cases, judges will not be able to avoid ruling on substantive scientific questions. Haack concludes her article by raising both practical and policy questions about the presentation of scientific evidence and expert testimony in the courtroom.

In “Proof and Policy from Medical Research Evidence,” Cynthia D. Mulrow, professor of medicine at the University of Texas Health Science Center–San Antonio and director of the San Antonio Evidence-Based Practice Center (EPC), describes the evolution of what physicians take to be evidence for the practices they adopt and reviews the principles now accepted in the medical research community for evaluating medical research evidence. In expanding her paper for publication here, Mulrow has been joined by Kathleen N. Lohr, chief scientist at Research Triangle Institute (RTI) and director of the RTI–University of North Carolina EPC, as coauthor. Lohr has worked extensively in conceptualizing and establishing evaluative criteria for clinical practice guidelines. The article reflects this work as well as Mulrow’s analysis of the evolution of EBM concepts and practice as presented at the workshop.

Daniel W. Shuman, professor of law at Southern Methodist University, whose research interests include scientific evidence and the law, was asked to consider the courts’ use of scientific evidence, particularly the terms under which expert witnesses testify. Shuman’s essay, “Expertise in Law, Medicine, and Health Care,” contrasts two models of the judge’s role in deciding what scientific evidence is presented to a jury: the traditional adversarial approach and the more recent “gatekeeper” approach. In particular, Shuman assesses how judicial gatekeeping has been influenced by the Supreme Court’s trilogy of decisions on the admissibility of expert testimony under the Federal Rules of Evidence: *Daubert v. Merrell Dow Pharmaceuticals* (509 U.S. 579 [1993]), *General Electric Co. v. Joiner* (522 U.S. 136 [1997]), and *Kumho Tire Co. v. Carmichael* (526 U.S. 137 [1999]). Shuman concludes that these landmark decisions have had less of an impact on judges’ prescreening of scientific evidence and expert testimony in medicine than might have been expected and offers several explanations for this. He further suggests several strategies for strengthening the scientific literacy of the legal profession and for enhancing the ability of lawyers and judges to identify the quality of scientific findings offered as evidence as an issue meriting judicial attention.

Peter D. Jacobson, professor in the Department of Health Policy and Management in the University of Michigan School of Public Health, presented the third paper at the workshop. Here his article “Cost-Effectiveness Analysis in the Courts: Recent Trends and Future Prospects,” with

coauthor Matthew L. Kanna, addresses the question of whether and how courts consider the use of cost-effectiveness analysis (CEA) and related evaluative techniques such as cost-benefit analysis and risk-utility analysis by health plans to establish coverage policies and justify medical treatment decisions. Finding little health care litigation that explicitly involved the application of CEA, Jacobson examines its role in product liability cases and how juries have reacted to its application by manufacturers in making product safety design and recall decisions. Finally, Jacobson considers factors that might precipitate the explicit use of techniques such as CEA by health plans in developing benefit packages and making medical necessity determinations.

In the final presentation of the workshop, "Evidence-Based Medicine and the Law: The Courts Confront Clinical Practice Guidelines," Arnold J. Rosoff, professor of legal studies and health care systems at the University of Pennsylvania's Wharton School, reviewed courts' treatment of clinical practice guidelines (CPGs) as evidence of a standard of care in contrast with the traditional use of customary (local) medical practice as the standard of care applied in malpractice cases. Rosoff draws lessons from the consideration given to CPGs in the courts for the likely fate of science-based medical evidence when presented in litigation of health plan coverage policies and medical necessity determinations. His discussion gives particular attention to a variety of professional cross-cultural conflicts and communication gaps. Conflicts arise not only between physicians or medical researchers testifying as to treatment standard of care and the lawyers prosecuting a malpractice case or health plan coverage dispute, but also between the "artful practitioner" and the physician-scientist or medical researcher regarding the application of research to individual patient care decisions. Rosoff emphasizes the need in such conflicts for a common vocabulary and understanding of the questions at issue in such cases, a task that this workshop only began to address.

The remainder of this introductory essay highlights issues raised by participants in discussions that followed the authors' presentations. Several workshop participants were asked to serve as "first respondents." Barbara S. Hulka, Kenan Professor in the Department of Epidemiology at the School of Public Health, University of North Carolina, and Judge Sam C. Pointer Jr., formerly chief judge of the U.S. District Court for the Northern District of Alabama, commented on Mulrow's and Shuman's presentations. Drummond Rennie, adjunct professor of medicine at the University of California—San Francisco, and David M. Eddy, senior advisor for health policy and management at Kaiser Permanente Southern

California, followed with comments on Jacobson's and Rosoff's presentations. The discussion summary that follows is organized thematically, following the general order of the presentations.

The Meaning of Evidence and the Practice of Evidence-Based Medicine

In his introductory remarks, Kenneth I. Shine, president of the IOM, characterized medicine as "the largest cottage industry in the United States," one in which evidence-based practice is still a relatively young and controversial concept. He noted that science should be applied not only *in* the practice of medicine but *to* that practice as well, as the IOM had recently done in a report to the Health Care Financing Administration on defining "medical necessity." This turned out to be an extremely difficult task, Shine said, because outcomes data are sparse.

Shine observed that the distinction between "efficacy," evidence of an effect under ideal conditions, such as double-blind, randomized controlled trials, and "effectiveness," evidence of what actually works in practice, is a subtle and important one for all users of evidence. Last, he noted the challenge that time constraints posed, not only for bringing science-based evidence to bear on legal proceedings but also for its utilization by clinicians overwhelmed with new information.

Following Shine, John M. Eisenberg described his agency's mission as the sponsorship of research that produces evidence about effectiveness of health care practices and the translation of that research into practice. Although physicians believe that their practices have always been evidence based, many in the research community do not concede that this is so. Eisenberg identified three different levels at which issues of applying evidence in health care arise:

1. the clinical level, as practitioners make patient care decisions;
2. the level of a health care system as, for instance, in selecting particular drugs as part of a formulary or deciding which treatments will be covered under a health plan; and
3. the level of public policy, both after the fact, in legal standards established in court cases, and prospectively, as in Medicare coverage policy determinations.

Eisenberg questioned whether the rules of evidence, that is, the ways in which evidence is brought to bear on the question at hand, are the same in each context. How research-based medical evidence is characterized

200 Journal of Health Politics, Policy and Law

and how it is used to address different types of questions in different settings were recurring themes over the course of the day.

Mulrow acknowledged that medical research evidence is just one type of evidence that clinicians take into account. She pointed out that even with the explosion over the past decade of published studies (over 2 million annually), of biomedical journals (30,000), and of controlled trials of medical therapies (perhaps a quarter million), research evidence remains limited and spotty; it cannot answer all questions of medical practice policy for all patients and conditions. EBM emphasizes a structured and critical examination of the medical research literature. Mulrow noted that research evidence can be ambiguous and requires interpretation and judicious weighting of its significance. Just how such evidence is assembled and interpreted depends on the use to which it will be put (see Mulrow and Lohr's article in this issue).

The process of changing medical practice in response to EBM is gradual and irregular, many participants noted. Eddy remarked that despite physicians' assumption that their practice is rooted in empirical science, the past three decades have produced incontrovertible evidence that clinical practice deviates from research-based recommendations: "All the studies of variations in practice patterns, . . . of inappropriate care, when you look at what doctors actually do, compared with what we know does and doesn't work, we found we missed the mark not 2, 3, or 5 percent of the time, but 10, 20, 57 percent of the time. It is all over the place. . . . We have to drop the old assumptions."

Mulrow introduced another theme that was echoed in the discussions following her presentation: the relationship of research-based evidence to CPGs and medical standards of care. These concepts, particularly the notion of standards of practice or standard of care, were perhaps the most problematic in terms of what physicians and health care researchers, on the one hand, and lawyers, on the other, understood them to be and do in the context of medical care, as Mulrow observed at the close of the workshop:

CPGs ideally incorporate the findings of clinical and epidemiological research: Guidelines are systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances.¹ Methods of formulating guidelines may differ in several respects, including methods for identifying,

1. This definition was formulated in Field and Lohr 1990.

appraising, and ranking relevant research evidence; models for integrating indirect evidence; methods for incorporating experience and opinion; whether harms, costs, and values are explicitly considered; and sponsorship. . . . Most guidelines are an amalgam of clinical experience, expert opinion, and research evidence. . . . Guidelines most often apply to the general and not the particular. They require extrapolation to individual circumstance. . . . Following evidence-based guidelines may generally but not always assure good medical care, and diverging from guidelines does not always signal poor care. (Mulrow and Lohr in this issue)

Standards are another way in which recommendations based on research evidence might be expressed:

Unlike a guideline, which is a recommendation for best practices, standards are practices that are medically necessary and services that any practitioner under any circumstance would be required to render. . . . Formulating standards rather than guidelines requires a higher bar. . . . Evidence-based guidelines that focus on single conditions likely will inform, but not determine, standards of medical care that our society deems necessary. . . . research evidence in and of itself will invariably be inadequate to establish standards because standards will require priority setting based on cost and value judgments. (Mulrow and Lohr in this issue)

“Standards” were thus understood by the discussants to be enforceable by courts in malpractice cases and other legal disputes. Courts have traditionally established legal standards applicable to health care by reference to customary medical practice and prevailing medical opinion, as testified to by medical experts. Although some medical expert testimony may reflect the expert’s awareness of scientific literature as well as his or her knowledge of customary practice, courts have seldom treated such testimony as primarily scientific in character.

As Mulrow and Lohr indicate in their article, legal standards differ from scientific findings of efficacy and safety in potentially incorporating value judgments and cost considerations. However, as long as the legal system looks to medical custom as the principal source of standards, it can incorporate cost-benefit trade-offs only to the extent that physicians make such trade-offs in their clinical choices—as they may be doing increasingly under pressure from payers and managed care plans.

The discussants considered the extent to which CPGs might be used

202 Journal of Health Politics, Policy and Law

to narrowly prescribe physician practices in a particular health plan as, for instance, in the case of a plan that instructs its medical group to follow certain guidelines. Researchers in the field of EBM did not seem comfortable, however, with such prescriptive use of CPGs, which they characterize as nonbinding recommendations that reflect developments in EBM. The gap between thinking of CPGs as advice to clinicians and using them as prescriptive legal standards was not bridged by the discussants. CPGs, while potentially valuable in establishing standards of care in malpractice cases, have not often been employed in this fashion, as Rosoff notes in his article.

Rules of Evidence, Claims of Medical Expertise, and the *Daubert* Trilogy

The sea change in the treatment of scientific evidence and expert witnesses anticipated in the wake of the Supreme Court's 1993 *Daubert* ruling has not yet been realized, Shuman argues in his essay. The traditional "adversarial" model of the American legal system, he writes, "assumes we are more likely to uncover the truth about a contested event as the result of the efforts of the parties who have a self-interest in the outcome of the investigation than from the efforts of a judge charged only with an official duty to investigate the case." Under this model, the lower court trial judge's role of ruling on the admissibility of expert testimony has focused primarily on the expert's qualifications, leaving assessment of the expert's methods and procedures, the substance of his or her testimony, to be determined by the jury. The trial judge has significant discretion in evaluating the expert's qualifications, Shuman noted, citing a 1977 Sixth Circuit Court decision: "But the only question for the trial judge who must decide whether or not to allow the jury to consider a proffered expert's opinions is 'whether his knowledge of the subject matter is such that his opinion will most likely assist the trier of fact in arriving at the truth.'" (*United States v. Barker*, 553 F.2d 1013, 1024 [6th Cir. 1977]).

This traditional approach was presumably supplanted, at least in the federal courts, by a new model grounded in the three Supreme Court decisions within the past decade on the admissibility of expert testimony under the Federal Rules of Evidence: *Daubert v. Merrell Dow, General Electric Co. v. Joiner*, and *Kumho Tire Co. v. Carmichael*. The new gatekeeper model requires the trial judge to impose a more rigorous standard for admitting expert testimony, "ensuring that an expert's testimony rests

on a reliable foundation and is relevant to the task at hand (*Daubert*).” Many federal courts had relied on the District of Columbia Court of Appeals’ 1923 decision in *Frye v. United States* (293 F. 1013 [D.C. Cir. 1923]), which relied on professional consensus or general acceptance within the relevant professional community to determine the admissibility of novel scientific evidence. *Daubert* held that the Federal Rules of Evidence revised this standard. Both *Joiner* and *Kumho Tire*, the second and third of the Supreme Court’s decisions in this trilogy, made it clear that the trial judge has broad and almost unreviewable discretion in applying *Daubert* standards for assessing the admissibility of all expert evidence, Shuman explained.

Shuman’s review of the impact of the *Daubert* decision in different areas of civil litigation revealed a mixed record. He finds that the threshold for admissibility of medical and other kinds of expertise has risen significantly in toxic tort and product liability cases but not in medical malpractice litigation. Although critics of medical malpractice suits had hoped that *Daubert* would eliminate unreliable expert testimony in such cases, Shuman argues that so long as expert (physician) testimony as to customary practice within the local community is accepted on its face (as it continues to be), *Daubert* challenges to admitting expert testimony as to the standard of care are not likely to be successful. For these courts, the scientific validity of practice standards is simply beside the legal point at issue. To put it more directly, if (as other discussants noted) medical practice does not routinely conform to the best and most current scientific evidence, the courts’ continued focus on professional custom for setting the legally recognized standard of care reinforces this disregard of scientific knowledge.

Shuman noted that a number of stratagems can be employed by judges to sidestep the issue of the validity of scientific evidence in deciding a case, but that these appear to some to be applied selectively. He argues that the standard of scientific rigor has consistently increased in product liability and toxic tort cases, where it works to the benefit of defendants. In contrast, in criminal cases, where the prosecution and defense tend to rely on the very same experts in different cases, the *Daubert*-based challenges to medical testimony have not been raised. If EBM is to be taken seriously in medical quality of care and coverage cases, lawyers and judges must come to believe that scientific validity is truly what matters. If rigorous, empirically based evidence is insisted upon only in cases where it is advantageous for one class of litigants, the demand for better science is unmasked as merely a (possibly biased) legal tactic.

204 Journal of Health Politics, Policy and Law

In responding to Daniel Shuman's argument as to the perception of bias in the application of *Daubert*-inspired standards of evidence, Richard O. Lempert suggested that the *Daubert* line of cases is a judicial response to docket pressure. The *Daubert* rule has become one way that the judiciary has, over the past twenty years, increasingly used summary judgment beyond its classic purpose, which was to resolve civil cases when there was no genuine issue of material fact. *Daubert* increases the number of situations in which summary judgment can be used to dispose of cases that would ordinarily be very lengthy and expensive to try.

In *Joiner* and *Kumho Tire*, the courts have said they will not question, except in the most extreme situations, the trial judges' exercise of discretion as to the admissibility of scientific evidence. According to Lempert, some early cases following *Daubert* suggest that when excluding "junk science" hurts tort plaintiffs, such questionable evidence will be excluded. By contrast, when the state offers junk science in criminal cases to better its chances of convicting a defendant, it will be admitted because that admission will seem to further justice. At some point the practice of using shifting standards of admissibility is going to become obvious and thus intolerable. A single standard must be consistently applied, Lempert argued. He also predicted that the bite of *Daubert* and its progeny would not be limited to junk science, but that it would be used to exclude good scientific evidence in situations where one party's scientific evidence, though valid, seemed overwhelmed by the evidence on the other side. The danger, arguably realized in *Kumho Tire*, is that courts, under the guise of deciding preliminary questions of admissibility, would be taking from juries issues that the Seventh Amendment gives to juries to resolve.

Lempert's final point was that although truth is the formal goal of the legal system, a plaintiff need not show his or her claims to be true to win a case and a defendant can prevail without proving the other party's claims are false. To prevail in a case, a party must demonstrate only that a preponderance of the evidence supports the claim. If there is weak scientific evidence for the plaintiff and no evidence for the defendant, is that enough to give the plaintiff a verdict? It depends either on whether the judge decides to exclude the plaintiff's weak evidence on the grounds that the science is not good enough to get to a jury, or on what the jury thinks of the evidence if it does hear it. Contrary to popular perception, overall tort juries show no noticeable pro-plaintiff bias, Lempert claimed, and plaintiffs who get to juries often lose there. A party can also prove its case almost indisputably with scientific evidence. These cases seldom

reach court reporters because they are usually dropped or settled without trial, depending on which side has the benefit of the science.

A final set of issues raised by Shuman's review of the *Daubert* decision's impact on malpractice and coverage cases concerned the extent and rate of uptake of *Daubert*. As long as medical experts are perceived to be testifying only on the matter of customary practice, they need not be qualified as expert scientists, since the opinions they express relate to what doctors actually do, not what scientific evidence suggests they should be doing. Although David Eddy's observations (quoted earlier) about the incongruence of EBM and actual medical practice might be understood to suggest that custom is a poor source of legal standards, the courts, adhering to a presumption that professionals generally practice according to scientific principles, have yet to look explicitly beyond medical practice directly to scientific evidence to define standards in malpractice cases. Likewise, coverage determinations based on "medical necessity" are often made by reference to professional opinion. For these reasons, courts have not subjected medical testimony in malpractice and coverage disputes to the closer kind of scrutiny that scientific testimony receives under *Daubert* and its progeny.

Joseph S. Cecil, of the Federal Judicial Center, suggested that federal judges have been caught off guard by the expectations that have arisen following the 1993 decision, and that they have made more than a good faith effort at trying to engage in an informed discussion about the basis of opinion. Acknowledging the limitations of the research he undertook in support of his presentation, Shuman pointed out that because settled cases are unreported, the extent to which courts employed the *Daubert* standard could not be learned from reported and appellate rulings alone, as these reflected only a small fraction of all cases. More detailed knowledge of judicial performance in malpractice cases and coverage disputes might therefore reveal some spillover from the federal *Daubert* ruling in state malpractice cases or in cases involving coverage issues.

This and other acknowledged limits of what we can know prompted several discussants to propose empirical research studies of how judges, lawyers, and juries evaluate, understand, and draw inferences from probabilistic and science-based evidence. Lohr suggested that probability and statistics might be applied to protect against making erroneous inferences and that links can be made back to the law from science and medical research in terms of how people think about reasonable doubt and preponderance of evidence. She noted that thresholds for statistical significance may be set for different purposes in health outcome measures, in

particular, self-report instruments that seek to measure quality of life. For example, the standard for making distinctions between groups is generally lower than the standard for individual patient decisions.

Coverage Policies, Medical Liability, and Cost-Effectiveness Analysis

Jacobson's presentation, "Cost-Effectiveness Analysis in the Courts: Recent Trends and Future Prospects," elicited questions and debates among the participants that revealed underlying differences in their assumptions and understandings about EBM and how it has been or should be incorporated into the design of health plans' benefit packages. Clinicians and researchers involved with the development and promotion of EBM tended to view cost-effectiveness as an evaluative criterion incorporated only into normative applications of research results, such as CPGs. Others were more likely to view cost-effectiveness analysis (CEA) as an implicit aspect of EBM policy and practice.

Jacobson acknowledged the wide variability in how CEA is defined, from "a method designed to assess the comparative impacts of expenditures on different health interventions" (Garber et al. 1996), to the determination of the lowest cost intervention that offers a clinically appropriate benefit (Eddy 1996). The simplest application of CEA is to compare two interventions aimed at the same outcome, although CEA may also be applied more broadly to design a health plan's overall benefit package, for example, where different services or preventive screening tests are considered for inclusion. Even in its more narrow applications, Jacobson noted, CEA is fraught with methodological difficulties and value judgments.

Jacobson was charged by the workshop planning group with a particularly challenging task: to extrapolate from very limited legal precedents and experience in other fields of liability litigation to the likely treatment of CEA in suits involving health plan coverage policies and medical necessity determinations. In his presentation at the workshop, Jacobson reported that, to date, health plans appear not to use CEA explicitly but, rather, rely on other cost containment tools. He offers several possible reasons for this and proposes that the most important one may be the limited evidence as to effectiveness of alternative interventions.

In responding, Drummond Rennie described CEA as a complex form of evidence, involving review of many individual studies and findings, which means that data, assumptions, and analytic models are chosen

according to subjective criteria. In many cases, Rennie contended, the evidence upon which CEAs are based is weak. He also argued that these methodological limitations, along with the publications bias that follows from commercial sponsorship of the research, severely compromises the scientific and practical value of such analyses. Recognizing that CEA is, nonetheless, an important tool, Rennie suggested—following the model of legal discovery—that all CEA study data, models, and assumptions be made public by putting them on the Web sites of the journals that publish them (see Rennie in this issue and Rennie and Luft 2000).

Eddy addressed the promise of CEA less skeptically than Rennie, arguing that almost any approach to this evaluative technique is better than the currently dominant approach in health care, which is not to take cost or opportunity costs into account at all. Eddy suggested that, in fact, health plans are using CEA, although not explicitly. Taking the term in its broadest sense, to the extent that plans have limited budgets for benefits, a responsible manager will think in terms of cost-effectiveness in allocating that budget. In addition, Eddy argued, individual physicians make cost-effectiveness decisions all the time. All aspects of medical practice employ thresholds, indications, and contraindications. With our growing ability to know what works and what does not in particular situations, these should be set more rationally than they have been in the past.

Transparency, openness, and consumer participation were themes of the panel discussion that followed. First, because value judgments are necessarily applied in constructing cost-effectiveness analyses, and because the choice of research-based evidence included in such analyses is subject to bias, several participants argued that the methodology and analytic choices made in any CEA should be public if the results are to be credible. Second, several panelists endorsed the notion of consumer involvement in the development of CPGs and health plan coverage rules, a process in which value-based choices are unavoidable. There was general agreement that involving consumer and patient advocates in weighing the evidence to make clinical practice recommendations is appropriate not only to broaden the considerations reflected in the guidelines but also to educate consumers and reform their expectations about what health care can realistically deliver in terms of risk reduction and life expectancy.

Jacobson traced the history of CEA in the courts (at least in the sense that risks and costs were weighed together) with respect to liability in general negligence cases and later in product liability. The practice of

evaluating risks and costs evolved from the pre-1930s “reasonable man” standard of common law negligence, through an explicit cost-benefit standard exemplified by Judge Learned Hand’s 1947 opinion in *United States v. Carroll Towing Co.* (159 F.2d. 169 [2d Cir. 1947]). This latter standard says that negligence occurs when the cost of investing in injury prevention is less than the expected liability, that is, the probability of injury multiplied by the damage wrought by that injury. In 1965 the American Law Institute adopted this standard in its *Restatement (Second) of Torts* (and is continued in its *Restatement [Third] of Torts: General Principles, Article 4* [1999]). Since the 1970s, strict liability cases for product design defects have been predominantly resolved according to explicit risk-utility analysis, which incorporates cost-benefit or cost-effectiveness tests (Jacobson and Kanna in this issue).

Jacobson used the performance of juries in product liability cases, when they are faced with evidence of corporations’ explicit consideration of cost-risk trade-offs, to speculate how juries might react to the application of CEA in deciding disputes about health plan benefits. Discussion of this question brought out that a jury’s perception as to who benefits from resource-conserving decisions is likely to be an important determinant of how it regards a corporation’s use of cost-benefit or cost-effectiveness analyses. If product design or health plan savings strategies are viewed as simply increasing corporate profits, these strategies will be judged less acceptable than if these efficiencies are seen as redounding to the consumers’ benefit in some way.

Medical liability diverges from general negligence primarily because courts have generally deferred to professional standards, supplied through expert testimony as to customary medical practice. Generally, only failure to adhere to customary practice constitutes negligence or malpractice. Jacobson noted that the customary practice standard is not intrinsically hostile to cost containment strategies, in that customary practice can evolve to incorporate efficiency-promoting treatment protocols or clinical practice guidelines. In assessing the response of courts to liability claims against a health plan or third-party payer whose coverage policies or utilization review procedures explicitly incorporate cost containment objectives, Jacobson argued that “judicial internalization of cost constraints” is gradually taking place in response to changes in the social and economic environment (Bovbjerg 1975).

Although much of his analysis focused on tort law, Jacobson noted that health plan coverage and benefit disputes also can, and often do, involve

contract claims. He observed that making contracts more explicit with regard to the use of CEA in coverage decisions would be one strategy for expanding its application. Discussants returned to the issue of more explicit and transparent contracts for health benefits as one approach to reducing consumer distrust and dissatisfaction with coverage and medical necessity decisions by health plans. Eddy argued that the grounds for coverage decisions by health plans, such as CEA, must be made explicit in contract language if the plan's coverage and utilization review policies are to hold up in court. Clark Havighurst observed that recent class action suits alleging that a health plan has systematically misrepresented the benefits the plan covers, or the degree of discretion the plan affords physicians in treating patients, make adequate disclosure and honest advertising central issues in health plan conduct and performance.

Participants found fuller disclosure of health plan coverage policies a promising strategy for increasing the acceptability to consumers of science-based medical practices and benefit packages designed according to cost-effectiveness criteria. At the same time, however, several voiced concerns about the capacity of consumers to understand complex and subtle health plan coverage rules, decision processes, and utilization review procedures. Alice Gosfield, among others, proposed that health plan members sit on advisory boards and review panels that design and interpret the plan's benefits and coverage policies. Various other commentators also suggested public regulation of health plan coverage policies and review and appeals procedures as a strategy to fix the limits both of health plan discretion and liability.

Using Evidence Appropriately in Medicine, Health Care, and the Law

At the end of the day the discussion centered on the uptake and assimilation of research findings into clinical practice and how the legal system might ease, or at least not impede, the adoption of evidence-based practices by clinicians and health plans. In his presentation Rosoff reasoned that if the goal of clinical practice guidelines is to change professional custom in the direction that outcomes-oriented research points, and if professional custom is the legal standard of care in medical liability cases, then courts must somehow recognize research-based clinical practice as a standard of care as well. Rosoff's essay here considers alternative mechanisms, private and voluntary as well as governmental, for

210 Journal of Health Politics, Policy and Law

the certification of CPGs as meeting standards of scientific validity. Such certification would aid courts in their review of multiple and potentially conflicting sets of such guidelines.

Discussants responded by pointing out that the IOM's Committee on Clinical Practice Guidelines had identified the attributes of good CPGs almost a decade ago and that these attributes had been operationalized in a way that allowed for evaluating the quality of sets of guidelines. The evaluation instrument has been refined in the United Kingdom and Europe over the past eight years, and a multinational study of evaluative criteria for CPGs is ongoing. Alice Gosfield proposed that just as there are attributes of good guidelines, attributes of evidence that are especially relevant in addressing public policy concerns such as fraud and abuse, medical necessity, and utilization review could be established by an agency such as AHRQ.

Eisenberg asked whether evidence must be couched in terms of prescriptive guidelines in order to be useful in legal situations or whether courts could use purely descriptive research findings directly. He noted that AHRQ evidence reports are issued without "shoulds" and that they are ideally factual statements, but he wondered whether this form of evidence would be adequate for judges and juries.

His questions elicited several responses. First, it is important to distinguish between evidence in systematic reviews and evidence reports, and CPGs. The former are the initial steps in developing a guideline, whereas a full guideline might involve consumer and patient inputs as well as legal counsel to formulate recommendations for clinical practice. The result of both steps is an evidence-based CPG. Second, although evidence can be used in its original form, in legal settings it may more commonly be used once it has been incorporated into normative CPGs. Third, basic evidence and professional guidelines will have distinctive functions within the courtroom. Guidelines can be used as presumptive of or exculpatory of negligence (regardless of whether or not their designers envisioned such a use for them). Evidence cannot by itself be used in this way. Guidelines thus must meet certain (higher) thresholds of incontrovertibility, if they are to be applied fairly.

As Lempert argued, by far the most important factor in jury trials is the weight of the evidence, rather than the identity of the parties or the quality of the legal counsel. Lohr asked just what was meant by evidentiary "weight" or "strength," wondering if some objective criteria, like those now being developed in health services research, might be introduced in court. The response was that, as in all other questions of evi-

dence, expert opinion will inevitably conflict and that if scientists themselves dispute the quality of evidence, what can we expect of judges and juries?

Eisenberg wrapped up the day with a recapitulation of the issues and questions raised and considered the potential for further research and resolution of them. His remarks are also included in this volume in the article “What Does Evidence Mean? Can the Law and Medicine Be Reconciled?” Several participants endorsed the notion of sustained examination and interaction among the legal and medical practice and research communities. This collection of articles and commentaries should help to extend and advance the conversations that need to occur among the professions.

212 Journal of Health Politics, Policy and Law

Appendix Participants in Attendance at the Workshop "Evidence: Its Meanings and Uses in Law, Medicine, and Health Care," 10 April 2000

Joseph S. Cecil, J.D., Ph.D.
Project Director
Program on Scientific and Technical Evidence
Division of Research
Federal Judicial Center
U.S. Department of Justice
Washington, DC

David M. Eddy, M.D., Ph.D.
Senior Advisor for Health Policy and Management
Kaiser Permanente Southern California
Washington, DC

John M. Eisenberg, M.D.
Director
Agency for Healthcare Research and Quality
Rockville, MD

Alice G. Gosfield, J.D.
Principal
Alice G. Gosfield and Associates, P.C.
Philadelphia, PA

Clark C. Havighurst, J.D.
Wm. Neal Reynolds Professor of Law
Duke University School of Law
Durham, NC

R. Brian Haynes, M.D., Msc., Ph.D.
Professor and Chair
Clinical Epidemiology and Biostatistics and Medicine
Faculty of Health Sciences
McMaster University
Hamilton, Ontario, Canada

Barbara S. Hulka, M.D., M.P.H.
Kenan Professor
Department of Epidemiology
University of North Carolina—Chapel Hill
Chapel Hill, NC

Peter Barton Hutt, LL.M.
Partner
Covington & Burling
Washington, DC

Appendix Continued

Peter D. Jacobson, J.D., M.P.H.
Associate Professor
School of Public Health
University of Michigan
Ann Arbor, MI

Douglas B. Kamerow, M.D., M.P.H.
Director
Center for Practice and Technology Assessment
Agency for Healthcare Research and Quality
Rockville, MD

Richard O. Lempert, J.D., Ph.D.
Francis A. Allen Collegiate Professor of Law and Professor of Sociology
School of Law
University of Michigan
Ann Arbor, MI

Kathleen N. Lohr, Ph.D.
Senior Program Manager for Health Services Research
Research Triangle Institute
Research Triangle Park, NC

Barbara J. McNeil, M.D., Ph.D.
Head
Department of Health Care Policy
Harvard Medical School
Boston, MA

Cynthia D. Mulrow, M.D., Msc.
Professor
Department of Medicine
University of Texas Health Science Center
San Antonio, TX

Mark A. Peterson, Ph.D.
Professor of Policy Studies and Political Science and Editor, *Journal of Health
Politics, Policy and Law*
UCLA School of Public Policy and Social Research
Los Angeles, CA

Honorable Sam C. Pointer Jr.
Judge
United States District Court for the Northern District of Alabama
Birmingham, AL

214 Journal of Health Politics, Policy and Law

Appendix Participants in Attendance at the Workshop "Evidence: Its Meanings and Uses in Law, Medicine, and Health Care," 10 April 2000 (continued)

I. Drummond Rennie, M.D.
Adjunct Professor
Department of Medicine
Institute for Health Policy Studies
University of California–San Francisco
San Francisco, CA

Arnold J. Rosoff, J.D.
Professor of Legal Studies and Health Care Systems
Department of Legal Studies
The Wharton School
University of Pennsylvania
Philadelphia, PA

Kenneth I. Shine, M.D.
President
Institute of Medicine
Washington, DC

Daniel W. Shuman, J.D.
Professor of Law
School of Law
Southern Methodist University
Dallas, TX

Harold C. Sox Jr., M.D.
Professor and Chair
Department of Medicine
Dartmouth–Hitchcock Medical Center
Lebanon, OH

Miles J. Zaremski, J.D.
Piper Marbury Rudnick & Wolfe
Chicago, IL

Alan R. Zwerner, M.D., J.D.
Senior Vice President, Strategic Planning
PersonalMD.com
Pleasanton, CA

References

- Bovbjerg, Randall. 1975. The Medical Malpractice Standard of Care: HMOs and Customary Practice. *Duke Law Journal* 1975:1375–1414.
- Eddy, David M. 1996. Benefit Language Criteria That Will Improve Quality While Reducing Costs. *Journal of the American Medical Association* 275:650–657.
- Field, Marilyn J., and Kathleen N. Lohr, eds. 1990. *Clinical Practice Guidelines: Directions for a New Program*. Washington, DC: National Academy.
- Garber, Alan M., Milton C. Weinstein, G. W. Torrance, and M. S. Kamlet. 1996. Theoretical Foundations of Cost-Effectiveness Analysis. In *Cost-Effectiveness in Health and Medicine*, ed. M. R. Gold, J. E. Siegel, L. B. Russell, and M. C. Weinstein. New York: Oxford University Press.
- Rennie, Drummond, and Harold S. Luft. 2000. Pharmacoeconomic Analyses: Making Them Transparent, Making Them Credible. *Journal of the American Medical Association* 283:2158–2160.

