

Commentary

From the Clinics to the Courts: The Role Evidence Should Play in Litigating Medical Care

E. Haavi Morreim

University of Tennessee–Memphis

Throughout this collection of essays, the Institute of Medicine and the Agency for Healthcare Research and Quality have identified an issue whose importance and nuances we are only beginning to appreciate. Although medicine has long claimed to be rooted in science, actual clinical care has often had only a limited scientific basis, resulting in inexplicably wide variations of care (Wennberg 1996). The past few years have witnessed a marked interest in evidence-based medicine (EBM), stemming from several concerns.

First, decades of double-digit health care inflation led to a recognition that enormous amounts of money have been wasted on interventions with little proven value. Health plans facing pressures to keep premiums down and profits up have moved aggressively to curb wasteful practices such as excessive hospitalizations and needless surgeries. The ruling norm under lavish insurance—“If it might help and probably won’t harm, do it”—has given way to a leaner norm: “Don’t do it, unless you can demonstrate its value.” Under this new rule even common, widely accepted clinical routines have met coverage denials, and irate providers are scrambling to gather the kind of data necessary to document the value of their care (Morreim 1994).

Second, the more forward-looking health plans aim, not just to cut costs, but to render care more rational. In many cases physicians’ clinical routines are based not so much on empirical evidence as on local

The author is grateful to Peter D. Jacobson and to Clark C. Havighurst for their very helpful comments on earlier drafts of this manuscript.

Journal of Health Politics, Policy and Law, Vol. 26, No. 2, April 2001. Copyright © 2001 by Duke University Press.

habits, malpractice fears, facilities availability, or even advertising.¹ And in other instances, physicians are failing to provide important, scientifically well-grounded interventions, such as those for ongoing management of chronic illnesses like diabetes, asthma, and hypertension.²

Third, some plans' rather drastic cost-cutting measures have occasioned concerns that basic quality of care is suffering. Thus, just as health plans question physicians' practices, many providers and purchasers have stridently challenged the scientific credibility of the guidelines by which plans have tried to enforce their preferred clinical practices.

Finally, beyond simply avoiding poor-quality care, many purchasers, particularly large employers, seek affirmative value for their dollars. No longer willing to pour money into care that may or may not produce good outcomes, or whose outcomes might be achieved much more efficiently, many buyers now expect health plans to demonstrate that premium dollars are well spent, including important forms of preventive care and disease management.³

Perhaps it is no mere coincidence that courts have likewise begun to demand more and better evidence from litigants demanding large sums of compensation for alleged damages. In both settings, large sums of money have been requisitioned, sometimes on no better basis than "junk science" (Huber 1991; Angell 1996). Whatever the connection, the essays in this collection explore important questions about the ways in which courts' quest for more and better evidence in other contexts may dovetail with health plans', providers', and purchasers' demands that clinical practices, and the guidelines sometimes imposed on them, reflect an adequate scientific foundation.

Mainly the questions explored in this collection are empirical: how has evidence-based medicine in fact affected clinical practice; how do judges understand and weigh scientific claims; how will courts address evidence-based medicine and cost-effectiveness analysis in coverage disputes; and so forth (see the Introduction by Clark C. Havighurst and others in this issue). This excellent foundation makes it possible to launch into some related normative (i.e., evaluative) issues that will be the focus

1. Burnum 1987: 1220–1222; Hardison 1979: 193; Reuben 1984: 592; Holoweiko 1995: 180–181; Woosley 1994: 250; Avorn, Chen, and Hartley 1982: 4–8; Lange and Hillis 1998: 1839; Avanian et al. 1998: 1901).

2. Legorreta et al. 1998: 457–464; Hartert et al. 1996: 386–394; Weiner et al. 1995: 1503–1508; Leape 1995: 1534–1537; Harris 1996: 838–842; Newcomer 1998: 32–35; Moser 1998: 1813–1816.

3. For example, the Pacific Business Group on Health, a coalition of over thirty large employers, requires its health plans to meet certain quality standards, including customer service, preventive care, and data collection (see Bodenheimer and Sullivan 1998: 1003–1007; Schauffler, Brown, and Milstein 1999: 134–142).

of this essay: how should courts respond to health plans' demands for evidence-based medicine; how should courts respond to plans' efforts to balance costs against benefits; how should courts screen testimony about physicians' alleged malpractice; what kinds and amounts of evidence should courts expect from litigants in medical cases. Thus this commentary represents, not so much a reflection on the core essays, as an exploration of some of the further issues those writings have prompted.

This move from empirical description into a normative discussion is important, because we cannot determine what courts ought to do simply by examining what they have done thus far. A powerful example comes from recent litigation concerning the use of high-dose chemotherapy with autologous bone marrow transplant (HDC/ABMT) for breast cancer. For well over a decade, women with advanced breast cancer were told this treatment offered hope, even though there was never any credible science behind the claim, just some theoretical promise alongside physicians' desperate desire to do something—anything—to help their patients. Indeed, some early studies had already indicated the treatment provides no benefit over standard chemotherapy and actually diminishes patients' prognosis in certain categories (see ECRI 1995). In other instances, studies allegedly showing benefit were methodologically deeply flawed.⁴

When health plans tried to deny coverage on the ground there was no evidence that HDC/ABMT is effective for breast cancer, desperate patients replied that the practice was well-accepted by physicians. In fact, both sides were correct. There was no good evidence, but physicians nevertheless did widely accept it. Hence, although a number of courts sided with health plans,⁵ a large number sided with patients.⁶ Between judicial

4. For instance, treatment-related deaths (death within one month after transplant) were "frequently disregarded; these patients were reported as 'unevaluable' because they 'did not survive long enough to exhibit a clinical response.' In many cases, omitting these patients led to higher response rates. Eliminating early deaths is inappropriate (particularly when they may have been caused by the treatment) and not standard for trial design or analysis." (ECRI 1995: 7).

Other methodological problems were rampant. Some studies lacked controls entirely, others included only those patients who had already shown they were responsive to chemotherapy, and still other studies neglected to keep track of key patient characteristics, such as the number of metastatic sites or estrogen receptor status (*ibid.*).

5. *Harris v. Mutual of Omaha Companies*, 992 F.2d 706 (7th Cir. 1993); *Sweeney v. Gerber Products Co. Medical Benefits Plan*, 728 F. Supp. 594 (D. Neb. 1989); *Thomas v. Gulf Health Plan, Inc.*, 688 F. Supp. 590 (S.D. Ala. 1988); *Arrington v. Group Hospitalization and Medical Service*, 806 F. Supp. 287 (D.D.C. 1992); *Fuja v. Benefit Trust Life Insurance Co.*, 18 F.3d 1405 (7th Cir. 1994); *Bechtold v. Physicians Health Plan*, 19 F.3d 322 (7th Cir. 1994); *Holder v. Prudential Insurance Co. of America*, 951 F.2d 89 (5th Cir. 1992); *Graham v. Medical Mutual of Ohio*, 130 F.3d 293 (7th Cir. 1997); *Turner v. Fallon Community Health Plan, Inc.*, 127 F.3d 196 (1st Cir. 1997).

6. *Pirozzi v. Blue Cross-Blue Shield of Virginia*, 741 F. Supp. 586 (E.D. Va. 1990); *Wilson v. Group Hospitalization*, 791 F. Supp. 309 (D.D.C. 1992); *Calhoun v. Complete Health Care, Inc.*, 860 F. Supp. 1494 (S.C. Ala. 1994); *Henderson v. Bodine Aluminum, Inc.*, 70 F.3d 958 (8th Cir.

412 Journal of Health Politics, Policy and Law

injunctions mandating insurance coverage, wrongful death verdicts imposing enormous damages,⁷ insurance companies' acquiescence to threats of litigation, and government mandates to cover the procedure (Hoffman 1999), the treatment proliferated rapidly (Peters and Rogers 1994). Indeed, although the National Institutes of Health (NIH) had major research under way, results were exceedingly slow in coming. Because so many women had access to the treatment through their insurers, it became difficult to recruit enough women willing to enter controlled scientific trials in which only half the subjects would receive the treatment.⁸ When the NIH studies finally concluded, results indicated that HDC/ABMT had no significant advantage over standard chemotherapy.⁹ By that time, some 30,000 women had received the treatment, at a cost estimated around \$3 billion.¹⁰ This figure does not count what some health plans paid in compensatory and punitive damages, or in legal fees and courts costs, for making coverage denials that turned out, in fact, to be correct.

1995); Dahl-Eimers v. Mutual of Omaha Life Insurance Co., 986 F.2d 1379 (11th Cir. 1993); Wilson v. CHAMPUS, 65 F.3d 361 (4th Cir. 1995); Bucci v. Blue Cross-Blue Shield of Connecticut, 764 F. Supp. 728 (D. Conn. 1991); Mattive v. Healthsource of Savannah, Inc., 893 F. Supp. 1559 (S.D. Ga. 1995); Killian v. Healthsource Provident Administrators, 152 F.3d 514 (6th Cir. 1998); Bailey v. Blue Cross/Blue Shield of Virginia, 67 F.3d 53 (4th Cir. 1995).

7. See, e.g., *Fox v. HealthNet*, No. 219692, 1993 WL 794305 (Riverside County Super. Ct./Central Cal. Dec. 23, 1993), discussed in Furrow 1997: 447; Meyer and Murr 1994; *Newsweek* 1994 10 Jan. 123(2):36–38.

8. As noted in one commentary, problems arose in efforts to recruit women for trials designed to assess the benefit of HDC/ABMT in the treatment of metastatic breast cancer. Because of the availability of HDC/ABMT outside of clinical trials, many women with metastatic breast cancer were not willing to accept the chance of being randomized to a control group in a trial designed to evaluate the effectiveness of HDC/ABMT. As a result, it took much longer than expected to obtain an adequate number of participants in these studies to resolve the uncertainty over the value of this technology. Steinberg, Tunis, and Shapiro 1995: 150. See also Kolata 1995; Kolata and Eichenwald 1999: A-1.

9. Of five studies released in 1999, four indicated that high-dose chemotherapy with bone marrow transplant was no better for breast cancer than conventional chemotherapy. A fifth study, done in South Africa, suggested some benefit. However, several months later, as scientists looked at this study more closely in an effort to replicate its results, the principal investigator admitted to having falsified some of the data "out of a foolish desire to make the presentation more acceptable" to the scientific meeting sponsored by the American Society of Clinical Oncology. Weiss et al. 2000: 1003. See also Antman, Heitjan, and Hortobagyi 1999: 1701–1703; Gradishar 1999: 1378–1380; Rowlings et al. 1999: 1335–1343; Horton 2000: 942–943; Bergh 2000: 944–945.

Another study completed even more recently reached the same conclusion, namely, that bone marrow transplant offers no advantage over conventional chemotherapy. See Stadtmauer et al. 2000: 1069–1076; Lippman 2000: 1119–1120.

Interestingly, the wide availability of insurance coverage effectively precluded completion of scientific trials on ABMT for ovarian cancer. See Kolata and Eichenwald 1999: A-1.

10. Associated Press 2000: A-5. It might also be noted that during the heyday of using ABMT for breast cancer, many hospitals and physicians made enormous sums of money from the treatment. See Kolata and Eichenwald 1999: A-1.

Such desperation- or sympathy-guided rulings are not merely expensive. They set a terrible legal precedent if we want empirical judgments to be guided by empirical evidence.¹¹ And yet such judicial aberrations from empirical realities are not unique.¹² Indeed, comparable cases from product liability and toxic tort litigation prompted the Supreme Court's mandate that judges screen empirical testimony more rigorously (*Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 43 F.3d 1311 [9th Cir. 1995]). Accordingly, it is appropriate to consider carefully what the approach of the courts ought to be regarding the uses of scientific evidence in litigation regarding health plans and providers alike. Health plans will be considered first, then physicians.

The Role of Evidence in Judging Health Plans

An Important Distinction

We begin with a distinction. When courts consider whether a health plan committed a tort or breached a contract in its attempts to trim costs and reshape clinical care, only some of the issues are empirical—that is, only some issues will be resolvable by appeal to the evidence found through sensory observation and experience. In the HDC/ABMT example, plans' claim that the therapy had little scientific support for treatment of breast cancer was an empirical claim. So are claims about whether or not mothers and infants experience greater mortality and morbidity when discharged within twenty-four hours after an uncomplicated vaginal delivery, or claims about the comparable effectiveness of generic versus brand-name drugs in the treatment of this or that infectious organism.

In contrast, a health plan makes normative claims when it determines, for example, whether certain expenditures produce enough benefit to

11. The issue is actually more complicated than this, because it is not simply a matter of courts permitting bad science (or no science) to dictate medical standards. As discussed below, the matter is complicated by health plan contracts that actually look for physician acceptance as their criterion of medical necessity.

12. In a fairly well-known phenomenon of which the ABMT episode is but one example, in "judge-made insurance" courts require payers and sometimes even physicians to provide even very high-cost, relatively unproven technologies to patients who otherwise lack much hope for survival. See Abraham 1981: 1155; Ferguson, Dubinsky, and Kirsch 1993: 2116; Kalb 1990: 1118; Morreim 1995: 251.

In another context, it has been observed that courts ruling on managed care cases can be remarkably ill-informed about the economic arrangements and financial structures of the health plans whose actions they judge, with the result that important decisions are sometimes based on badly misinformed factual premises. See Morreim 2000a: 699–728.

414 Journal of Health Politics, Policy and Law

warrant their cost. A plan might agree, empirically, that annual mammography for women under forty provides some benefit. And yet it might decide, normatively, that this benefit does not merit funding, given the more pressing alternative needs for that plan's limited funds in serving its large population (Eddy 1994). These value decisions are sometimes explicitly embedded in contractual terms of coverage and exclusions, but they can also be implicit in individual coverage decisions.

A third distinction would note that still other claims are conceptual. Health plans' decisions are heavily based on contractual provisions, and those provisions' terms must be interpreted. A plan may exclude coverage for "custodial care," for instance, but in a given instance it may require careful interpretation of linguistic concepts to decide whether a patient's extended home care counts as "medical treatment" or merely "custodial" care. This third distinction, while important, will not be discussed further here.

The significance of distinguishing empirical from normative issues is that, as courts consider various challenges to health plans' decisions about care and coverage, they must determine what sort of issue is at stake, and bring the right sort of evaluation to it. Courts cannot resolve normative issues by gathering empirical evidence, nor vice versa. We begin, then, by discussing courts' approach to the empirical issues, before turning to normative issues just below.

Empirical Dimension

If courts expect plans and providers to base their empirical decisions on more and better evidence, health plans have a formidable task. Outcomes studies attempting to document the actual effects of ordinary clinical care are a relatively new phenomenon. During the post-World War II era of lavish third-party health insurance, medical science focused mainly on the development and testing of high-technology new drugs and devices. There was little reason to evaluate new products' and procedures' best uses, or even their most efficient production modes because, so long as FDA approval plus professional acceptance ensured good sales, it would be foolish for manufacturers to do research that could ultimately reduce sales (Garber 1992). By the same token, fee-for-service rewarded physicians and hospitals for maximizing services, not for studying which ones to delete.

Only recently has an urgent need to cut costs and maximize value-for-dollars prompted serious attempts to connect inputs with outcomes and

to identify the most effective, and cost-effective, modes of care. However, although thousands of clinical practice guidelines (CPGs) have proliferated in recent years,¹³ many have at best only a limited scientific basis. The problems, detailed elsewhere,¹⁴ include a dearth of studies, inadequacy of databases, unstandardized methodologies, and biases and conflicts of interest. Nevertheless, outcomes research and health technology assessment (HTA) have become crucial to intelligent health care planning, and the quality of such research and the guidelines to which it gives rise are improving steadily.

If courts are to bring *Daubert* standards to evaluate the adequacy of the guidelines by which plans shape clinicians' care and make their coverage decisions, those CPGs should be anchored in "a reliable foundation" (*Daubert*, 509 U.S. at 597) not just the vague "general acceptance" of the standard set by *Frye v. United States* (293 F. 1013 [D.C. Cir. 1923], as discussed by Shuman in this issue). Hence, if a plan's CPG says "patients with condition X should generally be hospitalized only two days," there should be a credible empirical basis for choosing two days rather than some other number. Indeed, several courts have already held that plans' guidelines, utilization review programs, and coverage decisions must be made on a medically reasonable basis.¹⁵

Two caveats should be noted. First, in addition to their empirical bases, such choices will also reflect value judgments about where best to draw the lines between benefits and costs, as noted below under "normative" considerations. Lavishly funded plans will naturally have more liberal CPGs, while leaner ones will be less generous. Second, plans should not be required to use the "best" empirical evidence nor, as Daniel W. Shuman observes in his article here, should courts preclude differing schools of thought, reputable minorities, or the other kinds of allowance already permitted when courts appraise individual physicians' practices. Indeed,

13. For examples of guidelines, see the following Web sites: www.clinicalevidence.org (Clinical Evidence, published by *British Medical Journal*); www.guideline.gov (National Guideline Clearinghouse); www.ama-assn.org (Clinical Practice Guidelines Directory); 1ww.com (U.S. Preventive Services Task Force Guide to Clinical Preventive Services).

14. See Cynthia D. Mulrow and Kathleen N. Lohr's article in this issue. See also Morreim 1999: 983–989; Morreim 2000c: 65–70.

15. In *Shannon v. McNulty* (718 A.2d. 828, 835 [Pa. Sup. Ct. 1998]), a Pennsylvania superior court noted that "these decisions may, among others, limit the length of hospital stays, restrict the use of specialists, prohibit or limit post-hospital care, restrict access to therapy, or prevent rendering of emergency room care. While all of these efforts are for the laudatory purpose of containing health care costs, when decisions are made to limit a subscriber's access to treatment, that decision *must pass the test of medical reasonableness*" (emphasis added). See also *Crum v. Health Alliance Midwest, Inc.*, 47 F. Supp. 2d 1013 (C.D. Ill. 1999); *Wickline v. State of California*, 192 Cal. App. 3d 1630, 1645 (1987);

the very same dearth of research, of adequate databases, and of standardized methodologies that makes outcomes research so difficult would arguably require considerable flexibility from courts examining a CPG's validity.

Perhaps even more importantly, if courts do not leave reasonable room for differences of opinion, then the courts themselves would potentially engage in the practice of medicine by dictating too closely to health plans which clinical guidelines to adopt. If courts today feel ill-prepared even to assess which testimony is sufficiently "expert" to admit regarding toxic torts and products liability, they would quite surely be unprepared to dictate the nation's medical standards by permitting too narrow a range of CPGs to wear the mantle of judicial acceptability.

Nevertheless, a *Daubert* standard applied to CPGs would probably demand a better scientific pedigree than some health plans' guidelines currently appear to offer. According to some observers, "Most health insurers and managed care plans rely on ad hoc opinion by experts; only in a few instances are there HTA programs or structured processes for coverage decision making" (Perry and Thamer 1999: 1870). Moreover, "materials such as the practice guidelines prepared by Milliman and Robertson, a well-known actuarial firm, often rely on insurers' own decisions rather than on well-designed scientific research" (Rosenbaum et al. 1999: 231).¹⁶ In other cases, plans have relied on "an administrator who 'asked friends who are doctors,' or an insurance company's employee-physician (usually not a specialist in the field in question) who reads textbooks and discusses the issue with other insurance company physicians" (Holder 1994: 19).

In this context, many health plans have created a serious problem for themselves by defining "medical necessity"—the contractual cornerstone criterion of most health plans' coverage (Havighurst 1995: 15; Hall et al. 1996: 1055)—in terms of physician acceptance or general recognition by the medical profession.¹⁷ If health plans want to insist that physi-

16. Recently two physicians filed suit against Milliman and Robertson. "Two pediatricians, whom M&R cites as the authors of its pediatric hospitalization guidelines, allege not only that they did not write the guides, but that the guides are dangerous and inappropriate for pediatric care." See Page 2000: 1, 29, 35.

17. See, e.g., *McGraw vs. Prudential Insurance Co. of America*, 137 F.3d 1253, 1256 (10th Cir. 1998) (citing Prudential's definition of medical necessity partly in terms of being "recognized throughout the Doctor's profession as safe and effective"); *Healthcare America Plans, Inc., v. Bossemeyer*, 953 F. Supp. 1176, 1188 (D. Kan. 1996), affirmed without opinion, 166 F.3d 347 (10th Cir. 1998) (defining "experimental" treatments as those "not generally accepted by the medical community"); *Miller v. Whitburn*, 10 F.3d 1315, 1320 (7th Cir. 1993) (discussing Medicare's definition of medical necessity in terms of "whether the service has come to be gen-

cians base their practices on scientific evidence rather than on local habits, malpractice fears, facilities availability, and the like, then plans must rewrite their contracts to reflect that outlook. So long as “medical necessity” definitions hinge largely on physician acceptance, then the only “evidence” even the most *Daubert*-loyal court can demand will be evidence of a practice’s popularity among physicians, not its grounding in science.

Normative Dimension

As noted just above, it is one thing for a health plan to make an empirical claim—for example, that medication generally works better than surgery for certain patients’ heart problems. It is quite another matter for plans to make normative judgments, such as whether the marginal improvement in quality of life that comes from a new drug (with once-a-day convenience and fewer side effects) is worth paying considerably higher cost. Such value judgments permeate CPGs, right alongside scientific elements. Thus science can tell us that computed tomography (CT) will not harm patients and will occasionally detect a hitherto undiagnosed problem. Cost-effectiveness analysis (CEA) can further tell us how much each new CT-made diagnosis has cost and compare that with the costs and burdens of alternative approaches. But value judgments are required for a health plan to conclude that it will not cover annual head-to-toe CT scans (Barnard 2000: A-20) for all its enrollees as a preventive care routine, on the ground that this cost is too high for its anticipated benefits and that the money can be better spent elsewhere.

There are many ways to draw the myriad cost-value trade-offs that permeate health care, and each has its merits and drawbacks. As Justice David Souter points out in *Pegram v. Herdrich* (120 S Ct. 2143 [2000]), a cost-conscious plan is more likely to witness ruptured appendixes from too few appendectomy surgeries, while an open-ended fee-for-service approach is more likely to see unnecessary appendectomies. Health plans must constantly weigh the benefits that an expenditure will bring to a few individuals against its needs to serve all its other enrollees. These deci-

erally accepted by the professional medical community as an effective and proven treatment for the condition for which it is being used” [citing *Rush v. Parham*, 625 F.2d 1150 (5th Cir. 1980)]; Horvath 1999: 455 (citing the National Blue Cross and Blue Shield Association’s definition of “experimental and investigational treatment” partly in terms of whether it is generally “recognized by the medical profession as tested and accepted medical practice”).

418 Journal of Health Politics, Policy and Law

sions in turn are set within the limits of both the available premium dollars and the uncertainties of future needs.

Health plans' values choices are not always obvious. As Jacobson and Kanna note in their essay in this issue, plans tend to keep cost-value trade-offs "below the radar screen." Commonly they are hidden in plans' judgments about medical necessity. The term sounds precise enough, with tones of science and imperatives of necessity. But in fact there are enormous variations in how plans determine what is "necessary" and what isn't, even though virtually all plans promise to cover all, but only, necessary care.¹⁸

Courts could manage these tensions in a variety of ways. One option, as Jacobson and Kanna note, is Judge Hand's formula in *Carroll Towing*, balancing the probability that a given adverse event will happen, the gravity of injuries that might cause, and the burdens of preventing them (United States v. Carroll Towing, 159 F.2d 169, 173 [2d Cir. 1947]). However, it would be difficult to mandate this or any other one-size-fits-all approach. Health care is marked by an extraordinary diversity of goals and values, a diversity rooted partly in the fact that most of day-to-day health care is geared not toward life and death and the catastrophes of *Carroll Towing* but toward quality of life, the management of uncertainty, and other highly nuanced, deeply personal factors. There are many ways besides health care for citizens to improve or preserve their quality of life, and a variety of legitimate ways to decide what price is worth paying to reduce what kinds of uncertainty (Morreim 2000b).

Accordingly, in deference to the wide diversity of human goals and budgets, courts arguably should permit health plans to vary in the kinds and levels of coverage they provide. But once that is granted, we must ask afresh how courts should evaluate disputes about the value choices a health plan has implemented through its decisions about care and coverage.

The answer must begin with the observation that in this normative sphere, courts cannot demand scientific evidence, because this is the realm of values not facts. No amount of data can tell us the scientifically

18. For an extended discussion of the ways in which the concept of medical necessity figures in health care contracts, and the problems inherent in that approach, see Morreim 1999: 1010–1025; Morreim 2000b: 144–158. Even in Medicare plans that ostensibly provide the same benefits package to everyone, a recent study showed that Medicare payment for a chest Xray was 451 times more likely to be denied in Illinois than in South Carolina, and payment for real-time echocardiography was nearly one hundred times more likely to be denied by Transamerica Occidental than by Blue Shield of California (Pretzer 1995: 92–93). See also Gleason 1995: 1483; Mariner 1995: 236–246.

“correct” priority to place on Viagra, or how much money should be spent to reduce someone’s chance of fatal heart attack by 1 percent. Hence *Daubert*, with its focus on adequacy of empirical claims, does not apply to this normative realm.

Instead, courts might inquire whether the health plan has made its values clear—whether it informed its subscribers adequately, up front, just how it allocates resources, and whether it has acted in accordance with those stated values. A clear contract will tell prospective buyers “if you buy this plan, here is what you’ll receive, and here are the rules by which we will decide the borderline cases”: “yes, the evidence shows that treatment T works well enough, but in this plan T is not deemed sufficiently cost-effective to merit coverage.”¹⁹ This is a contract-focused approach, but not the same sort that Jacobson and Kanna identify in discussing contract-based approaches to CEA. As they describe the option of “abandon tort altogether in favor of contract” (Jacobson and Kanna this issue), courts would have little or no opportunity to evaluate the quality of the evidence on which plans base the empirical aspects of their CPGs. They would simply determine whether the plan followed its contract.

In contrast, the alternative approach suggested here would have courts address empirical disputes by considering the quality of the empirical evidence. *Daubert*-based challenges would be appropriate if, for instance, a plan constructed its CPGs, or applied them to individual cases, in a scientifically unacceptable manner. Reciprocally, courts would address normative disputes primarily as a contract issue by inquiring whether the health plan made its values clear and acted reasonably in accordance with them in the instant case. The important point for present purposes is that we must not confuse plans’ empirical claims about which interventions work best for which problems, with their value choices about what is worth paying for. The former, but not the latter, permits evidence-oriented scrutiny.

Fairness Issues

As courts determine how best to hold health plans accountable, they should avoid unfairly holding plans to a stricter standard than physicians, in cases where comparable questions are at stake. Physicians, after all, are not expected to achieve perfect results or even to exercise optimal skill and judgment. They are obligated only to provide ordinary and rea-

19. For further discussion of guidelines-based contracting, see Morreim 1999: 983–989; Morreim 2000c: 144–158.

sonable care. If patients nevertheless do poorly, that is not malpractice but only an unfortunate outcome.

Health plans likewise should not be judged by whether a patient has fared well or poorly. Rather, they too should be judged according to the adequacy of their policies and implementation. Plans could be easy targets otherwise, to the detriment of their fiscal stability and ultimately their patients' care. For example, many health plans now cover screening tests for early detection of major illnesses. Whether the test is mammography for breast cancer or PSA for prostate cancer, no matter what frequency of testing is chosen, some patients' illness will be missed (Leahy 1989). Those patients might correctly say, "If only the policy hadn't been so stingy, my illness would have been diagnosed at a more curable stage." But just as courts do not judge physicians simply on the basis of whether some other treatment might have worked, neither should they judge plans according to whether some other policy would have averted this patient's unfortunate outcome. Health plans must create policies to meet a broad diversity of needs, and for courts the important question should be whether the policy was well-conceived, adequately disclosed, and properly applied *ex ante*, not whether it served every person well *ex post*. As Arnold J. Rosoff notes in his article in this issue, plans must take a prospective viewpoint as they arrange for many people to receive good care within a budget. Although courts must resolve adverse incidents in retrospect, the proper approach for assessing the policies that lead to those incidents is essentially prospective.

The Role of Evidence in Judging Physicians

The movement toward evidence-based medicine is a bit overdue. As noted above, until quite recently scientific research has focused mainly on costly new technologies, to the relative neglect of ordinary care and its outcomes. And even where good studies are available to guide care, physicians' clinical practices have sometimes been woefully out of step.²⁰ Underuse, overuse, and misuse of diagnostic and therapeutic modalities are now recognized as a major problem throughout medicine (Chassin and Galvin 1998), as are avoidable errors that lead to morbidity and mortality (Institute of Medicine 1999). Even care whose well-documented value should be obvious, like aspirin after a heart attack, is often neglected (Newcomer 1998; Donohoe 1998). Thus the bare fact that physicians

20. For a detailed discussion of this point with numerous cites, see Morreim 1999: 989–998.

commonly practice or accept a particular pattern of care unfortunately tells us too little about whether that pattern is salutary for patients, let alone affordable for plans. As noted by Havighurst et al., “courts’ continued focus on professional custom for setting the legally recognized standard of care reinforces this disregard of scientific progress” (see the Introduction to this special issue by Clark C. Havighurst, et al.).

The poor correlation between what physicians actually do, and the emerging evidence showing what they should do, is just one of several reasons why courts should no longer determine what physicians ought to do by looking primarily at what they actually do. In addition, many prevailing practices often do not even reflect what physicians believe they should do. Defensive medicine, patients’ demands for the latest technologies, physicians’ own fascination with new devices and procedures, and economic pressures can quickly skew physicians’ most thoughtful, scientifically well-founded preferences. Beyond this, physicians are neither morally nor financially entitled to do literally whatever they ideally wish to do for their patients. Health care is costly, and the days of handing physicians a blank check to spend unlimited amounts of other people’s money are over (Morreim 1997: 16–27).

Accordingly, the courts should move away from their tradition of setting physicians’ professional standard of care by referring to their prevailing practices—even with the usual caveats for differing schools of thought, reputable minorities, and the like. Instead of asking for evidence about what members of the profession (say they) do, courts should begin, just as they should with health plans, to expect evidence of what actually works for patients. Indeed, this expectation of reliable scientific evidence rather than “general practice or acceptance” is precisely the transition that the Supreme Court has asked courts to make, in the move from *Frye* to *Daubert*, as Shuman indicates in this issue. Although this move was not entirely feasible for health care only a few years ago, the emergence and improving methodologic quality of outcomes research now makes such a transition increasingly plausible and desirable. Hence if, per Shuman’s analysis, courts have been reluctant to bring *Daubert* standards to medical litigation, it is time to begin that admittedly difficult transition.

This is not to say that courts should look exclusively at scientific evidence in judging physicians’ performance. For one thing, it was noted above that outcomes research and technology assessment, and the guidelines to which they sometimes lead, still need considerable improvement and expansion, and they must be kept up to date as technologies and information change (a daunting if not virtually impossible task). Hence,

422 Journal of Health Politics, Policy and Law

flexibility is at least as important here as in the assessment of health plans' performance. For another thing, even the best CPGs cannot possibly dictate each patient's course of care. They are based on generalities that hold true on average, but have only limited room to accommodate the natural variations among individuals in any population.²¹ Moreover, they do not accommodate important patient preferences that can thwart even the best treatment plans. Thus good medicine unavoidably requires sophisticated professional evaluations of patients' signs and symptoms, and careful judgments to determine whether this patient's situation is sufficiently typical to warrant following rather than deviating from the applicable CPG.

The upshot of these observations is that, while courts should expect various CPGs to be defended as scientifically reasonable, even though not demanding lockstep adherence to any particular CPG, nevertheless they should expect that, if a medically reasonable CPG suggests conduct from which the physician has deviated, the physician should be able to explain that deviation with something more than a flat assertion that "in my professional judgment, the guideline did not apply."

Judges will face a major challenge at this point. As noted above, CPGs abound, many of them with dubious scientific credentials, and physicians may often be right to deviate from them. At the same time, emerging outcomes research suggests that many of physicians' current clinical routines (and therefore, quite likely, many of their deviations from the CPGs that aim to improve those routines) are not scientifically adequate either. Judges will thus have a difficult time sorting out whose evidence is more credible. Fortunately, *Daubert* and progeny do not ask judges to determine which side is right. Juries ordinarily take a case from evidence to answers. But the task of screening the evidence will be daunting enough, as the authors in this collection have shown so well.

21. Indeed, the most pristine kind of science—the randomized, double-blind, controlled trial—can sometimes be least applicable in the clinical setting. In order to test strictly for the effects of the specific drug or procedure under investigation, scientific study design must be restricted to patients fitting a narrow set of eligibility criteria. Typically they must suffer exclusively from the particular disease whose treatment is being studied, with a minimum of other diseases or medications that could confound the results. Once the study is complete, however, its results are applied in clinical practice to all those complex patients who would never have been eligible to be test subjects in the study.

One result of this misfit between the original study's narrowly identified participants, and the ordinary folk who later use the drug, is that sometimes even well-researched new drugs and procedures must be quickly withdrawn from the market, as they suddenly produce undesirable results and side effects that were not seen during the research period. Between September 1997 and September 1998, five FDA-approved drugs were removed from the market because of unexpected side effects or interactions with other drugs. Friedman et al. 1999: 1729.

Conclusion

Medicine is in the midst of an extraordinary transition. Economic pressures, the quest for quality, and a host of other factors have finally begun to force medicine to become truer to the scientific roots it has always claimed to have. But that transition is also an upheaval. Cherished assumptions and long-entrenched habits have been castigated or supplanted, though not always with improved alternatives. And the upheaval will undoubtedly continue for years to come. If it is a good thing to bring better science to health care, then surely it will also be a good thing, at least in principle, to bring *Daubert's* demand for greater rigor to the legal proceedings that sort out the consequences of health care and, in many ways, shape the future of health care itself. The move toward greater rigor both in the courts and in the clinic will be difficult but ultimately should bring major improvements for the care of patients and for the fair treatment of litigants.

References

- Abraham, K. S. 1981. Judge-Made Law and Judge-Made Insurance: Honoring the Reasonable Expectations of the Insured. *Virginia Law Review* 67:1151–1191.
- Angell, M. 1996. *Science on Trial: The Clash of Medical Evidence and the Law in the Breast Implant Case*. New York: W.W. Norton.
- Antman, K. H., D. F. Heitjan, and G. N. Hortobagyi. 1999. High-Dose Chemotherapy for Breast Cancer. *Journal of the American Medical Association* 282:1701–1703.
- Associated Press. 2000. Breast Cancer Procedure Bogus. *Memphis Commercial Appeal*, 11 March, A-5.
- Avanian, J. Z., M. B. Landrum, S. T. Normand, E. Guadagnoli, and B. McNeil. 1998. Rating the Appropriateness of Coronary Angiography—Do Practicing Physicians Agree with an Expert Panel and with Each Other? *New England Journal of Medicine* 338:1896–1904.
- Avorn, J., M. Chen, and R. Hartley. 1982. Scientific versus Commercial Sources of Influence on the Prescribing Behavior of Physicians. *American Journal of Medicine* 73:4–8.
- Barnard, A. 2000. Scans Could Allay Health Fears, or Spark New Ones. *Wall Street Journal*, 30 September, A-20.
- Bergh, J. 2000. Where Next with Stem-Cell-Supported High-Dose Therapy for Breast Cancer? *Lancet* 355:944–945.
- Bodenheimer, T., and K. Sullivan. 1998. How Large Employers Are Shaping the Health Care Marketplace (first of two parts). *New England Journal of Medicine* 338:1003–1007.

424 Journal of Health Politics, Policy and Law

- Burnum, J. F. 1987. Medical Practice a la Mode. *New England Journal of Medicine* 317:1220–1222.
- Chassin, M. R., and R. W. Galvin, 1998. National Roundtable on Health Care Quality. The Urgent Need to Improve Health Care Quality. *Journal of the American Medical Association* 280:1000–1005.
- Donohoe, M. T. 1998. Comparing Generalist and Specialty Care: Discrepancies, Deficiencies, and Excesses. *Archives of Internal Medicine* 158:1596–1608.
- ECRI. 1995. High-Dose Chemotherapy with Autologous Bone Marrow Transplantation and/or Blood Cell Transplantation for the Treatment of Metastatic Breast Cancer. Healthy Technology Assessment Information Service, Executive Briefings, February.
- Eddy, D. M. 1994. Rationing Resources while Improving Quality: How to Get More for Less. *Journal of the American Medical Association* 272:817–824.
- Ferguson, J. H., M. Dubinsky, and P. J. Kirsch. 1993. Court-Ordered Reimbursement for Unproven Medical Technology: Circumventing Technology Assessment. *Journal of the American Medical Association* 269:2116–2121.
- Friedman, M. A., J. Woodcock, M. M. Lumpkin, J. E. Shuren, A. E. Hass, and L. J. Thompson. 1999. The Safety of Newly Approved Medicines: Do Recent Market Removals Mean There Is a Problem? *Journal of the American Medical Association* 281:1728–1734.
- Furrow, B. R. 1997. Managed Care Organizations and Patient Injury: Rethinking Liability. *Georgia Law Review* 31:419–509.
- Garber, A. M. 1992. No Price Too High? *New England Journal of Medicine* 327:1676–1678.
- Gleason, S. C. 1995. Health System Deregulation: Some Aspects of Health Care System Reform Need Not Be Held Hostage. *Journal of the American Medical Association* 274:1483–1486.
- Gradishar, W. J. 1999. High-Dose Chemotherapy and Breast Cancer. *Journal of the American Medical Association* 282:1378–1380.
- Hall, M. A., T. R. Smith, M. Naughton, and A. Ebbers. 1996. Judicial Protection of Managed Care Consumers: An Empirical Study of Insurance Coverage Disputes. *Seton Hall Law Review* 26:1055–1068.
- Hardison, J. E. 1979. To be complete. *New England Journal of Medicine* 300:193.
- Harris, H. M. 1996. Disease Management: New Wine in New Bottles? *Annals of Internal Medicine* 124:838–842.
- Hartert, T. V., H. H. Windom, S. Peebles Jr., L. R. Freidhoff, and A. Togias. 1996. Inadequate Outpatient Medical Therapy for Patient with Asthma Admitted to Two Urban Hospitals. *American Journal of Medicine* 100:386–394.
- Havighurst, C. C. 1995. *Health Care Choices: Private Contracts as Instruments of Health Reform*. Washington, DC: AEI.
- Hoffman, S. 1999. A Proposal for Federal Legislation to Address Health Insurance Coverage for Experimental and Investigational Treatments. *Oregon Law Review* 78(1):203–274.
- Holder, A. R. 1994. Medical Insurance Payments and Patients Involved in Research. *IRB* 16 (1–2):19–22.

- Holoweiko, M. 1995. When an Insurer Calls Your Treatment Experimental. *Medical Economics* 72(17):171–182.
- Horton, R. 2000. After Bezwoda. *Lancet* 355:942–943.
- Horvath, K. A. 1999. Infertility Treatment: An Argument for Mandated Coverage. *Journal of Health Law* 32(3):445–474.
- Huber, P. S. W. 1991. *Galileo's Revenge: Junk Science in the Courtroom*. New York: Basic.
- Institute of Medicine. 1999. *To Err Is Human: Building a Safer Health System*. Washington, DC: National Academy.
- Kalb, P. E. 1990. Controlling Health Care Costs by Controlling Technology: A Private Contractual Approach. *Yale Law Journal* 99:1109–1126.
- Kolata, G. 1995. Women Rejecting Trials for Testing a Cancer Therapy. *New York Times*, 15 February, C8.
- Kolata, G., and K. Eichenwald. 1999. Hope for Sale: Business Thrives on Unproven Care, Leaving Science Behind. *New York Times*, 3 October, A-1.
- Lange, R. A., and L. D. Hillis. 1998. Use and Overuse of Angiography and Revascularization for Acute Coronary Syndromes. *New England Journal of Medicine* 338:1838–1839.
- Leahy, R. E. 1989. Rational Health Policy and the Legal Standard of Care: A Call for Judicial Deference to Medical Practice Guidelines. *California Law Review* 77:1483–1528.
- Leape, L. L. 1995. Translating Medical Science into Medical Practice: Do We Need a National Medical Standards Board? *Journal of the American Medical Association* 273:1534–1537.
- Legorreta, A. P., J. Christian-Herman, R. D. O'Connor, M. M. Hasan, R. Evans, and K. M. Leung. 1998. Compliance with National Asthma Management Guidelines and Specialty Care. *Archives of Internal Medicine* 158:457–464.
- Lippman, M. E. 2000. High-Dose Chemotherapy Plus Autologous Bone Marrow Transplantation for Metastatic Breast Cancer. *New England Journal of Medicine* 342:1119–1120.
- Mariner, W. K. 1995. Business vs. Medical Ethics: Conflicting Standards for Managed Care. *Journal of Law, Medicine, and Ethics* 23:236–246.
- Meyer, M., and A. Murr. 1994. Not My Health Care. *Newsweek*, 10 January, 123(2):36–38.
- Morreim, E. H. 1994. Redefining Quality by Reassigning Responsibility. *American Journal of Law and Medicine* 20(1–2):79–104.
- . 1995. Moral Justice and Legal Justice in Managed Care: The Ascent of Contributive Justice. *Journal of Law, Medicine, and Ethics* 23:247–265.
- . 1997. Medicine Meets Resource Limits: Restructuring the Legal Standard of Care. *University of Pittsburgh Law Journal* 59(1):1–95.
- . 1999. Playing Doctor: Corporate Medical Practice and Medical Malpractice. *Michigan Journal of Law Reform* 32(4):939–1040.
- . 2000a. Confusion in the Courts: Managed Care Financial Structures and Their Impact on Medical Care. *Tort and Insurance Law Journal* 35(3):699–728.

426 Journal of Health Politics, Policy and Law

- . 2000b. Quality of Life: Erosions and Opportunities under Managed Care. *Journal of Law, Medicine, and Ethics* 28:144–158.
- . 2000c. Saving Lives, Spending Money: Shepherding the Role of Technology. In *Ethical Issues in Health Care on the Frontiers of the Twenty-First Century*, ed. S. Wear, J. J. Bono, G. Logue, and A. McEvoy. Dordrecht: Kluwer Academic, 65–112.
- Moser, M. 1998. Why Are Physicians Not Prescribing Diuretics More Frequently in the Management of Hypertension? *Journal of the American Medical Association* 279:1813–1816.
- Newcomer, L. N. 1998. Physician, Measure Thyself. *Health Affairs* 17(4):32–35.
- Page, L. 2000. Lawsuit Puts Spotlight on Hospital Discharge Criteria. *American Medical News*, 27 March, 1, 29, 35.
- Perry, S., and M. Thamer. 1999. Medical Innovation and the Critical Role of Health Technology Assessment. *Journal of the American Medical Association* 282:1869–1872.
- Peters, W. P., and M. C. Rogers. 1994. Variation in Approval by Insurance Companies of Coverage for Autologous Bone Marrow Transplantation for Breast Cancer. *New England Journal of Medicine* 330:473–477.
- Pretzer, M. 1995. Hate Those Medicare Denials? Try Moving. *Medical Economics* 72(7):92–101.
- Reuben, D. B. 1984. Learning Diagnostic Restraint. *New England Journal of Medicine* 310:591–593.
- Rosenbaum, S., D. M. Frankford, B. Moore, and P. Borzi. 1999. Who Should Determine When Health Care Is Medically Necessary? *New England Journal of Medicine* 340:229–232.
- Rowlings, P. A., S. P. Williams, K. H. Antman, et al. 1999. Factors Correlated with Progression-Free Survival after High-Dose Chemotherapy and Hematopoietic Stem Cell Transplantation for Metastatic Breast Cancer. *Journal of the American Medical Association* 282:1335–1343.
- Schauffler, H. H., C. Brown, and A. Milstein. 1999. Raising the Bar: The Use of Performance Guarantees by the Pacific Business Group on Health. *Health Affairs* 18(2):134–142.
- Stadtmauer, E. A., A. O'Neill, L. J. Goldstein, P. A. Crilley, K. F. Mangan, J. N. Ingle, I. Brodsky, S. Martino, H. M. Lazarus, J. K. Erban, C. Sickles, J. H. Blick, and Philadelphia Bone Marrow Transplant Group. 2000. Conventional-Dose Chemotherapy Compared with High-Dose Chemotherapy Plus Autologous Hematopoietic Stem-Cell Transplantation for Metastatic Breast Cancer. *New England Journal of Medicine* 342:1069–1076.
- Steinberg, E. P., S. Tunis, and D. Shapiro. 1995. Insurance Coverage for Experimental Technologies. *Health Affairs* 14(4):143–158.
- Weiner, J. P., S. T. Parente, D. W. Garnick, et al. 1995. Variation in Office-Based Quality: A Claims-Based Profile of Care Provided to Medicare Patients with Diabetes. *Journal of the American Medical Association* 273:1503–1508.
- Weiss, R. B., R. M. Rifkin, F. M. Stewart, R. L. Thanaule, L. A. Williams, A. A. Her-

- man, and R. A. Beveridge. 2000. High-Dose Chemotherapy for High-Risk Primary Breast Cancer: An On-Site Review of the Bezwoda Study. *Lancet* 355:999–1003.
- Wennberg, J. E., ed. 1996. *The Dartmouth Atlas of Health Care in the United States*. Chicago: American Hospital.
- Wosley, R. L. 1994. Centers for Education and Research in Therapeutics. *Clinical Pharmacological Therapy* 55:249–255.

