Virtual Mentor
American Medical Association Journal of Ethics

January 2013, Volume 15, Number 1: 1-97.
Ethical Issues in Evidence-Based Medicine

Upcoming Issues of *Virtual Mentor*

February: The Hospital: Business or Public Service?
March: Conscience Rights and Responsibilities
April: Effects of Individual Choice on Health Status
May: Reality and Perception in the Management of Pain

Virtual Mentor
American Medical Association Journal of Ethics
January 2013, Volume 15, Number 1: 4-8.

FROM THE EDITOR
Evidence-Based Medicine: A Science of Uncertainty and an Art of Probability

A science of uncertainty and an art of probability [1]—that is how William Osler portrayed medicine as he practiced it at the turn of the last century, but he may as well have been describing the current era of "evidence-based" medicine.

The moniker "evidence-based" made its debut in the early 1990s. As Gordon Guyatt, a physician at McMaster University, first described the term:

> Clinicians were formerly taught to look to authority (whether a textbook, an expert lecturer, or a local senior physician) to resolve issues of patient management. Evidence-based medicine uses additional strategies, including quickly tracking down publications of studies that are directly relevant to the clinical problem, critically appraising these studies, and applying the results of the best studies to the clinical problem at hand [2].

The rest is history. In two decades, the skills laid out by Guyatt have become an integral component of medical practice and training. Today, nearly all U.S. medical schools report teaching evidence-based medicine as part of a required course [3], and the Accreditation Council for Graduate Medical Education (ACGME) has incorporated EBM into U.S. residency training requirements [4].

During this time, the methods of EBM have evolved. Guyatt's first article, for example, used the term "microcomputer" and notes that the cost of retrieving a few citations from MEDLINE was $0.79. Today, the National Library of Medicine makes MEDLINE citations of published research freely available through PubMed; research abstracts in many journals have been restructured for efficient appraisal; and numerous secondary resources, such as *ACP Journal Club* and Dynamed, summarize and review original research for clinical relevance [5].

The ideas underlying EBM have evolved, too. For instance, notions about shared decision making have been refined. Recent discussions about EBM particularize the importance of "decision aids" and other means by which patients participate in their own treatment [6, 7]. A common current definition of EBM is "the integration of best research evidence with clinical expertise and patient values" [8]—a description that seems not so distant from Osler's science and art.

However, EBM is not old hat [9]. Systematic research evidence is more abundant and accessible than ever before, and EBM provides an original framework for

integrating the results of this research into clinical practice [10]. It also proposes new methods to assess how research evidence should be applied in clinical practice, focusing on transparency and explicitness in its interpretation and role in making recommendations [11].

The rationale for EBM seems obvious: if results of clinically relevant research may be readily available and may benefit patients, we should consider such research in making medical decisions [12].

Yet, the practice of EBM is not always so straightforward. The articles in this issue of *Virtual Mentor* illustrate how implementing EBM in clinical practice, policy, and education can be complicated. They shed light on many of EBM's evolving strengths, but also bring into focus the contours and boundaries of this new tool of modern medicine.

Many aspects of EBM concern values in addition to facts. They are intimately linked to customary considerations of medical ethics, including autonomy, justice, beneficence and nonmaleficence, and they raise important questions regarding these principles.

Autonomy: *What should be the patient's role in interpreting and applying research evidence in clinical decisions? Moreover, if the way in which evidence is conveyed to patients can alter their decisions [13], then how should research evidence be communicated?*

In this issue, a thoughtful commentary by Lauris Kaldjian and Paul Christine addresses the complex and interrelated decisions that physicians face in discussing research evidence with patients. Valerie Reyna and Evan Wilhelms describe what has been learned from their research and that of others about effective strategies for communicating risks and benefits.

Thomas LeBlanc reflects on his experiences as an oncologist to address a related question*: Can providing research evidence to a patient be harmful?* Reprinted alongside his contribution is an essay by the late Stephen Jay Gould about making sense of the statistics and probabilities he encountered in his own struggle with cancer.

Justice: *Who should have the authority to prioritize research questions and funding? Whose interests should be considered in the interpretation and dissemination of research evidence?*

Chetan Huded, Jill Rosno, and Vinay Prasad provide an excellent summary of John Ioannidis's essay "Why Most Published Research Findings are False," one of the most cited articles in the history of the journal *PLoS Medicine* [14]. They examine the biases of medicine's "evidence base" and contribute suggestions to rectify these

shortcomings based upon insights from their own research into "medical reversals" [15].

Jodi Halpern and Richard Kravitz offer a discussion on the role of "health advocacy organizations" in the dissemination and interpretation of research evidence and consider the special case when interpretations of research evidence conflict. Joanna Siegel at the Agency for Healthcare Research and Quality (AHRQ) and her colleagues at the American Institute for Research (AIR) outline methods of "public deliberation" to engage the public in decisions about health research. They note a recent study conducted by AHRQ to learn about ways to gather public input about the use of medical evidence in guiding health care practice [16].

Beneficence and nonmaleficence: *Under what conditions, if any, should a physician disregard codified "best evidence" for the benefit of an individual patient? And how should physicians make decisions in the best interests of their patients in the absence of good research evidence?*

Concerns about implementing EBM become particularly pronounced when clinical practice guidelines, quality measures, and reimbursements are linked to scientific research evidence. William Dale and Erika Ramsdale elaborate on how these concerns apply to providing care for older patients, whose health is often particularly complex and difficult to generalize and who are often excluded from participation in research for these reasons. Valarie Blake examines a related issue: the role of clinical practice guidelines in the courts.

Salima Punja and Sunita Vohra contribute a description of "n-of-1 trials," which have been used in clinical practice to obtain patient-specific research evidence. The method was first implemented by Gordon Guyatt and David Sackett in the 1980s [17] and is currently used in clinical services around the world, including one at the University of Alberta headed by Dr. Vohra [18].

Evidence-based medicine is, at its foundation, about medical education—not only does EBM emphasize the role of traditional medical training to disseminate its new methods [19], but it requires all doctors to engage in a continual process of education in order to make use of current research evidence. Therefore, issues surrounding the use of EBM should be the concern of trainees, educators, and anyone else interested in the education of today's physicians.

The contribution to this issue by Ariel Zimerman, a medical historian and physician, traces EBM's roots in medical education at McMaster University. And an essay written by Martha Carvour (who, I should note, taught me much of what I know about evidence-based medicine) provides suggestions for teaching medical students how to think critically about incorporating research evidence into clinical practice. Dien Ho provides a brief account of some ways that "evidence" has been conceptualized by philosophers of science throughout history.

Finally, Ross Upshur's commentary brings to a point the message that readers should take away from the whole of this issue: evidence-based medicine education should be integrated with an education in clinical ethics.
As the articles in this issue make clear, EBM is a powerful tool with the potential to improve clinical decision making and, ultimately, the health of patients. But, as Guyatt and his colleague Victor Montori have noted, EBM can also be dangerous when used inappropriately [7].

Evidence-based medicine can be thought of like a scalpel or a potent drug with possible adverse effects. The effect of an "evidence-based" approach to medicine depends upon accurate and appropriate integration of research evidence into patient care. Using EBM requires precision, attention, and humility.

Like the other tools and techniques of modern medicine's armamentarium, the use of EBM should require deliberate, thoughtful, and mentored experience. This is not a new idea [19]. But it is an important consideration for those entrusted with teaching the next generation of physicians both the science and the art of medicine.

Matthew Rysavy
Medical Scientist Training Program
Carver College of Medicine
University of Iowa
Iowa City, IA

References

1. Bean RB, Bean WB. *Sir William Osler: Aphorisms from his Bedside Teachings and Writings*. New York: H. Schuman; 1950.
2. Guyatt GH. Evidence-based medicine. *ACP J Club*. 1991;114:A-16.
3. Association of American Medical Colleges. Basic science, foundational knowledge, and pre-clerkship content: inclusion of topics in required and/or elective courses. https://www.aamc.org/initiatives/cir/curriculumcontent/310728/topicsinrequiredandelectivecourses.html. Accessed December 1, 2012.
4. Nasca TJ, Philibert I, Brigham T, Flynn TC. The next GME accreditation system--rationale and benefits. *N Engl J Med*. 2012;366(11):1051-1056.
5. Guyatt G, Cook D, Haynes B. Evidence based medicine has come a long way. *BMJ*. 2004;329(7473):990-991.
6. Haynes RB, Devereaux PJ, Guyatt GH. Clinical expertise in the era of evidence-based medicine and patient choice. *ACP Journal Club*. 2002;136(2):A11-A14.
7. Montori VM, Guyatt GH. Progress in evidence-based medicine. *JAMA*. 2008;300(15):1814-1816.
8. Sackett DL, Straus SE, Richardson WS, Rosenberg WM, Haynes RB. *Evidence-based Medicine: How to Practice and Teach EBM*. 2nd ed. Edinburgh: Churchill Livingstone; 2000; 1.

9. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ*. 1996;312(7023):71-72.
10. Norman GR. Examining the assumptions of evidence-based medicine. *J Eval Clin Pract*. 1999;5(2):139-147.
11. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336(7650):924-926.
12. Goodman KW. Comment on M.R. Tonelli, "the challenge of evidence in clinical medicine." *J Eval Clin Pract*. 2010;16(2):390-391.
13. Fagerlin A, Zikmund-Fisher BJ, Ubel PA. Helping patients decide: ten steps to better risk communication. *J Natl Cancer Inst*. 2011;103(19):1436-1443.
14. Freedman DH. Lies, damned lies, and medical science. *The Atlantic*. November 2010. http://www.theatlantic.com/magazine/archive/2010/11/lies-damned-lies-and-medical-science/308269/. Accessed December 12, 2012.
15. Prasad V, Cifu A, Ioannidis JP. Reversals of established medical practices: evidence to abandon ship. *JAMA*. 2012;307(1):37-38.
16. Agency for Healthcare Research and Quality (AHRQ). AHRQ Community Forum. http://effectivehealthcare.ahrq.gov/index.cfm/who-is-involved-in-the-effective-health-care-program1/ahrq-community-forum/. Accessed December 1, 2012.
17. Guyatt G, Sackett D, Taylor DW, Ghong J, Roberts R, Pugsley S. Determining optimal therapy—randomized trials in individual patients. *N Engl J Med*. 1986;314(14):889-892.
18. Kravitz RL, Duan N, Niedzinski EJ, Hay MC, Subramanian SK, Weisner TS. What ever happened to N-of-1 trials? Insiders' perspectives and a look to the future. *Milbank Q*. 2008;86(4):533-555.
19. Evidence-Based Medicine Working Group. Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA*. 1992;268(17):2420-2425.

Virtual Mentor
American Medical Association Journal of Ethics
January 2013, Volume 15, Number 1: 9-17.

ETHICS CASES
Communicating Evidence in Shared Decision Making
Commentary by Paul J. Christine, MPH, and Lauris C. Kaldjian, MD, PhD

Dr. Garrison is a third-year resident in internal medicine who takes pride in helping patients make educated decisions regarding their health care, informing them of recent research and answering their questions.

One afternoon in the outpatient medicine clinic, Dr. Garrison spoke with Mr. Mendez, a 62-year-old man who had come in for a yearly physical. In reviewing Mr. Mendez's chart, Dr. Garrison noted that his total cholesterol was high at 242 mg/dL, while his HDL cholesterol was low at 35 mg/dL. He was otherwise healthy, having normal blood pressure, no other symptoms or signs of cardiovascular disease, no family history of cardiovascular disease, and no history of smoking. According to the Framingham 10-year risk calculator, this picture corresponded to a 15 percent risk of a myocardial infarction in the next 10 years ("moderate risk").

After discussing the findings of the physical exam with Mr. Mendez, Dr. Garrison entered into a discussion of the lab results. "Mr. Mendez, your lab results indicate that you have high cholesterol. Based upon your current state of health and lab results, you have about a 15 percent risk of experiencing a heart attack in the next 10 years. For patients in your situation, we typically recommend treatment with a drug known as a statin, which helps to lower your bad cholesterol." Dr. Garrison proceeded to explain the benefits and side effects of statin therapy to Mr. Mendez, who stated that he was willing to try whatever the doctor thought was best.

Dr. Garrison added, "You should also know that, for the average person with your medical history and state of health, the number needed to treat—that is, the number of individuals who must be treated with a statin to prevent one death from a cardiovascular event such as a heart attack or stroke—is generally between 60 and 100, which means that if I treated 60 people in your situation, 1 would benefit and 59 would not. As these numbers show, it is important for you to know that most of the people who take a statin will not benefit from doing so and, moreover, that statins can have side effects, such as muscle pain, liver damage, and upset stomach, even in people who do not benefit from the medication. I am giving you this information so that you can weigh the risks and benefits and then make an informed decision."

Mr. Mendez looked surprised. He said that he felt fine and, having heard the statistics, did not wish to start statin therapy at this time.

Later that day, Dr. Garrison met a fellow third-year internal medicine resident, Dr. Parra, for coffee in the hospital cafeteria. Discussing some of their cases from the day, Dr. Garrison remarked that several of her patients had elected to try lifestyle modifications rather than initiate drug therapy for hyperlipidemia and some other conditions. "That's remarkable," Dr. Parra observed, thinking that most of her patients elected pharmacologic treatment despite her encouragement to consider behavioral changes. Dr. Garrison continued, "I've found that providing patients more evidence regarding the effectiveness of drugs, and giving them information such as the number needed to treat, encourages them to think more realistically about the benefits and risks of pharmacologic treatment."

Dr. Parra wondered about this and responded with concern: "Really? I don't think our patients can make sense of all this information. They're not trained in statistics and don't know how to interpret scientific data. That's our job." Dr. Garrison asserted that it is an ethical obligation for physicians to provide information such as the number needed to treat and asked Dr. Parra, "How can our patients make informed decisions regarding their health care if they don't know the evidence?"

Commentary

The disagreement between Dr. Parra and Dr. Garrison highlights one of the fundamental ethical issues surrounding Mr. Mendez's care: How much information about evidence should physicians communicate to patients to enable them to make informed decisions? The answer will most likely depend upon the varying assessments of different physicians, the varying preferences of different patients, and the varying professional knowledge available in different clinical situations. Yet there is a general recognition in contemporary medicine that something needs to be said about the available research evidence related to a specific clinical decision in order to enable patients to make informed decisions about their health—even if wisdom is needed to understand how that evidence should be interpreted and communicated.

Shared Decision Making and the Need for Communicating Evidence

Communicating information about prognosis and treatment is recognized as one of the clinical cornerstones of respecting patient autonomy. The patient's right to informed self-determination implies a corresponding obligation for physicians to provide relevant and understandable information. This is no small task when dealing with the complexities of risk communication. Patient numeracy, physician framing of risks, and embedded mental shortcuts that simplify complex information and decisions (i.e., heuristics) are but a few of the challenges physicians face when sharing treatment information. One critical issue is the degree to which Mr. Mendez understands his prognosis and the risks and benefits of his treatment options. Such challenges have been discussed in a recent *Virtual Mentor* article [1] and need not be recapitulated here.

Instead, we focus on fundamental questions of professional ethics that lie beneath the surface of Mr. Mendez's case. Dr. Garrison's and Dr. Parra's disagreement about

their perceived ethical obligations to share evidence reflects different conceptions about the balance of control over decision making in the patient-physician relationship, which can tilt more towards the patient or more towards the physician in any given patient-physician dyad.

These issues are given greater relevance in the context of shared decision making, which is characterized by the involvement of both the patient and the physician in the process of selecting among treatment options [2]. Through patient-physician dialogue, shared decision making aims to promote the ethical principle of respect for patient autonomy by empowering, or "conferring agency" on, patients [3]. It explicitly involves patients in decision making by eliciting their preferences and communicating information about treatment options [4]. The communication of evidence is therefore a necessary pillar of the shared decision-making model.

If it is agreed that respect for patient autonomy means that patients need to be informed about evidence pertaining to treatment options, then physicians have a professional obligation to provide such information. But if this obligation is viewed within the complex network of accountability that connects patients, physicians, and society [1], fulfilling it requires us to balance the principle of respect for patient autonomy with the ethical principles of beneficence and justice [3].

Because the complexities of striking such a balance may be underappreciated, we discuss below each of these principles and attempt to show how they can inform our understanding of the nature and extent of a physician's obligation to disclose to patients evidence related to treatment options. Our purpose is to suggest that the decision regarding how much information a patient should receive depends on more than patient numeracy and lessons learned from the science of risk communication. More fundamentally, we believe this decision depends on a physician's professional understanding of his or her ethical obligations that flow from the roles of patient and physician in the context of shared decision making.

Relational Autonomy and the *Shared* in Shared Decision Making

Respect for patient autonomy refers to a physician's obligation to respect those patient preferences and decisions that accord with that patient's values and beliefs [5]. Communicating treatment options, eliciting patient preferences, and explicitly recognizing the authority of the patient to make treatment decisions are all practices that promote patient autonomy. These practices stand in contrast to more paternalistic approaches in which the locus of decision making tilts more toward the professional authority of the physician.

While promoting patient involvement is a necessary corrective to paternalistic approaches [6], it would be misguided for physicians to assume that shared decision making is only about patient autonomy. As one author put it, "That one party becomes more responsible does not necessarily make the other party less responsible" [7]. This is particularly relevant when patient autonomy is framed as *relational autonomy*. Relational autonomy recognizes that an individual's identity

and values do not exist in a vacuum, but are "constituted in and by" interpersonal relationships and the broader social environment [8]. In this view, a patient's social relationships (especially their relationships with physicians) inform his or her medical preferences and treatment decisions.

Seeing patients and their autonomy as existing in relation to other persons, including physicians, can help us think about how much evidence related to treatment options we should communicate to patients. If social relationships and circumstances shape a patient's autonomy, physicians ought to be sensitive to how these factors may bear on treatment decisions [8]. For example, if difficult or negative prior encounters with health professionals or treatments have caused a patient to discount or prematurely dismiss what a physician perceives to be a good treatment option, part of the "relational work" to enhance autonomy may involve providing more detailed evidence, or more carefully explained evidence, in order to show respect through patient explanation and listening and, thereby, gain trust.

Relational autonomy also helps remind us to consider the question of who decides *which* treatment options should be offered in the first place, a role largely assumed by physicians. We acknowledge that access to information on the Internet about traditional and alternative therapies can have an extraordinary and independent impact on a patient's understanding of treatment options. Nevertheless, it is important to appreciate that, when physicians communicate about evidence pertaining to treatments and define the treatment options available, they are contributing to the relational autonomy of their patients.

The principle of respect for patient autonomy is relevant to Mr. Mendez's case in several ways. First, considering the notion of relational autonomy, we should ask: What social circumstances inform Mr. Mendez's understanding of his high cholesterol? Are there circumstances that strongly discourage Mr. Mendez from taking a statin, even if his primary goal is to avoid a future cardiac event at all costs? How is Mr. Mendez's capability to enact his goals for care shaped by these social circumstances? Secondly, what treatment options did Dr. Garrison present to Mr. Mendez, and how did these shape his decision-making autonomy? Exploring the answers to the above questions would have enabled Dr. Garrison to appreciate Mr. Mendez's circumstances and how they shaped his ultimate treatment decision. They would have informed her approach to providing evidence such as the number needed to treat and provided an ethical rationale with which to engage Dr. Parra's objections.

Beneficence

The principle of beneficence refers to a physician's obligation to act for the good of the patient [5]. Implicit within this definition is the recognition that a physician must act in accordance with what *he or she believes* is best for the patient, as informed by his or her expertise and clinical judgment. With respect to deciding what treatment options to present to patients, beneficence plays a vital role. What "option set" a

physician selects is inherently influenced by what the physician believes to be the best and most appropriate treatments for promoting the patient's well-being [6].

With regard to communicating evidence, beneficence also supports efforts to promote a patient's understanding of the treatment options offered (here one sees evidence of the interrelationship between autonomy and beneficence, since one of the ways to act for a patient's good is to enhance his or her autonomy). For example, it has been well documented that people are often overly optimistic regarding health risks. In one study in which participants were informed that 4 out of every 100 persons would suffer a poor outcome, some people were unreasonably optimistic that they would not be among the unlucky 4 percent [9]. Thus, even if a patient selects a treatment that the physician favors, beneficence may oblige the physician to offer more information to ensure that the patient takes both the risks and benefits seriously.

Dr. Garrison proposed statin therapy as a reasonable treatment for someone in Mr. Mendez's state of health. What she thinks is good for Mr. Mendez most likely influenced the information she shared about taking statins and the treatment alternatives she offered (if any). In describing the number needed to treat, Dr. Garrison may have been trying to make sure Mr. Mendez fully understood how the benefits and risks of treatment applied to him. Dr. Parra's objection to Dr. Garrison's approach most likely reflected her own notion of beneficence, which could include an interpretation of the evidence and making a recommendation for a preferred treatment (a recommendation that could be changed if side effects happened to occur).

Justice

Justice in the setting of health care usually pertains to questions of fair distribution ("distributive justice") or to questions about what each person deserves, or is "due," as an individual endowed with human dignity [5]. However, what exactly a person is due with respect to evidence pertaining to treatments is not clear, especially when that evidence pertains to populations as well as the individuals who comprise them. Physicians are traditionally trained to focus on the needs and preferences of individual patients, one at a time. But there are also times when physicians are called upon to weigh simultaneously the needs of an individual patient and the needs of a population or society. Such situations raise challenging questions of distributive justice.

Due to cost concerns, we appear to be in a time of increased attention to the needs of society, as can be seen in the new emphasis on "value-based" and "cost-conscious" health care that takes into account the health outcomes achieved per dollar spent [10]. According to a society-focused view, the success of a treatment is evaluated not only by its effect on the individual patient's well-being but also on the cost-effectiveness of the treatment over the long term (e.g., prevention of hospitalization from a myocardial infarction) [11]. With limited health care resources, there is an increasing premium placed on treatments that are backed by quality evidence, offer

more than marginal health gains for individuals, and come at acceptable costs to society.

We would suggest that there is often a deep tension at play when a physician is expected to act simultaneously in the interests of an individual patient and a population or society. A discussion of this tension is beyond the scope of this essay. Suffice it to say that there appear to be increasing expectations that physicians become more cost-conscious about the care they authorize and should see cost-consciousness as part of their accountability to society.

But it should be noted that the effort to balance the needs of individuals with populations (and societal costs) is not a new concern for physicians. Take vaccination programs, for example, and consider the similarities between statins and vaccines. Both statins and vaccines require a financial investment to prevent the occurrence of potentially fatal and costly outcomes. Both have possible side effects and adverse reactions. Both are less than 100 percent effective in preventing the disease in people who receive the treatment.

In the case of vaccines, there is also a societal expectation that patients should be encouraged to receive treatment to benefit society, not only themselves. The physician's obligation is to encourage patients, insofar as it is medically and personally acceptable, to contribute to this societal benefit by being vaccinated. One could think about statin therapy in a similar way. A study evaluating the cost-effectiveness of expanding statin therapy to all low-risk patients found that such an expansion could prevent nearly 14,000 deaths from coronary heart disease and save approximately $1.4 billion per year [12]. One could imagine such cost savings translating into funding for other pressing health issues. On this view, does Mr. Mendez's decision not to take a statin undermine the well-being of society by potentially diverting funds away from other causes? Does society expect Mr. Mendez to choose statin therapy as a cost-effective treatment plan? Most importantly for our discussion, how does (or should) a physician weigh these considerations when he or she thinks about an individual patient's care? Will a physician's thoughts about societal benefits impact what information he or she decides to communicate with patients about treatment options and their benefits?

We make the comparison of statins and vaccines not to advocate the use of statins, but to highlight how a physician's perceived responsibility to society could have an impact on shared decision making in the patient-physician relationship. To the extent that calls to incorporate cost-consciousness in clinical decision making are justified by considerations of distributive justice, it is understandable that physicians should be encouraged to recommend treatments that promote individual health and are cost-effective for society. But such encouragements should not lose sight of the double responsibility this involves and the potential or actual tension between the well-being of individual patients and the cost savings for society.

In sharing the number needed to treat for statin therapy, Dr. Garrison may have provided evidence that discouraged Mr. Mendez from accepting a pharmacologic treatment that could have been personally beneficial and cost-effective for society. Alternatively, Dr. Garrison could be seen as having encouraged lifestyle modifications that, if successful, could also be beneficial to the patient and cost-effective for society. While Dr. Parra's objection to Dr. Garrison's practice appeared to focus upon the interpretability of the evidence, she might also have objected to Dr. Garrison's disclosure of information that discouraged the acceptance of a potentially cost-saving therapy. One wonders whether Dr. Garrison should have considered communicating information to Mr. Mendez about potential cost savings to society, either as a matter of disclosure (since it could possibly have influenced her recommendations) or as a means of encouraging of civic cooperation. This is a very significant question, one we raise not to promote a position but to stimulate discussion about the tension between individual and societal benefit.

Evidence and Uncertainty

Our discussion would be incomplete if we did not acknowledge that evidence always carries a degree of uncertainty. Statistics such as the number needed to treat are derived from population studies, and probabilities from such studies do not map perfectly onto individual experiences. For Mr. Mendez, the treatment effects of the statin cannot be predicted with complete accuracy. This uncertainty is amplified when the evidence and opinions surrounding a certain treatment are conflicting, as is the case for prescribing statins to low-risk individuals [13]. Without a shared understanding of the uncertainty surrounding both the evidence provided and the decision to pursue a course of treatment, the potential for true patient involvement in decision making may be limited [6].

Shared Decision Making and the Balancing of Ethical Principles

How physicians understand their professional obligations will influence what information they communicate to patients and their approach to clinical decision making. To engage in shared decision making honestly requires that physicians acknowledge (1) their responsibility to the patient (respect for patient autonomy), (2) their responsibility to be true to their own clinical judgment about the best interests of the patient (beneficence), (3) their accountability to society (distributive justice), and (4) the uncertainty of the evidence. While it should also be acknowledged that these ethical principles interact within a broader web of treatment guidelines, local institutional practices, and payer policies—all of which influences the context of shared decision making [6]—how these principles are balanced in a given clinical scenario should be expected to influence what information about evidence a physician decides to share with his or her patient.

The complex interplay of ethical principles in this case reminds us that information about risks and benefits of treatment should not be seen as a self-interpreting package of data ready for delivery. Rather, the information we share with our patients reflects our own ethical values and our interpretations of the evidence. It also reflects our implicit understanding and balancing of our responsibilities to

individual patients, ourselves, and society. Appreciating that a patient's autonomy is relational and shaped in part by the information a physician provides does not weaken the ethical imperative for shared decision making. It strengthens it. Through the process of patient-physician communication, patient preferences can be elicited to help shape the evidence that physicians share. By recognizing the ethical principles and value judgments at stake in such dialogue, physicians should be in a better position to decide what evidence to share with patients and, we may hope, more likely to truly engage in shared decision making.

References

1. Schwartz PH. Discounting a surgical risk: data, understanding, and gist. *Virtual Mentor*. 2012;14(7):532-538.
2. Charles C, Gafni A, Whelan T. Shared decision-making in the medical encounter: what does it mean? (Or it takes at least two to tango). *Soc Sci Med*. 1997;44(5):681-692.
3. Elwyn G, Frosch D, Thomson R, et al. Shared decision making: a model for clinical practice. *J Gen Intern Med*. 2012;27(10):1361-1367.
4. A set of "competences" has been elaborated for the model of shared decision making that includes: explicit involvement of the patient in the decision making process, providing information in a patient's preferred format, providing treatment options using equipoise statements, checking for patient understanding, eliciting patient preferences, and formulating a consensus about treatment. For further discussion, see: Elwyn G, Edwards A, Kinnersley P, et al. Shared decision-making and the concept of equipoise: defining the competences of involving patients in health care choices. *Br J Gen Pract*. 2000;50(460):892-899.
5. Kaldjian LC, Weir RF, Duffy TP. A clinician's approach to clinical ethical reasoning. *J Gen Intern Med*. 2004;20(3):306-311.
6. Wirtz V, Cribb A, Barber N. Patient-doctor decision-making about treatment with the consultation—a critical analysis of models. *Soc Sci Med*. 2006;62(1):116-124.
7. Holm S, Davies M. Ethical issues around shared decision-making and evidence-based patient choice. In: Edwards A, Elwyn G, eds. *Shared Decision-Making in Health Care: Achieving Evidence-Based Patient Choice*. New York: Oxford University Press; 2009:59-64.
8. Mackenzie C. Relational autonomy, normative authority and perfectionism. *J Soc Philosophy*. 2008;39(4):512-533.
9. Weinstein ND. Unrealistic optimism about susceptibility to health problems: conclusions from a community-wide sample. *J Behav Med*. 1987;10(5):481-500.
10. Patel MS, Davis MM, Lypson ML. The VALUE Framework: training residents to provide value-based care for their patients. *J Gen Intern Med*. 2012;27(9):1210-1214.
11. Owens DK, Qaseem A, Chou R, Shekelle P; Clinical Guidelines Committee of the American College of Physicians. High-value, cost-conscious health

care: concepts for clinicians to evaluate the benefits, harms, and costs of medical interventions. *Ann Intern Med*. 2011;154(3):174-180.
12. Lazar LD, Pletcher MJ, Coxson PG, Bibbins-Domingo K, Goldman L. Cost-effectiveness of statin therapy for primary prevention in a low-cost statin era. *Circulation*. 2011;124(2):146-153.
13. Debates about the use of statins in low-risk individuals are rampant in the literature and often involve contrary interpretations of the same clinical trial data. For an opinion piece supporting statin use in low-risk individuals, see: Blaha MJ, Nasir K, Blumenthal RS. Statin therapy for healthy men identified as "increased risk." *JAMA*. 2012;307(14):1489-1490. For an opinion piece against statin use in low-risk individuals, see: Redberg RF, Katz MH. Healthy men should not take statins. *JAMA*. 2012;307(14):1491-1492.

Paul J. Christine, MPH, is an MD-PhD student at the University of Michigan Medical School and School of Public Health in Ann Arbor. He is pursuing a doctoral degree in epidemiology and is broadly interested in the relationship between population health and clinical practice.

Lauris C. Kaldjian, MD, PhD, is director of the Program in Bioethics and Humanities and a professor in the Department of Internal Medicine at the University of Iowa Carver College of Medicine in Iowa City. Dr. Kaldjian practices general internal medicine, and his research has focused on goals of care and patients' end-of-life treatment preferences, physician disclosure of medical errors, ethics education, and the role of philosophical and religious beliefs in clinical decision making.

Related in VM

Effective Ways to Communicate Risk and Benefit, January 2013

Discounting a Surgical Risk: Data, Understanding, and Gist, July 2012

When Patients Seem Overly Optimistic, July 2012

The people and events in this case are fictional. Resemblance to real events or to names of people, living or dead, is entirely coincidental.

The viewpoints expressed on this site are those of the authors and do not necessarily reflect the views and policies of the AMA.

Virtual Mentor
American Medical Association Journal of Ethics
January 2013, Volume 15, Number 1: 18-22.

ETHICS CASES
Authority, Health Advocacy Organizations, and Scientific Evidence
Commentary by Jodi Halpern, MD, PhD, and Richard L. Kravitz, MD, MSPH

Dr. Sanders, a second-year oncology fellow, arrived for work carrying a sheath of bright pink pamphlets. He set them down on a table in the staff room and proclaimed to his colleagues: "It's that time of year again—the annual charity walk is coming up! The sign-up sheet is on the table here for anyone who wants to participate. The work this group does is so important in the fight against breast cancer!"

A new oncology fellow, Dr. Wang, walked over to the table and picked up one of the pamphlets. "These guys want to increase access to mammography for all women under 50. That's the opposite of the new U.S. Public Health Task Force recommendations." The evidence was pretty good, he thought, for their guideline that mammography for women under 50 be limited to high-risk groups because of the risks associated with false positive results—unnecessary biopsies, anxiety, and so on.

Dr. Sanders replied, "Well, I think that the work the group does to help vulnerable women avoid and treat breast cancer is really important. I have a lot of respect for what they do."

"But," Dr. Wang interjected, "isn't it problematic for you, a physician, to support a group that promotes health care activities that go against the evidence? This seems like bad medicine. Plus, doesn't this group just encourage even more people to think of evidence-based medicine as rationing? You are supposed to be a steward of our profession and of the best evidence-supported medicine. I don't think it's right for someone in your profession to support a group like this—it seems bad for medicine and bad for patients."

Commentary
In the vignette, Dr. Sanders and Dr. Wang disagree about what professionalism requires of them. While "professionalism" refers to doctors' conduct, medical leaders have wisely avoided telling physicians how to behave and instead defined the goals the profession should be serving. These goals are to safeguard patient welfare, to respect patient autonomy and to promote social justice [1, 2]. In practice, professionalism often comes down to discerning how best to meet these goals, given that goals can conflict with each other and meet up with organizational and resource constraints. Here, Dr. Sanders supports a patient advocacy organization that promotes women's health and excellence in cancer care that may be overexuberant with regard to screening. Dr. Wang believes that physicians have a stronger, perhaps

even absolute, obligation to reject any organization whose recommendations appear to skirt the best evidence.

In conflicts over professionalism, a good first step is to clarify the facts. An apparent ethical conflict can dissolve when the circumstances are better understood. Thus, we first reflect on the nature of the evidence regarding mammography and the societal context in which disease advocacy groups operate. Then, in the second half of our commentary, we turn to the possibility that the doctors' disagreement might reflect conflicting conceptions of professional responsibility.

The Context of the Case

Let us begin by considering the evidence behind the guidelines for mammography. As John Ioannidis has pointed out, many published research findings are likely to be false [3]. All evidence is not created equal. Among the reasons for misleading research findings are fraud, error, bias, and chance. Beyond the raw data, guidelines produced by reputable organizations frequently disagree. In the case of breast cancer screening, the American Cancer Society continues to recommend that women under 50 should discuss the value of mammography with their personal physicians [4], whereas the United States Preventive Services Task Force (USPSTF) recommends against *routine* screening in this age group [5]. Because scientific evidence is inevitably context-dependent and often subjective it should not be ignored but neither should it be worshipped.

Importantly, even if a therapeutic or preventive intervention is of limited or no benefit for most patients, the intervention might still be worthwhile for selected subgroups. This is what statisticians and epidemiologists call heterogeneity of treatment effects (HTE) [6]. In the case of breast cancer, women between 40 and 49 may obtain greater-than-average benefit from mammography if: (1) they have higher-than-average risks of developing breast cancer; (2) they have less dense breasts, allowing easier diagnosis; or (3) they are already very worried about cancer or are phlegmatic enough to be unperturbed by false positives and the downstream risk of unnecessary biopsy and overtreatment.

Reflecting on the facts also includes considering the range of services provided by the women's advocacy group. Let's say that, while overselling the benefits of mammography, this advocacy organization also fights cancer-associated stigma, educates women who need testing and might not otherwise seek it, provides crucial support for women with breast cancer, and supports research efforts that may help ultimately prevent or cure breast cancer. Given how little societal support there is for people grappling with illness in our country, advocacy groups can make a profound difference in women's lives. Just as people frequently support political parties without endorsing their entire platforms, perhaps physicians can ethically support a health care advocacy organization without agreeing with everything it says and does. Thus, Dr. Sanders might be justified in supporting the advocacy organization if he concludes that the group does, on balance, more good than harm.

On the other hand, supporting a group that ignores important scientific evidence can contribute to a very dangerous "antiscience" societal trend. At a time when 40 percent of Americans don't believe in evolution and think that humans were created in the "last 10,000 years or so" [7], the standing of science in America is perilous. Furthermore, as Dr. Wang implies, the rising cost of health care threatens our national well-being but any attempt to rationalize use of medical resources is painted by regressive forces as "rationing" or "death panels." Under these circumstances, it could be argued that physicians, trained in science, have a duty to defend the scientific method. What is to be given greater priority: patient groups that support and improve women's health now or scientific integrity, which is critical to medical progress in the future?

Models of Professional Responsibility

To address such questions requires analyzing our ethical as well as factual assumptions. Note that we have already been considering the ratio of good to harm, which suggests a utilitarian approach. Utilitarianism focuses on the outcomes of actions and defines right actions as those that an impartial spectator would calculate as having the greatest aggregate benefit. Should doctors attempt to guide their professional conduct by such calculations? How could doctors accurately weigh the long-term social costs of downplaying evidence-based recommendations against the current benefits of supporting an advocacy group that works on behalf of women's health?

Not only would attending to every such issue make practicing medicine impossibly demanding, it would distract doctors from their primary responsibility, which is to pay attention to their patients. It is not just that, as John Rawls famously argued, utilitarianism is blind to the individual, favoring aggregate outcomes even when this sacrifices individual welfare and autonomy [8]. Rather, it is that doctors are not morally positioned to be impartial social planners, as they are not independent of particular relationships and duties. Patients trust physicians with their lives based on the expectations that physicians are committed to them, to putting their interests first. Recall the patient-centered focus of the principles of professionalism: physicians are responsible for serving each patient's welfare, and for respecting each patient's autonomy.

How does taking a patient-centered approach resolve the conflict between Drs. Sanders and Wang? It reminds these physicians that practicing evidence-based medicine is not a good in itself; it is a good when it enables better care of each individual patient, present or future. In this case, better care includes social and psychological support as well as the information provided by advocacy groups.

According to their roles as trusted fiduciaries, Drs. Sanders and Wang should redirect their attention back to their patients' lives. They might ask: how can I act today in a way that best serves Mrs. Jones, Ms. Arquette, and Ms. Martinez? Do they each truly depend on this health advocacy organization? Do they have other sources of support? Can I educate them to avoid unnecessary mammograms despite the role

of emotions in driving decision making? If not, are there other organizations that provide accurate information as well as social support?

Additionally, doctors' responsibilities extend beyond their known patients. The third principle of professionalism, to promote social justice, reminds doctors that their role responsibilities extend to broader populations. Moreover, doctors are responsible for sustaining the public's trust in physicians. In addition to the questions above, then, doctors might also ask themselves: Am I considering the unmet needs of women from disadvantaged communities? Am I upholding the scientific integrity that sustains public trust? Am I promoting the science that may discover new treatments?

Addressing these questions reveals that in this case there is a strong link between evidence-based care and promoting the welfare of individual women. Yet it is crucial for women to be informed about advocacy groups that can empower them and promote their health. The challenge for physicians is to balance both of these obligations to best serve the goals of medical care.

Thus, while it may be fine to bring brochures to the medical staff room, we would recommend against distributing brochures in the patient waiting room. Absent discussion with their doctors, women might take this to be an endorsement of this particular group's views of mammography. Instead, Dr. Sanders should inform individual patients about the organization and its pros and cons as part of educating them both about the need for discernment regarding mammography and about the health value of social support. Assuming that Dr. Wang does not know of another comparable source of social support, we think that, despite his misgivings, he too should inform women about this organization as one source of possible social support even as he shares his concerns with them. Perhaps Dr. Sanders and Dr. Wang might also support the development of more evidence-based advocacy groups (through speaking or writing about the science, for example).

In summary, physicians' strongest moral obligation is not to impartially protect science, but rather to fulfill their role responsibilities to patients. Still, practicing evidence-based medicine is one of the most powerful ways to serve individual patients, so professionalism will most often coincide with favoring scientific practice. The ethical issue is to see that science—as it constitutes the "evidence base"—is a means to the end of helping real people. Good clinical practice integrates clinical research, professional experience, and knowledge of the individual patient.

References

1. ABIM Foundation, the ACP–ASIM Foundation, and the European Federation of Internal Medicine. Medical professionalism in the new millennium: a physician charter. *Ann Intern Med*. 2002;136(3):243-246.
2. Epstein, RM, Hundert, EM. Defining and assessing professional competence. *JAMA*. 2003; 287(2):226-235.
3. Ioannidis JP. Why most published research findings are false. *PLoS Med.* 2005;2(8):e124.

4. Smith RA, Cokkinides V, Brawley OW. Cancer screening in the United States, 2009: a review of current American Cancer Society guidelines and issues in cancer screening. *CA Cancer J Clin*. 2009;59(1):27-41.
5. Woolf SH. The 2009 Breast Cancer Screening Recommendations of the US Preventive Services Task Force. *JAMA.* 2010;303(2):162-163.
6. Kravitz RL, Duan N, Braslow J. Evidence-based medicine, heterogeneity of treatment effects, and the trouble with averages. *Milbank Q*. 2004;82(4):661-666.
7. Gallup Politics. In U.S., 46% hold creationist view of human origins. http://www.gallup.com/poll/155003/hold-creationist-view-human-origins.aspx. Accessed October 1, 2012.
8. Rawls J. "Justice as Fairness." *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971.

Jodi Halpern, MD, PhD, is an associate professor of bioethics and medical humanities in the Joint Medical Program and the School of Public Health at the University of California, Berkeley. A psychiatrist with a doctorate in philosophy, she is the author of *From Detached Concern to Empathy: Humanizing Medical Practice* (paperback 2011, Oxford University Press).

Richard L. Kravitz, MD, MSPH, is professor and co-vice chair (research) in the Department of Internal Medicine at University of California, Davis. He is also co-editor in chief of the *Journal of General Internal Medicine*.

Related in VM
The Debate over Prostate Cancer Screening Guidelines, January 2011

The people and events in this case are fictional. Resemblance to real events or to names of people, living or dead, is entirely coincidental.

The viewpoints expressed on this site are those of the authors and do not necessarily reflect the views and policies of the AMA.

Virtual Mentor
American Medical Association Journal of Ethics
January 2013, Volume 15, Number 1: 23-27.

MEDICAL EDUCATION

Teaching Critical Appraisal of Medical Evidence

Martha Carvour, MD, PhD

Teaching medical learners how to evaluate evidence for medical decision making represents both a critical and daunting task for curriculum designers. It is, after all, a desirable outcome for any curriculum in evidence-based medicine to produce graduates who not only understand the medical literature, but who can appraise, discuss, and apply what they read in a sound, thoughtful, and efficient manner [1].

When critical appraisal skills remain underdeveloped, students tend to resort to other, less useful approaches. One of these might be called *convenient appraisal*, in which the reader merely accepts, at face value, what the literature claims. Another, equally concerning, approach might be called *cynical appraisal*, in which the reader simply rejects the bulk of the literature, perhaps citing the potential for bias in all research studies. The critical reader, by contrast, takes a balanced approach to the medical literature, seeking to glean the most from each piece, approaching it with reason (not reflex), and supplementing the evidence from one article with data from other sources to build a thoughtful and defensible assessment of a study's content.

Instruction in critical appraisal skills is sometimes deferred in favor of teaching students *what* they read rather than demonstrating *how* to read it. Below are three practical methods for reversing this trend—that is, for actively teaching critical appraisal skills in an evidence-based medicine curriculum.

Teaching Critical Appraisal Skills

1. Start with what makes sense. It is hard to imagine a student who believes that basing the practice of medicine on evidence is *un*important. There are certainly many, however, who question whether incorporating the skills of evidence-based medicine into clinical decision making is practicable and who doubt that they can ever become proficient in these skills. Indeed, understanding of evidence-based medicine concepts among medical learners tends to be lower than desired [2].

Yet, the fact is that many principles of evidence-based medicine—including some of the most complex—mirror the way that many clinicians *already* think about medicine. For instance, the likelihood that a treatment will benefit a patient often depends on the characteristics of the patients in whom it is used. This parallels the evidence-based medicine concepts of "effect modification" or its practice of differentiating between effectiveness and efficacy.

Similarly, in clinical practice, a patient's symptoms may be mistakenly attributed to one diagnosis when the circumstances surrounding the case conceal the genuine cause. This is analogous to the evidence-based medicine concept of confounding. Likewise, individual patients and their experiences are, unsurprisingly, not all the same. This nonuniformity can be linked to discussions in evidence-based medicine about diagnostic validity and precision or about variation and central tendency. For instance, how might a particular diagnostic result be interpreted in a population at high risk for disease compared to a population at lower risk? How should the concept of error in diagnostic tests influence conversations with patients about diagnoses? When evaluating a research study, should clinicians direct more focus to the experience of the entire study population (the distribution or curve representing all patients in the study) or to the average experience (indicated by a mean or similar measure)?

Principles like these have the appeal of being both relevant and intuitive. Instruction in evidence-based medicine should start with what intuitively makes sense (e.g., the general principles of prognostic research) and use this context to explain more difficult concepts (e.g., the critical interpretation of particular prognostic statistics).

2. Don't treat continuous concepts as though they were categorical concepts. Evidence-based medical instruction should not shy away from complexity when complexity makes more sense than simplicity. Although there may not always be a single correct interpretation of a research article so much as several reasonable ones, the temptation for curriculum designers is to create simplicity where it may not be very useful in the long run—to create false dichotomies between good and bad studies, for instance, rather than encouraging learners to critically appraise what they read.

The problem is not that students start simple, but that they learn to think this way. Many students have been trained, intentionally or otherwise, to evaluate study designs, statistical analyses, and research findings reflexively.

Many students, at all levels of medical education, evaluate studies simply on the basis of their design, perhaps by an oversimplified dichotomization of randomized controlled trials and a lesser, nontrial, category, or by identifying the position of a study design on a hierarchical map. These approaches fail to evaluate the content or quality of any individual study and render students unprepared to evaluate important subsections of the medical literature. Critical readers of the evidence should have a basic understanding of the utility, purpose, and value of many designs, including nontrial studies, and be able to engage in meaningful discussions about them [3, 4].

Similarly, students may reflexively evaluate a study on the basis of a single number, such as a p-value, regardless of where that value originated, whether the design or analysis makes any sense, or what the statistical test was meant to assess in the first place. This kind of assessment also falls short of the standard of critical appraisal. Instead, students should learn—from the start—how to interpret the concepts of

evidence-based medicine, including p-values, *within their appropriate contexts*. For example: What was the purpose of the study? Was it appropriately designed to answer the question? What biases, confounding factors, and other considerations influence the interpretation of study findings? Do these factors strengthen or weaken the findings? Then (and only then), should they ask: How should I interpret the p-value in the context of a medical decision?

Notably, many educators have recognized the need to expand upon knee-jerk approaches to the medical literature by creating standardized algorithmic methods for students to follow when they read an article from the literature. This is a step in the right direction. But, here again, a temptation exists to conflate algorithmic assessments of research studies with critical evaluations of their content. These are simply not the same. Completing a checklist or calculating a score satisfies neither the intellectual rigor nor the thoughtful independence of a critical evaluation any more than adding Likert-based scores to a student evaluation offers a constructive view of a student's ability to practice medicine.

3. Adopt a SOAP approach. Admittedly, offering context-based learning opportunities in a concept-heavy curriculum presents a considerable challenge. It may be easier, after all, to define, even abstractly, a p-value or a statistical test and then imagine, also abstractly, that students will eventually learn to think critically about it. Here again, curriculum designers should start with what makes sense—that is, with the way clinician educators already think about medical education.

Consider the way that students are taught to approach the diagnosis of disease. Typically, the process begins with a chief concern—a problem to be solved or a question to be answered. This must then be supported by a series of questions, examination techniques, and laboratory or radiological analyses developed in a thoughtful, systematic, yet patient-specific manner in order to refine a differential diagnosis, identify a working diagnosis, and outline a plan of action together with the patient.

This process is expressly predicated on the notion that a single laboratory value—an international normalized ratio (INR), for instance—derives its practical meaning from the rest of the story that precedes and encompasses it. An INR value—2.5, let's say—is not very useful unless the practitioner knows where it originated and how it may influence decision making. (Is this a therapeutic INR in a patient on warfarin, or is it too elevated to justify nonurgent paracentesis in a patient with cirrhosis?)

The process by which students learn to think about diagnosis may also be applied in evidence-based medicine to promote contextual appraisal of the literature. One possible approach is shown below. While this outline may be used as an example, it does *not* represent an algorithm or formula for critical evaluation. Rather, just as the process of diagnosis must be based on standards of care tailored to individual circumstances, the process of critical evaluation should be founded on standards of

quality adapted to the clinical question of interest and the studies designed to address it.

Subjective

- What is the question to be answered? What information is needed to answer the question? How does the study approach the question? Is the design well matched with the question, and is the measured outcome relevant to it?
- What kind of information was collected? Where did the information originate?
- What analyses were used? What are the assumptions underlying the analyses?

Objective

- What are the results? How are these presented (e.g., p-values, confidence intervals, figures)?
- Is all of the necessary information provided to make an assessment? What is missing?
- How accurate or precise is the information (e.g., estimates of error, sensitivity analyses)?

Assessment

- What are the biases, confounding factors, and other considerations influencing appraisal of the subjective and objective information? Do these strengthen or weaken the findings?
- What relevant contextual questions are not answered by this study? How might these be addressed?
- Taking all of this information together, what is a reasonable interpretation of the study findings (even if it differs from that of others who read the same study)?

Plan

- Despite any limitations of the study, can this information be useful? If so, how and when might it be applied? What are the limitations or alternatives of this plan?
- What other information should be sought to aid in answering the original question? How can this information be obtained?

Conclusion

There is no substitute for genuine critical appraisal of the medical evidence. Medical learners who lack training in this skill may learn to rely instead on convenience (e.g., choosing a randomized trial to present to the group and, upon finding a low p-value, accepting its results at face value) or cynicism (e.g., identifying a shortcoming of a study without offering reasoned explanation about how it affects the results and disregarding anything to be gained from the research). However, critical appraisal skills may be successfully incorporated into evidence-based medical curricula by starting with what makes sense and using this as a context for more challenging concepts, by limiting oversimplification of the appraisal process, and by encouraging students to develop a systematic, yet nonalgorithmic, approach to evidence appraisal.

References

1. Bayley L, McLellan A, Petropoulos JA. Fluency-not competency or expertise-is needed to incorporate evidence into practice. *Acad Med*. 2012;87(11)1470.
2. Windish DM, Huot SJ, Green ML. Medicine residents' understanding of the biostatistics and results in the medical literature. *JAMA.* 2007;298(9):1010-1022.
3. Baird JS. Journal clubs: what not to do. *Acad Med*. 2012;87(3):257.
4. Rawlins M. De testimonio: on the evidence for decisions about the use of therapeutic interventions. *Lancet.* 2008;372(9656):2152-2161.

Martha Carvour, MD, PhD, is a first-year resident in internal medicine at the University of Texas Southwestern Physician Scientist Training Program in Dallas. Prior to residency, Dr. Carvour completed her MD and PhD (epidemiology) degrees at the University of Iowa, where she first became involved in the design, development, and delivery of an evidence-based medicine curriculum.

Related in VM

Rating Evidence in Medical Literature, January 2011

Insights from Teaching Evidence-Based Medicine, January 2011

Authority, Health Advocacy Organizations, and Scientific Evidence, January 2013

Virtual Mentor
American Medical Association Journal of Ethics
January 2013, Volume 15, Number 1: 28.

THE CODE SAYS
The AMA *Code of Medical Ethics*' Opinion on Informing Patients about Treatment Options

Opinion 8.08 - Informed Consent
The patient's right of self-decision can be effectively exercised only if the patient possesses enough information to enable an informed choice. The patient should make his or her own determination about treatment. The physician's obligation is to present the medical facts accurately to the patient or to the individual responsible for the patient's care and to make recommendations for management in accordance with good medical practice. The physician has an ethical obligation to help the patient make choices from among the therapeutic alternatives consistent with good medical practice. Informed consent is a basic policy in both ethics and law that physicians must honor, unless the patient is unconscious or otherwise incapable of consenting and harm from failure to treat is imminent. In special circumstances, it may be appropriate to postpone disclosure of information, (see Opinion 8.122, "Withholding Information from Patients").

Physicians should sensitively and respectfully disclose all relevant medical information to patients. The quantity and specificity of this information should be tailored to meet the preferences and needs of individual patients. Physicians need not communicate all information at one time, but should assess the amount of information that patients are capable of receiving at a given time and present the remainder when appropriate.

Issued March 1981, updated November 2006, based on the report "Withholding Information from Patients (Therapeutic Privilege)," adopted June 2006.

Related in VM
Communicating Evidence in Shared Decision Making, January 2013

Effective Ways to Communicate Risk and Benefit, January 2013

Virtual Mentor
American Medical Association Journal of Ethics
January 2013, Volume 15, Number 1: 29-33.

JOURNAL DISCUSSION
When Research Evidence is Misleading
Chetan Huded, MD, Jill Rosno, MD, and Vinay Prasad, MD

Ioannidis JP. Why most published research findings are false. *PLoS Med*. 2005;2(8):e124.

Each year, millions of research hypotheses are tested. Datasets are analyzed in ad hoc and exploratory ways. Quasi-experimental, single-center, before and after studies are enthusiastically performed. Patient databases are vigorously searched for statistically significant associations. For certain "hot" topics, twenty independent teams may explore a set of related questions. From all of these efforts, a glut of posters, presentations, and papers emerges. This scholarship provides the foundation for countless medical practices—the basis for the widespread acceptance of novel interventions. But are all of the proffered conclusions correct? Are even most of them?

John P. A. Ioannidis's now famous work, "Why Most Published Research Findings Are False," makes the case that the answer is no [1]. Ioannidis uses mathematical modeling to support this claim. The core of his argument—that a scientific publication can be compared to a diagnostic test, and that we can ask of it, "how does a positive finding change the probability of the claim we are evaluating?"—has a simple elegance.

Ioannidis asks us to think broadly about the truth of a claim, contemplating not only the study in question but the totality of the evidence. He uses the concept of positive predictive value (PPV)—that is, the probability that a positive study result is a true positive—as the foundation of his analysis. He demonstrates that, under common statistical assumptions, the PPV of a study is actually less than 50 percent if only 5 of 100 hypotheses relating to a field or area of study are true. This means that, under these circumstances—which seem reasonable in many medical fields—a simple coin flip may be as useful as a positive research finding in determining the validity of a claim.
Bias, however, is a ubiquitous threat to the validity of study results, and testing by multiple independent teams can further degrade research findings. As bias, such as that resulting from the use of poor controls or study endpoints, increases, the PPV of a study decreases. Similarly, the global pursuit of positive research findings by multiple investigators on a single subject also decreases the likelihood of true research findings due to a phenomenon called multiple hypothesis testing.

From this framework, Ioannidis draws 6 conclusions:

1. The smaller the sample size of a study, the lower its PPV.
2. Fields with lower effect sizes (e.g., the degree of benefit/harm afforded by a treatment or diagnosis) suffer from lower PPV.
3. The prestudy probability of a true finding influences the PPV of the study—thus, fields with many already tested hypotheses are less likely to yield true findings.
4. Increased flexibility in study design, endpoints, and analysis affords more opportunities for bias and false results.
5. The greater the financial and other external influences, the greater the bias and lower the PPV.
6. There is a paradoxical relationship between the number of independent investigators on a topic and the PPV for a given study in that field—the hotter the field, the more spurious the science.

The consequences of all this bad science are not felt by researchers—whose careers may be propelled by these erroneous results—but by the patients who are subject to medical practices the validity of which is uncertain. Certain medical practices have risen to prominence based on false preliminary results only to be contradicted years later by robust, randomized controlled trials. Examples from recent history include the use of hormone replacement therapy for postmenopausal women [2], stenting for stable coronary artery disease [3], and fenofibrate to treat hyperlipidemia [4].

We have previously called this phenomenon "medical reversal" [5]. Reviewing all original articles published in the *New England Journal of Medicine* during 2009, we found that 46 percent (16 of 35) of articles challenged standard of care and constituted reversals. These reversals spanned the range of medical practice, including screening practices (prostate-specific antigen testing), medications (statins in hemodialysis patients), and even procedural interventions (such as percutaneous coronary intervention in stable angina).

Ioannidis has also provided an empirical estimate of contradiction in medical research [6]. He reviewed the conclusions of highly cited medical studies from prominent journals over 13 years and tracked research on those same topics over time. He found that 16 percent (7 of 45) of the highly cited papers were later contradicted and another 16 percent found stronger effects than subsequent studies. Observational studies were most likely to be later contradicted (5 of 6 observational studies versus 9 of 39 randomized trials). Among randomized trials, the only predictor of contradiction was sample size (the median sample size was 624 among contradicted findings as opposed to 2,165 among validated findings), a finding that supports Ioannidis's conclusions 1 and 2. Sufficient sample size and statistical power are invaluable for reproducibility.

How can we fix the problem of all of this incorrect science in medicine's "evidence base"? With Ioannidis, we have proposed [7] that trial funding for phase-3 studies—

which test the efficacy of medical products—be pooled. Scientific bodies without conflicts of interest should prioritize research questions and design and conduct clinical trials. This recommendation would address several problems in medicine [8, 9] including forcing trials to address basic unanswered questions for common diseases rather than simply advancing the market share of specific products.

Additionally, our proposal would dramatically reduce medical reversal by favoring large-scale randomized trials over countless, scattered lesser ones [10]. Not all RCTs are the same. When RCTs have small sample sizes and financial conflicts of interest and there are multiple studies investigating a single intervention, they are more likely to be in error. (For an illustration of this, look at the number of RCTs that have been conducted for the drug bevacizumab, for which the FDA withdrew its indication for metastatic breast cancer last year [11].)

For observational studies, careful selection of study hypotheses is crucial to increasing the truth and durability of research findings. Hypotheses must be clearly defined prior to data collection and analyses to minimize bias. Krumholz has proposed that such studies document not only the final methods but the history of the methods, accounting for how they were adjusted or changed during the study detailing any and all exploratory analyses [12]. Registration of protocols for observational studies currently remains optional, unlike clinical trial registration, which is required prior to patient enrollment. We agree with suggestions to formally establish a registry of observational analyses [13]. Bias in observational analyses currently presents challenges in reliably basing medical practices on this type of work alone. Only time will tell if this is surmountable.

Finally, Ioannidis's conclusions pertain to concrete ethical choices that budding physicians and researchers make, although they may not see them as such. Our system of medical education and postgraduate medical training rewards the accumulation of publications (abstracts, posters, presentations, and papers) rather than the pursuit of truths. Even students who ultimately pursue private practice careers often engage in research to build their curriculum vitae. Many educators feel that this process is acceptable—better for physicians to gain exposure to research, even if they don't ultimately pursue it, and, anyway, what's the harm? Here, we show the harm. The current system increases the number of publications in a given field while muddying true associations. Ioannidis has criticized medical conferences for promulgating poor science by selecting—on the basis of several-hundred-word abstracts—research that is often not published after more extensive peer review [14]. Students and trainees should consider the inevitable ethical question: if most research findings are false—how vigorously should I advertise my own?

The prevalence of false research findings and their adoption into mainstream practice carries heavy consequences. In our era of soaring health care costs, we cannot afford to implement unnecessary and costly interventions in the absence of sound evidence. More importantly, the use of unfounded interventions puts millions of patients at risk for harm without benefit. Ioannidis's piece reminds us that we have an ethical

responsibility to adhere to high-quality clinical investigation in order to eliminate waste and promote the health and safety of our patients.

References

1. Ioannidis JP. Why most published research findings are false. *PLoS Med*. 2005;2(8):e124.
2. Rossouw JE, Anderson GL, Prentice RL, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the Women's Health Initiative randomized controlled trial. *JAMA*. 2002;288(3):321-333.
3. Boden WE, O'Rourke RA, Teo KK, et al. Optimal medical therapy with or without PCI for stable coronary disease. *N Engl J Med*. 2007;356(15):1503-1516.
4. The ACCORD Study Group, Ginsberg HN, Elam MB, et al. Effects of combination lipid therapy in type 2 diabetes mellitus. *N Engl J Med*. 2010;362(17):1563-1574.
5. Prasad V, Gall V, Cifu A. The frequency of medical reversal. *Arch Intern Med*. 2011;171(18):1675-1676.
6. Ioannidis JP. Contradicted and initially stronger effects in highly cited clinical research. *JAMA*. 2005;294(2):218-228.
7. Prasad V, Cifu A, Ioannidis JP. Reversals of established medical practices: evidence to abandon ship. *JAMA*. 2012;307(1);37-38.
8. Prasad V, Rho J, Cifu A. The diagnosis and treatment of pulmonary embolism: a metaphor for medicine in the evidence-based medicine era. *Arch Intern Med*. 2012;172(12):955-958.
9. Prasad V, Vandross A. Cardiovascular primary prevention: how high should we set the bar? *Arch Intern Med*. 2012;172(8):656-659.
10. Hennekens CH, Demets D. The need for large-scale randomized evidence without undue emphasis on small trials, meta-analyses, or subgroup analyses. *JAMA*. 2009;302(21):2361-2362.
11. D'Agostino RB Sr. Changing end points in breast-cancer drug approval--the Avastin story. *N Engl J Med*. 2011;365(2):e2.
12. Krumholz HM. Documenting the methods history: would it improve the interpretability of studies? *Circ Cardiovasc Qual Outcomes*. 2012;5(4):418-419.
13. Ioannidis JP. The importance of potential studies that have not existed and registration of observational data sets. *JAMA*. 2012;308(6):575-576.
14. Ioannidis JP. Are medical conferences useful? And for whom? *JAMA*. 2012;307(12):1257-1258.

Chetan Huded, MD, is a resident physician in internal medicine at Northwestern Memorial Hospital in Chicago. He was last year's winner of the Rambach Award, given to the most meritorious resident of the intern class. He is interested in cardiology.

Jill Rosno, MD, is a resident physician in internal medicine at Northwestern Memorial Hospital in Chicago. A graduate of Dartmouth Medical School, she is interested in general internal medicine.

Vinay Prasad, MD, is chief fellow in medical oncology at the National Cancer Institute in Bethesda, Maryland. Dr. Prasad coined the term “medical reversal” and has published on the subject in the *Journal of the American Medical Association* and *Archives of Internal Medicine*, among others. He is interested in the adoption of rational methods in medical practice.

Related in VM

A Call to Integrate Ethics and Evidence-Based Medicine, January 2013

Teaching Critical Appraisal of Medical Evidence, January 2013

Rating Evidence in Medical Literature, January 2011

Virtual Mentor
American Medical Association Journal of Ethics
January 2013, Volume 15, Number 1: 34-41.

STATE OF THE ART AND SCIENCE

Effective Ways to Communicate Risk and Benefit

Evan A. Wilhelms and Valerie F. Reyna, PhD

Avoiding misunderstanding of health risk and benefit communications is a major challenge in medicine and public health [1]. Everyone is capable of reading or hearing every word presented, understanding each one, and yet grasping nothing of the meaning of the message [2]. Thus, it is critical that health messages are designed with an understanding of how people process health information and consequently make medical decisions.

Here, we describe some evidence-based recommendations for communicating about health risks. In doing so, we describe how people—even expert physicians—can misinterpret test results and misjudge probabilities of illness or other outcome [3]. We also consider theory-driven interventions that reduce these errors.

Recommendations for Successful Risk and Benefit Communication

Don't stop at the numbers. Consider the example of consenting to a surgical procedure (e.g., carotid endarterectomy) that carries with it a 2 percent chance of life-threatening complications [4]. A patient who consented to this procedure and then later recalled a 0 percent risk of complication would be reflecting greater verbatim accuracy than a patient who recalled 10 percent risk of complication, although the former represents a fundamental misunderstanding of the risks involved in the procedure. This misunderstanding is a demonstration of a failure to get the gist of the message, that there is some risk involved. Compared to the patient who recalled a 10 percent risk of complication—who understands that there is a surgical risk—the patient who recalls no risk is unable to give a truly informed consent.

The example illustrates that getting the gist is the essential element of informed consent. This is critical because gist representations of information—vague, qualitative representations that capture the meaning—are relied on in reasoning and decision making, in contrast to representations of verbatim facts that do not affect reasoning accuracy [1]. Despite the fact that verbatim and gist representations are understood in parallel, adults tend to rely on gist representations to make health and other decisions. For decades, decision-support tools for patients have been developed to guide decisions for costly and prevalent health problems [5, 6], and many researchers have assumed that providing calculators and emphasizing calculation of numerical values should improve health-related judgments and decisions [7]. However, even those who demonstrate the ability to understand and work with numbers process and decide based on categorical (e.g., good/bad) or ordinal (e.g., high-, medium-, and low-risk) gists.

Give reasons for facts. It is difficult to understand gists when relevant context is unavailable, as is often the case with novice patients (who are not experts concerning their diagnoses). For example, understanding that HIV/AIDS is a virus may help in understanding that, like other STIs that are viral, HIV/AIDS is also incurable. Giving the reasons helps useful information endure in memory beyond just the verbatim facts [8].

Begin with the message in mind. Before communicating risk to a patient, identify the "bottom line" [1]. This can include translating numerical facts into meaningful messages, as well as explaining the reasons behind these messages. Bearing in mind the bottom-line message is also critical when considering the possible effects of framing communications positively or negatively. People tend to be risk-averse when situations are described as gains (e.g., in terms of survival or other benefits), and risk-seeking when situations are described as losses (e.g., in terms of death or other harms). Consequently, communications are more effective when framed as gains if the patient is truly in a gain frame and stands to benefit compared to the status quo (for example, if the goal behavior is health-promoting). Loss frames can be more appropriate if the patient truly stands to lose (if, for example, there is uncertainty and risk of negative outcome such as death or reduction of quality of life) [9].

Use graphs that highlight the gist. Where possible, use graphical presentations of risk and benefit information with patients. Different graphical formats can be used to highlight different relationships. For example, to illustrate a monotonic trend, such as a survival or mortality curve or the effectiveness of a drug over time, line graphs are typically best because the gist of the trend is automatically understood. This is so because patients tend to ignore the numbers in favor of the relationship expressed—that the magnitude is going up or down [1].

These same principles can be applied to other graphs. Stacked bar graphs can be used to demonstrate absolute risks—for example, in comparing the frequencies of a given outcome among the total of those treated—by drawing attention to the denominator (i.e., the total treated, in this case). Simple bar charts work better for conveying relative risk, given that small differences can be highlighted. These emphases should be used for specific purposes. For example, it's better to use relative risks when the decision is between options with everything else being equal, or to highlight differences between options; absolute risk better conveys the essential meaning that risks do not differ significantly. Fraenkel et al. [10] demonstrated that pie charts could be effectively used to demonstrate frequencies of adverse drug side effects when the risk was greater than 1 percent, and icon-based pictographs for adverse events with a smaller than 1 percent risk, effectively communicating the essential gist of complex information. These display methods are illustrated in figures 1a-1d.

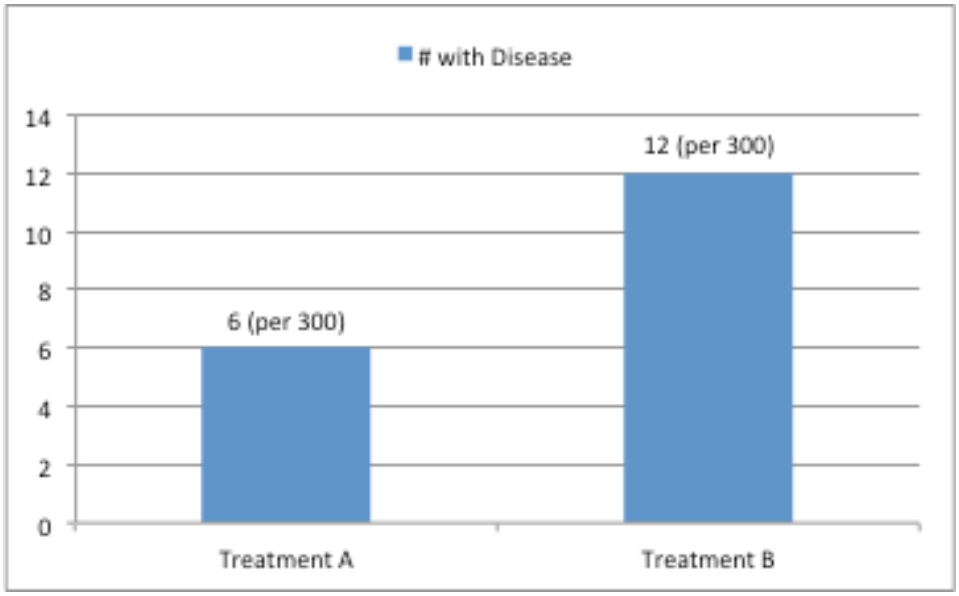

Figure 1a. Differences of relative risk (e.g., conveying that one treatment has increased adverse effects relative to another) can be best displayed in bar charts, in which the frequencies of disease outcomes for two treatments are indicated.

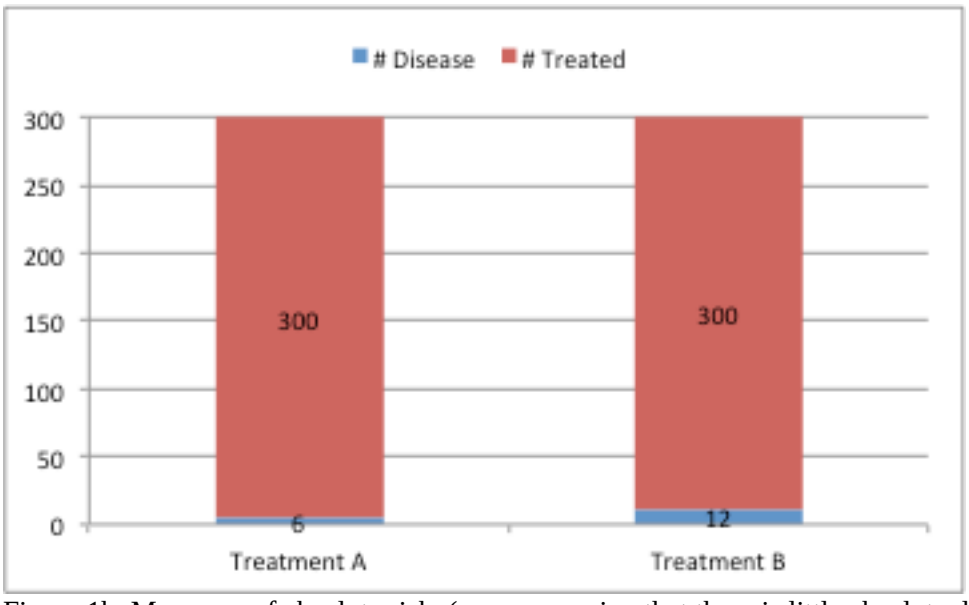

Figure 1b. Messages of absolute risk, (e.g., conveying that there is little absolute difference in effectiveness between two treatments) can be displayed using stacked bar charts, in which the frequencies of disease outcomes are displayed among the total treated.

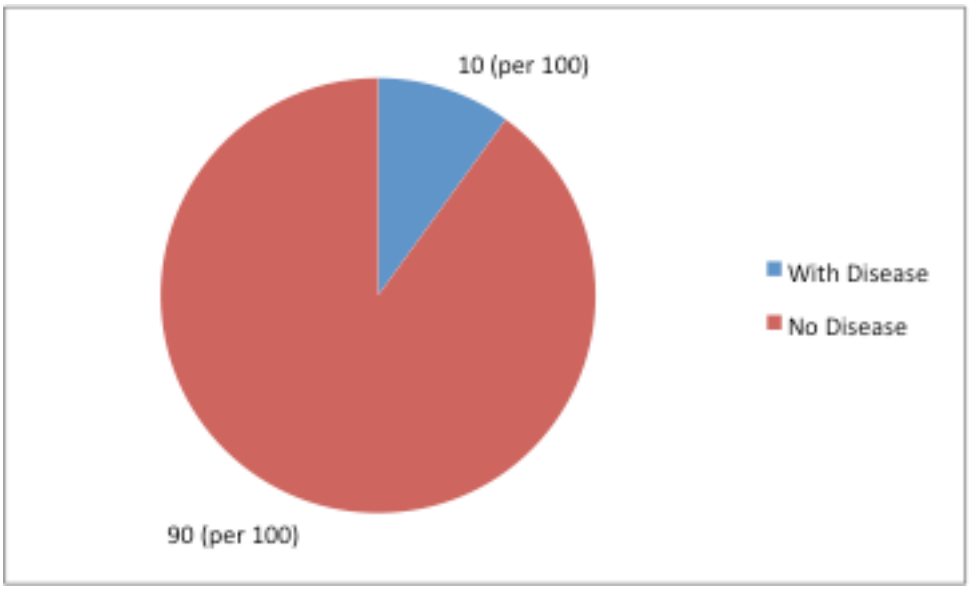

Figure 1c. Incidents of adverse events that occur more than 1 percent of the time are effectively communicated through pie charts.

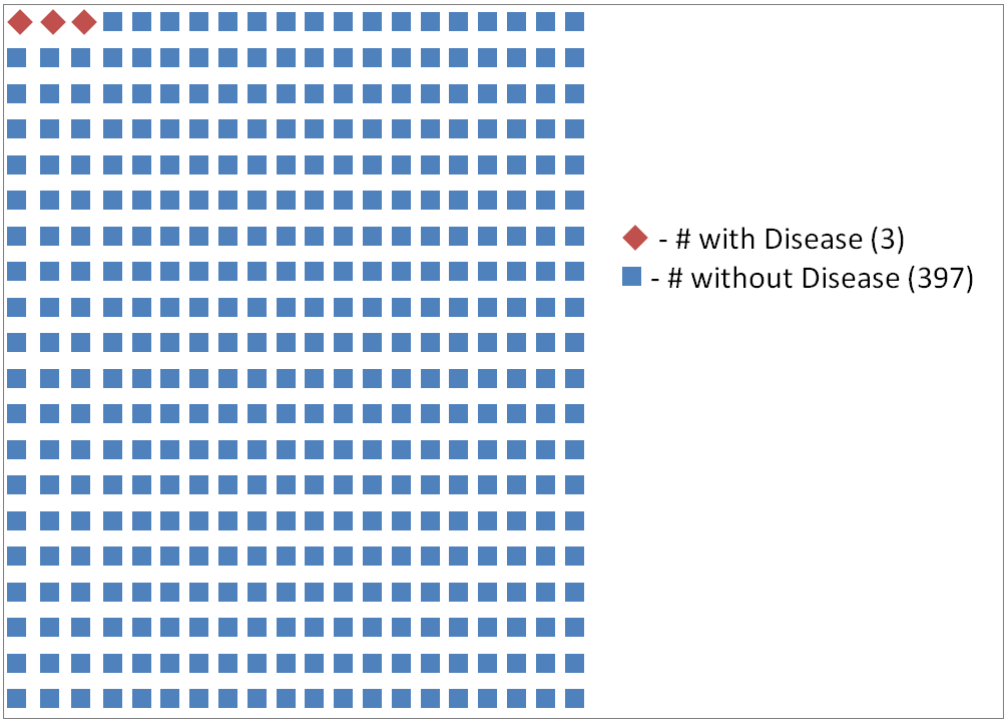

Figure 1d. Events that occur less than 1 percent of the time are effectively communicated through icon-based pictographs.

Remember that even experts are susceptible to reasoning errors. Errors in understanding among medical students, physicians, specialists and subspecialists can cause misjudgment of probabilities of risks and benefit [11-13]. Usually, these errors are not the result of lack of knowledge (specialized medical knowledge) or bad logic

[14]. For example, physicians and other health care professionals in one study were asked to predict the likelihood that a patient who tested positive for an unknown disease actually had the disease, given that the disease had a 10 percent base rate in the population and that the test had 80 percent sensitivity and 80 percent specificity, which were defined for them [15]. (Sensitivity is the probability that, if you truly had a disease, you would test positive for it, and specificity is the probability that, given an absence of a disease, you would test negative for it.) Participants merely had to select whether the correct likelihood was closer to 30 percent or 70 percent. Only 31 percent of physicians selected the correct answer, and the group of health care professionals that fared best—public health experts—only scored around chance (55 percent). Moreover, high school students selected the correct answer at roughly the same rate as physicians (33 percent [3]), demonstrating that this sort of processing error is not related to medical expertise. However, the fact that physicians' judgments were below the level of chance suggest that this was not merely an inability to reason correctly; it appeared to result from a systematic bias. Now that this susceptibility is identified, it can be addressed using verbal and visual communications, as we describe below.

Explain all combinations of test results and disease. The same principle that explains the previous error in interpreting diagnostic tests—that people confuse nested classes such as sensitivity and posttest probability—applies when probabilities must be combined to make diagnostic judgments, as in conjunction ("and") or disjunction ("or") judgments. This was illustrated by a study in which physicians were required to make diagnostic judgments of the probability that hypothetical patients had coronary artery disease (CAD), an imminent risk of myocardial infarction (MI), one or the other, or both [12]. The physicians in this group were all vulnerable to disjunction errors (ranging from 20-30 percent of judgments) in which the combined probability that the hypothetical patient had either CAD or imminent risk of MI was judged to be lower than the probability that the hypothetical patient had CAD or risk of MI individually (though specialists were better at discriminating high and low risk based on the hypothetical patient descriptions).

An evidence-based theory of reasoning, fuzzy trace theory, explains this as a result of confusion of overlapping classes, as in the class of having CAD but no risk of MI, of having risk of MI but not having CAD, and of both being at risk of MI and having CAD. Consequently, this is an *advanced* error that isn't a reflection of lack of knowledge. Yet, the susceptibility to this error is relevant to clinical practice.

This sort of error can be effectively addressed by using as a visual intervention a 10 x 10 grid in which each square represents a woman with potential illness. The grid demonstrates pretest information such as the base rate chance of the potential diagnoses (e.g., CAD and MI), as well posttest information expected based on the sensitivity and specificity of the diagnostic test [16]. This intervention is illustrated in figure 2. Using this intervention, diagnostic errors in posttest probability estimation were reduced. Moreover, this intervention performed better than the use of Bayesian clinical calculators, despite the participants' being taught how to use

Bayes's theorem. This contradicts the assumption that calculation of exact numbers should improve health-related judgments [6]. This visual intervention was successful because it represents each class discretely and accounts for all relevant classes, allowing for visual estimation and a reduction of interference from overlapping classes and thus accurate estimation of probability of disease given a positive or negative test result.

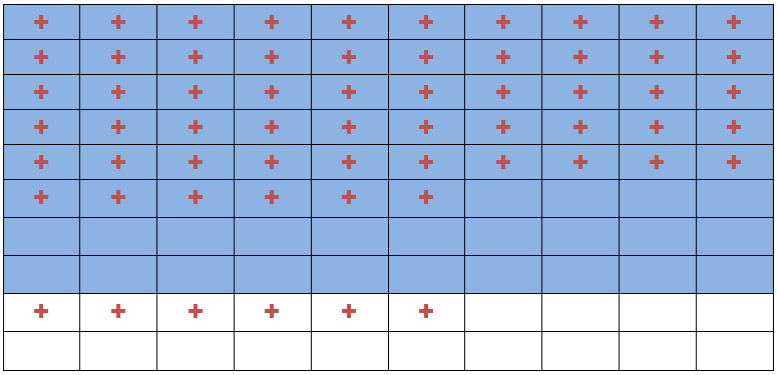

Figure 2. A sample grid indicating all posttest probabilities, indicating a base rate pretest probability of 0.8, a test sensitivity of 0.7, and a test specificity of 0.7. Patients with disease are represented with blue cells, while patients with no disease are represented with white cells. Patients with a positive test are represented with a red +. (Adapted from Lloyd & Reyna [16]).

Combining Communication Strategies for Better Patient Care

Using these evidence-based communications methods can improve judgments of risks and benefits of illness and treatment in clinical care. The same principles have been applied in many settings—to reduce unhealthy risk-taking in adolescence, to plan for and mitigate cognitive declines that occur with aging, and to provide a model of vaccination decisions in the proliferation of antivaccination messages, to name a few examples [13, 17, 18]. In all of these settings, informed medical decision making is dependent on communicating an essential bottom line: which option is more likely to provide life rather than death, relief rather than suffering, mobility rather than disability.

These conclusions about communicating health risks and benefits also illustrate that evidence-based theory is critical to designing interventions with practical applications. Relevant cognitive theories should be used to improve risk and benefit communication in relevant medical contexts. Effective communication provides a link between research and good medical outcomes.

References

1. Reyna VF. A theory of medical decision making and health: fuzzy trace theory. *Med Decis Making*. 2008;28(6):850-865.

2. Wilhelms EA, Reyna VF. Fuzzy trace theory and medical decisions by minors: Differences in reasoning between adolescents and adults. *J Med Philosophy*. In press.
3. Reyna VF. How people make decisions that involve risk: a dual process approach. *Curr Directions Psychol Sci*. 2004;13(2):60-66.
4. Reyna VF, Hamilton AJ. The importance of memory in informed consent for surgical risk. *Med Decis Making*. 2001;21(2):152-155.
5. Reyna VF. Theories of medical decision making and health: an evidence-based approach. *Med Decis Making*. 2008;28(6):829-833.
6. Bransford JD, Johnson MK. Contextual prerequisites for understanding: Some investigations of comprehension and recall. *J Verbal Learning and Verbal Behav*. 1972;11:717-726.
7. Reyna VF, Nelson WL, Han PK, Dieckmann NF. How numeracy influences risk comprehension and medical decision making. *Psychol Bull*. 2009;135(6):943-973.
8. Reyna VF, Mills BA, Estrada SM. Reducing risk taking in adolescence: effectiveness of a gist-based curriculum. Paper presented at the 30th Annual Meeting of the Society of Medical Decision Making; October 2008; Philadelphia, PA.
9. Fischhoff B, Brewer NT, Downs JS. Communicating risks and benefits: an evidence-based user's guide. Food and Drug Administration; 2011. http://www.fda.gov/downloads/AboutFDA/ReportsManualsForms/Reports/UCM268069.pdf. Accessed December 11, 2012.
10. Fraenkel L, Peters E, Charpentier P, et al. A decision tool to improve the quality of care in rheumatoid arthritis. *Arthritis Care Res*. 2012;64(7):977-985.
11. Reyna VF, Lloyd F, Whalen P. Genetic testing and medical decision making. *Arch Intern Med*. 2001;161(20):2406-2408.
12. Reyna VF, Lloyd FJ. Physician decision making and cardiac risk: effects of knowledge, risk perception, risk tolerance, and fuzzy processing. *J Exp Psychol Learn Mem Cog*. 2006;12(3):179-195.
13. Reyna VF, Brainerd CJ. Numeracy, ratio bias, and denominator neglect in judgments of risk and probability. *Learn Indiv Diff*. 2008;18(1):89-107.
14. Reyna VF, Lloyd FJ, Brainerd CJ. Memory, development, and rationality: an integrative theory of judgment and decision-making. In: *Emerging Perspectives on Judgment and Decision Research*. New York: Cambridge University Press; 2003: 201-245.
15. Reyna VF, Adam MB. Fuzzy-trace theory, risk communication, and product labeling in sexually transmitted diseases. *Risk Anal*. 2003;23(2):325-342.
16. Lloyd FJ, Reyna VF. A web exercise in evidence-based medicine using cognitive theory. *J Gen Intern Med*. 2001;16(2):94-99.
17. Reyna VF. Risk perception and communication in vaccination decisions: a fuzzy-trace theory approach. *Vaccine*. 2012;30(25):3790-3797.
18. Reyna VF, Estrada SM, DeMarinis JA, et al. Neurobiological and memory models of risky decision making in adolescents versus young adults. *J Exp Psychol Learn Mem Cog*. 2011;37(5):1125-1142.

Further Reading

Wolfe CR, Fisher CR, Reyna VF. Semantic coherence and fallacies in estimating joint probabilities. *J Behav Decis Making*. 2010;23(2):203-223.

Brust-Renck PG, Reyna VF, Wilhelms EA, Lazar AN. A fuzzy-trace theory of judgment and decision making in healthcare: explanation, prediction, and application. In: Diefenbach, MA, Miller SM, Bowen DJ, eds. *Handbook of Health and Decision Science*. New York: Springer. In press.

Evan A. Wilhelms is a PhD student in the Department of Human Development at Cornell University in Ithaca, New York. His research interests have focused on models of decision making in cognitive development and behavioral economics.

Valerie F. Reyna, PhD, is a professor at Cornell University and Weill Cornell Medical College and co-director of Cornell's University Magnetic Resonance Imaging Facility and its Center for Behavioral Economics and Decision Research. Her recent work has focused on numeracy, medical decision making, risk communication, risk taking, neurobiological models of decision making, and neurocognitive impairment.

Related in VM

Discounting a Surgical Risk: Data, Understanding, and Gist, July 2012

When Patients Seem Overly Optimistic, July 2012

The Message Isn't as Mean as We May Think, January 2013

Virtual Mentor
American Medical Association Journal of Ethics
January 2013, Volume 15, Number 1: 42-45.

STATE OF THE ART AND SCIENCE

N-of-1 Trials: Individualized Medication Effectiveness Tests

Sunita Vohra, MD, MSc, and Salima Punja, BSc

Evidence-based management of chronic diseases presents a unique set of challenges. Results from randomized controlled trials (RCTs), often considered the "gold standard" of research evidence, are often not well suited to the realities of clinical practice given patient heterogeneity, comorbidities, and the use of multiple concurrent therapies. In fact, RCTs may exclude the majority of patients seen in routine clinical practice [1]. In addition, evidence is lacking on the long-term effectiveness, comparative effectiveness, and additive effectiveness of many therapies for chronic conditions. This lack of relevant evidence can limit a clinician's ability to make evidence-based decisions.

In the absence of relevant research evidence to inform clinical care, routine clinical decisions are often made based on informal "trials of therapy" during which a patient tries various therapies over the course of multiple visits with a physician and continues with the one that seems to help the most. These so-called trials are unblinded, have no control, and involve no formal assessment of effectiveness, making them vulnerable to invalid conclusions about treatment response. N-of-1 trials straddle the divide between RCTs and trials of therapy by providing scientific rigor and an individualized approach to patient care simultaneously.

N-of-1 trials are multiple crossover trials conducted on single individuals [2]. They entail making *a priori* decisions about when and how outcomes will be measured. They have been widely used in psychology and other social sciences. Research on adopting them in clinical medicine was first introduced in 1986 by Gordon Guyatt and David Sackett [3]. Advantages of n-of-1 trials are that they:

1. Offer direct evidence about treatment benefit to a particular patient rather than the population-level outcomes yielded by RCTs, which may or may not be applicable to a specific individual;
2. Allow assessment of long-term therapy in chronic conditions to determine if treatment effectiveness is continuing to be achieved;
3. Can be used to establish comparative and additive treatment effectiveness for patients with various comorbid conditions and using concurrent therapies; and
4. Help reduce ineffective polypharmacy and thus promote patient safety by limiting therapies to those with demonstrated effectiveness.

N-of-1 trials allow for treatment evaluation in understudied populations, such as those with comorbid conditions and rare diseases, and in difficult-to-study fields

such as pediatrics or palliative care. Researchers and clinicians have used n-of-1 trials to evaluate interventions for a range of clinical conditions, including those in the neurological [4], behavioral [5], rheumatologic [6], pulmonary [7], and gastrointestinal [8] domains.

The success of an n-of-1 trial depends largely on the collaboration and commitment of the patient and his or her doctor. Clinicians must explain the process to their patients and collaborate with them to develop individualized outcome measures. Clinicians must also monitor them at regular intervals throughout the trial period, then evaluate and interpret what the results of the trial mean to that individual.

Generally, an n-of-1 trial involves multiple repetitions of 2 treatment options. These options often consist of: (a) active treatment and placebo, (b) low-dose and high-dose active treatment, or (c) active treatment A and active treatment B. The sequence of treatments can be randomized. The number of treatment pairs a patient undergoes does not necessarily have to be predetermined; however, conclusions of treatment effectiveness (or lack thereof) are less subject to bias if the number is specified in advance. Alternatively, treatment pairs can be replicated until the physician and patient are convinced that the treatment is effective, harmful, or has no effect. The length of treatment period depends on the amount of time it takes for a treatment to reach full effect and to cease effect after discontinuation; therefore, interventions with a quick onset and offset are most efficient for n-of-1 evaluation. Responses to treatment must be measured at least once during each period.

Clinicians and patients should determine which symptoms are most relevant prior to commencing the trial. Target outcomes form the basis of disease- and patient-specific questionnaires. A useful questionnaire, known as the Measure Yourself Medical Outcome Profile (MYMOP), allows patients to identify the most troubling symptoms or problems they would like alleviated with the treatment and score how they are feeling daily on a 7-point scale [9]. Analysis of the data can be visual (i.e., presented in a graph) or statistical (such as making comparisons using a paired-t test).

Imagine, for example, the parents of a 12-year-old boy diagnosed with attention deficit/hyperactivity disorder come into his physician's office concerned about his behavior. The patient has been taking 10 mg of methylphenidate daily for 2 years, and the physician suspects that the child may no longer be responsive to this dose or, perhaps, the drug at all. The physician decides the best course of action is to conduct a head-to-head comparison n-of-1 trial (listed above as option 3), in which the patient undergoes randomly alternating weeks of 20 mg/day methylphenidate and 10 mg/day dexamphetamine for 6 weeks. The physician tells the parents that neither he nor they will be aware of which treatment the child will be on each week. The parents and the child discuss and decide that symptoms of hyperactivity and irritability have been the most troublesome over the past few months. The parents are instructed to monitor these target symptoms, rate their severity each day using the MYMOP, and, if the child complains of any side effects throughout the 6-week trial,

to bring him into the office immediately. After the 6 weeks, the physician graphs the results of the MYMOP, explains the results to the parents, and together they decide how to proceed.

Despite the numerous advantages offered by n-of-1 trials and their potential to provide the strongest evidence for individual treatment decisions, it is reasonable to ask whether the use of randomization, blinding, and placebos requires that n-of-1 trials be considered research and therefore necessitates the various processes that this title would imply. An ethical assessment to the use of n-of-1 trials should be based on the trial's purpose: Is the trial intended to be research or clinical care? In research, the goal is to find an answer for a particular condition with the hope of having some generalizable result, and therefore any benefit gained by individual participants is secondary. In clinical care, however, the primary goal is to find an answer about treatment effectiveness for the individual patient. The two activities are fundamentally different in their intent, and therefore require different ethical considerations. In the above example, the n-of-1 trial is being used strictly for clinical reasons; the patient's health and well-being are of primary interest.

N-of-1 trials provide a structured, objective, and evidence-based framework for evaluating treatment effect on an individual. They can provide patients with *optimal* clinical care—care that is personalized and evidence-based—and reduce the potential for bias seen in routine clinical care.

References

1. Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?" *Lancet*. 2005;365(9453):82-93.
2. Guyatt G, Keller J, Jaeschke R, Rosenbloom D, Adachi JD, Newhouse MT. The n-of-1 randomized controlled trial: clinical usefulness. Our three-year experience. *Ann Intern Med.* 1990;112(4):293-299.
3. Guyatt G, Sackett D, Taylor D, Chong J, Roberts R, Pugsley S. Determining optimal therapy. Randomized trials in individual patients. *N Engl J Med*. 1986;314(14):889-892.
4. Forman AC, Vasey PA, Lincoln NB. Effectiveness of an adjustment group for brain injury patients: a pilot evaluation. *Int J Ther Rehabil*. 2006;13(5):223-228.
5. Nikles JN, Mitchell GK, Del Mar CB, Clavarino A, McNairn N. An n-of-1 trial service in clinical practice: Testing the effectiveness of stimulants for attention deficit/hyperactivity disorder. *Pediatrics*. 2006;117(6):2040-2046.
6. Nikles CJ, Yelland M, Glasziou PP, Del Mar C. Do individualized medication effectiveness tests (N-of-1 trials) change clinical decisions about which drugs to use for osteoarthritis and chronic pain? *Am J Ther.* 2005;12(1):92-97.
7. Smith BJ, Appleton SL, Veale AJ, McElroy HJ, Veljkovic D, Saccoia L. Eformoterol n-of-1 trials in chronic obstructive pulmonary disease poorly reversible to salbutamol. *Chron Respir Dis*. 2004;1(2):63-69.

8. Lashner BA, Hanauer SB, Silverstein MD. Testing nicotine gum for ulcerative colitis patients: Experience with single-patient trials. *Dig Dis Sci*.1990;35(7):827-832.
9. Paterson C. Measuring outcome in primary care: a patient-generated measure, MYMOP, compared to the SF-36 health survey. *BMJ*. 1996;312(7037):1016-1020.

Further Reading

Guyatt G, Sackett D, Adachi J, et al. A clinician's guide for conducting randomized trials in individual patients. *CMAJ*. 1988;139(6):497-503.

Guyatt G, Heyting A, Jaeschke R, Keller J, Adachi JD, Roberts RS. N-of-1 randomized trials for investigating new drugs. *Control Clin Trials*. 1990;11(2):88-100.

Larson EB. N-of-1 trials: A technique for improving medical therapeutics. *West J Med*. 1990;152(1):52-56.

Sunita Vohra, MD, MSc, a pediatrician and clinician scientist with a master's degree in clinical epidemiology and fellowship training in clinical pharmacology, is the leader of a Canadian Institutes of Health Research project investigating n-of-1 trials. A professor in the Faculty of Medicine and School of Public Health at the University of Alberta in Edmonton, Dr. Vohra is the founding director of Canada's first academic pediatric integrative medicine program, the Complementary and Alternative Research and Education (CARE) program at the Stollery Children's Hospital. Dr. Vohra is also the program director for Canada's first fellowship program in pediatric integrative medicine and the founding director of the Canadian Pediatric CAM Network (PedCAM).

Salima Punja, BSc, is a PhD candidate in the Department of Medicine at the University of Alberta in Edmonton.

Virtual Mentor
American Medical Association Journal of Ethics
January 2013, Volume 15, Number 1: 46-50.

HEALTH LAW
Medicine, the Law, and Conceptions of Evidence
Valarie Blake, JD, MA

Evidence-based medicine (EBM) calls for doctors to use the best available scientific evidence on safety, effectiveness, and cost in treating their patients [1]. While scientific evidence is increasingly influencing physician practice, courts traditionally have focused on physicians' "customary practice"—defined by what *most* physicians in the local geographic area or specialty *usually* do—in defining the legal standard for how a physician ought to practice. EBM and clinical practice guidelines are driving the need for a new approach in the judicial system that takes into account the role of evidence in medical malpractice and other legal claims implicating physicians [2]. For an analysis of the impact of EBM on medical malpractice and tort reform specifically, see the January 2011 *Virtual Mentor* health law piece [3].

This article will focus on special implications of EBM for the courts more generally, both the likelihood and challenges of widespread uptake into legal practice and the key distinctions in how doctors and lawyers consider the meaning of evidence. How the courts react to EBM will have great impact on how doctors use it. If courts embrace EBM and use scientific evidence as the standard for how a physician ought to have acted in a given case, physicians are more likely to adopt EBM practices to avoid liability [2]. Alternatively, if courts largely ignore EBM or fail to admit it into the court's body of evidence, physicians who choose to practice EBM may risk liability when their practice no longer mirrors customary practice [2]. Thus courts must take on the challenge of defining the role and scope of EBM in legal practice, which will consequently shape how EBM is used in medical practice.

Use of Evidence-Based Medicine in the Courts
Courts and health care have historically viewed evidence in fundamentally disparate ways, which may confound the legal system as it moves toward adopting EBM. The legal system is built on an adversarial model which "recognizes that evidence and even facts are disputable, that experts may disagree, and that therefore there is a political element to interpreting evidence" [1]. Disputes about fact are left to a jury or judge to decide, and the goal is to ensure fair process rather than fair outcomes—or truth [1]. Juxtaposed to this, empirical evidence in medicine seeks to define a single unimpeachable truth that can stand on its own. Moreover, medical evidence often focuses on populations, while at the court level, the evidence must be relevant to the single injured patient [4].

Prior to EBM, courts often sought medical expertise in the form of expert witnesses for four key questions in a lawsuit: the applicable standard of care or what the

physician ought to have done, causation or the connection between the wrongful conduct and the harm to the patient, assessment of damages or how much money is needed to adequately compensate the patient's injuries, and medical prognosis [5]. Such testimony has often depended on the expert's subjective analysis, and one lawsuit often may involve several experts who disagree and present opposing testimony on each side. EBM presents the possibility of empirical, objective guidance that can be used in addition to the information given by expert witnesses [5].

One significant question is how courts will allow the admission of such information, if at all. While traditionally plaintiffs and defendants have both been able to present expert testimony, courts might consider allowing only one side to use EBM. It might only be allowable by a physician or a health plan defending its decisions or, alternatively, only by the injured plaintiff as proof that the defendant failed to meet a standard of care [5]. Alternatively, some courts may not even allow EBM if the judge deems it hearsay, a statement made outside of the court while the speaker was not under oath [6]. The rationale for this is that the speaker is not available to be cross-examined by the opposing side.

In the 2006 *Hinlicky v. Dreyfuss* case, a New York court allowed a guideline from the American Heart Association and American College of Cardiology to be admitted because, according to the court, the purpose was to show the physician's decision-making methodology rather than to establish the standard of care [7]. The specific use allowed in this case says little about whether other courts would permit such guidelines as direct evidence of a standard of care or bar it as hearsay.

Another major question is what type of legal evidence can be admitted into a court and who should decide its quality. Courts have been critiqued in the past for allowing "junk science" to control such cases as silicon breast implant and drug and medical device litigation [5]. Based on the Supreme Court case *Daubert v. Merrell Dow*, a trial judge must prescreen all scientific evidence that is introduced to the court and admit only that which will inform the jury, not confound it [8]. Judges might consider the amount, complexity, and consistency of the evidence in deciding whether it well help or hurt the jury in reaching a decision, and the judge, through this standard, is the ultimate arbiter of whether certain evidence gets admitted. If admitted, there is the additional question of how much weight such evidence will carry: Does it equal, outweigh, or carry less weight than an opposing statement made by a single expert?

EBM and guidelines are applicable to two different types of legal cases: (1) medical malpractice lawsuits in which the injured patient seeks to recover monetary damages for an injury allegedly caused by a physician or other members of the health care team and (2) cases of patients contesting insurance coverage decisions of a health plan.

Malpractice. Traditionally courts have looked to customary practice to determine whether a physician fell below the standard of care in a medical malpractice case [2]. Expert witnesses may be sought from both sides to introduce evidence for the jury or the judge about the customary norm and whether the doctor acted in accordance with it [2].

A significant problem with a custom-based standard is that it is not necessarily good patient care, but merely what that has traditionally occurred. Some courts have moved away from a customary practice standard to a judicial risk-benefit analysis [2]. In *Helling v. Carey*, the court found two ophthalmologists negligent for failing to screen a patient under the age of 40 for glaucoma [9]. While such screening was not the customary practice at that time, the court found in the patient's favor because the risk of the glaucoma screening was minimal compared to the benefit of prevention [9].

Other courts have used a reasonable physician standard, which holds physicians to the same standard as parties in nonprofessional negligence suits—did the physician act with an ordinary or reasonable level of care in his or her dealings with the injured party—a fact-finding issue determined by the judge or jury [2].

Will or should courts use EBM to determine whether a physician in a medical malpractice suit has met the standard of care? Some EBM may get so well incorporated into medical practice that it becomes the custom, and some use of EBM in practice may satisfy either a reasonable physician standard or a risk-benefit analysis. The real challenge comes when a doctor defends an action by saying it is in line with EBM, even if it does not meet any of the above traditional standards.

Insurance coverage. Another potential use of EBM in the courts relates to cases which contest the benefits covered by a given health plan [5, 10]. Here the claim is against a health care organization or plan, rather than an individual physician. Patients may bring suit against their health plans, arguing that a given treatment is covered under the health plan contract, a claim that happens with frequency given the extremely vague language that often defines coverage, e.g., defining coverage as that which is "reasonable and necessary," "appropriate," "investigational," or "generally accepted medical practice" [10].

As in malpractice claims, EBM might provide a way to explore the empirical questions related to some cases—for example a plan's claim to cover only a generic drug based on data comparing the effectiveness of the generic and a brand name drug [11]. Other cases may involve difficult trade-offs that are normative, not empirical, in nature—for example, a decision by a plan not to cover a mammography for women under age 40 because, while beneficial, its benefits do not outweigh those of treatments the plan prioritizes more highly [11].

Challenges to Applying EBM in the Courts

Whether EBM has the potential to improve health care litigations or not, the shift in practice towards EBM makes it an unavoidable challenge that courts must face. Educational gaps, limitations of the evidence itself, and fundamental distinctions between the meaning of evidence in medicine and law all raise particular challenges for courts addressing EBM.

As the *Daubert* standard states, it is the role of the judge to decide whether a particular piece of scientific evidence should be admitted. Yet, understanding complex scientific evidence (and even more significantly, distinguishing quality studies from poor studies) is not a skill in which judges are appropriately trained. Because of the sheer quantity, diversity, and complexity of the evidence, adequate interpretation has been viewed as difficult even for physicians, let alone lawyers and juries, who frequently have less expertise than physicians in empirical questions, research methods, and health care [12]. Moreover, while clinical practice guidelines may serve to consolidate and apply much of the evidence, they are not always available, may be outdated, or may conflict with each other. Such cases may place the judge or jury in the difficult position of deciding which professional society or research group has more sound evidence—a political hotbed that they are ill-equipped and probably unwilling to traverse [5].

A transition by the courts to applying EBM in medical malpractice and insurance claim cases will inevitably confront and challenge much of the tradition in both medicine and law, but such a transition will be necessary to adapt to a changing practice of medicine in which physicians are expected to use an evidence base in treating their patients.

References

1. Rodwin MA. The politics of evidence-based medicine. *J Health Polit Policy Law*. 2001;26(2):439-446.
2. Williams CL. Evidence-based medicine in the law beyond clinical practice guidelines: what effect will EBM have on the standard of care? *Washington & Lee Law Rev* 2004;61(1):479-533.
3. Mackey TK, Liang BA. The role of practice guidelines in medical malpractice litigation. *Virtual Mentor*. 2011;13(1):36-41.
4. Mulrow CD, Lohr KN. Proof and policy from medical research evidence. *J Health Polit Policy Law*. 2001;26(2):249-266.
5. Rosoff AJ. Evidence-based medicine and the law: the courts confront clinical practice guidelines. *J Health Polit Policy Law*. 2001;26(2):327-368.
6. Rosoff AJ. The role of clinical practice guidelines in healthcare reform: an update. *Ann Health Law*. 2012;21:21-33.
7. *Hinlicky v Dreyfuss*, 848 NE2d 1285 (NY 2006).
8. *Daubert v Merrell Dow Pharmaceuticals*, 509 US 579 (1993).
9. *Helling v Carey*, 519 P2d 981 (Wash 1974).

10. Eddy DM. The use of evidence and cost effectiveness by the courts: how can it help improve health care? *J Health Polit Policy Law*. 2001;26(2):387-408.
11. Morreim EH. From the clinics to the courts: the role evidence should play in litigating medical care. *J Health Polit Policy Law*.2001;26(2):409-428.
12. Eisenberg JM. What does evidence mean? Can the law and medicine be reconciled? *J Health Polit Policy Law*. 2001;26(2):369-382.

Valarie Blake, JD, MA, is a senior research associate for the American Medical Association Council on Ethical and Judicial Affairs in Chicago. Ms. Blake completed the Cleveland Fellowship in Advanced Bioethics, received her law degree with a certificate in health law and concentrations in bioethics and global health from the University of Pittsburgh School of Law, and obtained a master's degree in bioethics from Case Western Reserve University. Her research focuses on ethical and legal issues in assisted reproductive technology and reproductive tissue transplants, as well as regulatory issues in research ethics.

Related in VM
The Role of Practice Guidelines in Medical Malpractice Litigation, January 2011

Virtual Mentor
American Medical Association Journal of Ethics
January 2013, Volume 15, Number 1: 51-55.

POLICY FORUM

Evidence-Based Guidelines and Quality Measures in the Care of Older Adults

Erika Ramsdale, MD, and William Dale, MD, PhD

Discussion of health care utilization, costs, and quality continues to pervade the public and political discourse. The basic message is that we have an unsustainable, dysfunctional system, with costs spiraling out of control and a poor return on our investment. Among 19 industrialized countries, the United States has the highest rate of preventable deaths despite the highest per capita expenditure on health care, mostly attributable to poorly controlled chronic conditions [1]. Health care reform initiatives aim to simultaneously increase access, improve quality, and contain costs. The most recently enacted iteration, the Patient Protection and Affordable Care Act of 2010, extends insurance coverage, promises to lower costs through the work of the Independent Payment Advisory Board, and (beginning in 2015) to cut payments to providers who do not report on selected quality measures.

In any year, approximately 10 percent of Americans account for more than two-thirds of health care costs [2]. Forty percent of this group are older than 65, and many of them have multiple chronic conditions. For these older patients with complex conditions and multimorbidity, how should "quality care" be defined? Is it possible to standardize quality measures for these patients based on the best current evidence? Should the measures be linked to provider incentives? Will taking these steps improve outcomes and quality of life for this group of patients?

What Counts as Evidence?

The dominant paradigm guiding current medical practice is evidence-based medicine (EBM) [3]. The gold standard for "evidence" in EBM is the blinded, prospective randomized clinical trial (RCT), which eliminates bias by: (1) random assignment of participants to interventions to control for confounding variables, and (2) concealing outcomes from data evaluators to allow them to assess the efficacy and effectiveness of interventions objectively. EBM based on RCTs has largely supplanted other forms of clinical evidence, such as individual expertise, anecdotal case series, case-control studies, and observational cohort studies. Ideally, EBM advocates the "conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients" [4]. Indeed, guidelines are preferentially based upon evidence from prospective RCTs and represent a systematic application of EBM to practice.

Detractors maintain that this approach encourages a defensive, algorithmic medical practice in an attempt to adhere to overly standardized, reductive guidelines [5]. In many cases, clinical trials fail to enroll subjects representative of patients seen in

practice. In fact, the most complex patients—elderly patients with multimorbidity—are typically excluded from RCT participation. Even when older patients are included, trials tend to enroll only the healthiest, generally excluding those with significant numbers of comorbid conditions, functional impairments, and cognitive dysfunction [6]. In cancer therapeutic trials, for example, elderly patients represent only 25 percent of trial participants, despite comprising greater than 60 percent of cancer cases overall [7], and rigorous exclusion criteria permit only the healthiest older patients to enroll.

Can EBM Be Used to Individualize Care for Older Patients?

This process of developing EBM-based guidelines and applying them to clinical care highlights the tension between generating unbiased knowledge based on statistical aggregation and the application of this information to individual patients. RCTs are designed to eliminate the "noise" of population heterogeneity from the measurement of true effects between an intervention and an outcome. Clinical practice, however, must respond to the specificity of the individual patient's condition, and clinical heterogeneity increases with age and multimorbidity.

Many older adults remain robust and clinically similar to their younger counterparts, but others are more vulnerable to stressors and still others are frail, generating a broad spectrum of older patients. Applying data from narrowly defined clinical trials that enroll mostly younger and healthier patients to the entire spectrum of older patients is inappropriate and possibly even harmful in many situations.

What about Quality Measures?

Should EBM-based clinical practice guidelines form the basis for quality and cost containment measures for older adults with complex conditions? Attractive as it might be for standardizing health care, there are pitfalls to this approach. Such a strategy both fails to account for the uniqueness of older individuals and dismisses the role of the physician as a judicious user of evidence. Encouraging defensive and algorithmic medical practice diminishes the "art of care" that has been a cornerstone of medicine. EBM-based practice was never intended to be a static, one-size-fits-all mandate, but a highly dynamic process that includes a trained physician's judgment in applying multiple sources of evidence to complex, emergent systems (i.e., the irreducible totality of the patient). Linking EBM to quality measures and pay-for-performance threatens to co-opt this process, eliminating the step between the measuring of population-based evidence and the weighing of its merit for specific individuals.

Furthermore, there is no definitive evidence that using EBM clinical guidelines as quality measures improves outcomes. Most quality measures focus on structures or processes, rather than directly on outcomes [8]. Outcomes are what really matter, but they can be heavily influenced by factors outside the health care system (e.g., poverty, education level, lifestyle choices). Processes, such as screening for cancer, giving aspirin after a heart attack, or giving adjuvant chemotherapy to stage III colon cancer patients, are easily measured and directly ascertain what is being done to the

patient in the health care setting. Processes are then "linked" to outcomes based on RCTs showing, for example, that aspirin improves mortality when given after an MI [9], or that adjuvant chemotherapy improves survival following resection of a stage III colon cancer [10]. Still, these linkages between process and outcome may not be generalizable, particularly to complex clinical situations. Little data exist showing that process measures definitively improve outcomes in many cases, and they may prove quite harmful for many older patients.

Furthermore, using aggregate evidence may lead to denial of reimbursement for the treatments deemed most useful for a particular patient. The debates that have raged the last few years over the use of mammography (especially in women ages 40-50) and prostate-specific antigen (PSA) screening for men illustrate this point; using these guidelines as a basis for insurance reimbursement, under the rubric of "quality care," often ignores individual risk-benefit ratios.

Conclusions

The substitution of overly simplistic guidelines for nuanced clinical judgment attuned to individual patient complexity can be dangerous. Unless we focus more on clinical judgment, care cannot be said to be "individualized" or "patient-centered." It is often said that we are shifting toward a "patient-centered" model of health care, one that involves patients in medical decision making and treats the "whole patient" [11]. The reality for many physicians can be very different: hemmed in by decreasing payments, increasing demands for reporting and documentation, and growing interference by insurers, doctors can scarcely afford to educate patients during their increasingly abbreviated clinic visits, elicit their opinions and insight, or have conversations about their goals.

The patients who account for the largest share of our health care dollars—older adults with multiple chronic diseases—are also those for whom EBM guidelines least often apply. In fact, applying EBM guidelines to such individuals can be an onerous or impossible challenge [12]. Our older patients with chronic conditions are the most vulnerable in a health care model driven solely by EBM-based quality measures, and perhaps have the most to gain from a truly patient-centered model of health care. Algorithmic guidelines sufficiently flexible to account for the wide heterogeneity of older adults are unlikely to be created. Bluntly applied guidelines may cause as much harm as good. Instead, we need to find better ways to support thoughtful, well-trained clinicians in applying evidence in a shared decision with patients. Reimbursing physicians for spending more time with and thinking harder about these patients, rather than checking off "quality" boxes, is one genuine way to improve quality in health care. This approach is more likely to produce more ethically sound, as well as more individualized, outcomes for older adults.

References

1. Nolte E, McKee CM. Measuring the health of nations: updating an earlier analysis. *Health Aff (Millwood)*. 2008;27(1):58-71.
2. Cohen S, Yu W. Statistical Brief #354: The concentration and persistence in the level of health expenditures over time: estimates for the U.S. population, 2008-2009. Agency for Healthcare Research and Quality. http://meps.ahrq.gov/mepsweb/data_files/publications/st354/stat354.shtml. Accessed December 13, 2012.
3. Guyatt G, Cairns J, Churchill D, et al. Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA*. 1992;268(17):2420-2425.
4. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ*. 1996;312(7023):71-72.
5. Feinstein AR, Horwitz RI. Problems in the "evidence" of "evidence-based medicine". *Am J Med*. 1997;103(6):529-535.
6. Zulman DM, Sussman JB, Chen X, Cigolle CT, Blaum CS, Hayward RA. Examining the evidence: a systematic review of the inclusion and analysis of older adults in randomized controlled trials. *J Gen Intern Med*. 2011;26(7):783-790.
7. Murthy VH, Krumholz HM, Gross CP. Participation in cancer clinical trials: race-, sex-, and age-based disparities. *JAMA*. 2004;291(22):2720-2726.
8. Donabedian A. Evaluating the quality of medical care. *Milbank Mem Fund Q*. 1966;44(3 Suppl):166-206.
9. Antithrombotic Trialists' Collaboration. Collaborative meta-analysis of randomised trials of antiplatelet therapy for prevention of death, myocardial infarction, and stroke in high risk patients. *BMJ*. 2002;324(7329):71-86.
10. Andre T, Boni C, Mounedji-Boudiaf L, et al. Oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment for colon cancer. *N Engl J Med*. 2004;350(23):2343-2351.
11. Agency for Healthcare Research and Quality. Expanding patient-centered care to empower patients and assist providers. *Research in Action*. 2002;(5):1-11. http://www.ahrq.gov/qual/ptcareria.pdf. Accessed December 13, 2012.
12. Boyd CM, Darer J, Boult C, Fried LP, Boult L, Wu AW. Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. *JAMA*. 2005;294:716-724.

Erika Ramsdale, MD, is in the final year of her hematology/oncology fellowship at the University of Chicago. She has also completed fellowships in geriatric medicine and clinical medical ethics and is planning to pursue a career in geriatric oncology. Her research focuses on clinical decision making in older adults with cancer, particularly those with complex comorbidities or frailty.

William Dale, MD, PhD, is the section chief of Geriatrics and Palliative Medicine at the University of Chicago, where he established the Specialized Oncology Care and

Research in the Elderly (SOCARE) clinic. He has published more than 50 articles on medical decision making, geriatric oncology, and quality of life for older adults, particularly older adults with cancer. Dr. Dale is a board-certified geriatrician with a PhD in health policy.

Related in VM

The Limitations of Evidence-Based Medicine—Applying Population-Based Recommendations to Individual Patients, January 2011

A Call to Integrate Ethics and Evidence-Based Medicine, January 2013

The Debate over Prostate Cancer Screening Guidelines, January 2011

N-of-1 Trials: Individualized Medication Effectiveness Tests, January 2013

Virtual Mentor
American Medical Association Journal of Ethics
January 2013, Volume 15, Number 1: 56-64.

POLICY FORUM

Public Deliberation in Decisions about Health Research

Joanna E. Siegel, ScD, Jessica Waddell Heeringa, MPH, and Kristin L. Carman, PhD

The idea of "patient-centeredness" receives much attention in today's health care environment. Increased patient engagement in health care decision making is posited as a promising path toward better quality, efficiency, and health. Importantly, this emphasis has extended beyond the setting of patient care to medical and health services research. Many, if not most, health research agencies and organizations have developed new ways to include patients and other members of the public in developing research agendas, identifying research priorities, and evaluating research proposals, and to support their involvement in all steps of the research process (e.g., the National Cancer Institute Director's Consumer Liaison Group, the Department of Defense Congressionally Directed Medical Research Programs, the Food and Drug Administration's Patient Representative Program, and the Agency for Healthcare Research and Quality Stakeholder Support [1]).

The views, preferences, and values of the public shape many of the most important determinants of health and the effectiveness of health care. Public perceptions determine the impact of programs ranging from obesity prevention to the success of immunization programs and cancer screening and from the use of new technologies to the design of insurance. Recently, interest has increased in the use of structured forums for eliciting public input on specific health care issues, particularly when—as is often the case—the problems faced cannot be resolved by technical information alone and involve values-based or ethical tensions. Public deliberation, a form of public consultation, is often central to these forums.

The literature on public deliberation offers a rich set of ideas and chronicled experiences about its goals, benefits, uses, and expected outcomes. This paper provides a brief overview of public deliberation and describes its emerging role in health and health care research.

Overview of Public Deliberation

Public deliberation is based on the premise that many of the important decisions faced by a society—particularly those that involve competing values and complex trade-offs—are best made by decision makers in partnership with the public [2-6]. In the practice of public deliberation, citizens are brought together to engage in a process of learning about, debating, and discussing an issue. Throughout the process, participants have the opportunity and are encouraged to describe their perspectives and articulate the reasoning behind their views [2, 5, 6, 7-12].

Deliberative methods encompass a range of approaches including citizens' juries or councils, deliberative focus groups, issues forums, deliberative polling, and others. These methods vary considerably in intensity, with sessions lasting anywhere from a few hours to several days; they may convene once or periodically over the course of a term. However, all of these methods contain four core elements of public deliberation. First, a *sponsor convenes* a group of people, either in person or using online technologies that connect people in remote locations [8, 13]. The active interest of the sponsor is essential for framing the questions of interest, motivating participants, and ensuring the effective implementation of the deliberative sessions.

Second, participants are *informed* about the relevant issue(s) through educational materials and/or the use of content experts. Education is critical to the thoughtful discussion that is central to deliberative methods [7, 10, 11]. Third, participants *deliberate* the issues presented. Participants are asked to give reasons for their opinions and preferences with the goal of clarifying underlying values; they are encouraged to listen and respond to the perspectives of others [2, 6, 8-10, 12].

Finally, the content of the deliberation is *reported* to assist a sponsor in understanding public perspectives and incorporating them into decisions. Public deliberation thus assembles a diverse group of people who learn about and debate issues surrounding a social topic and describe their perspectives and reasoning to one another. The result is a record of underlying values and ethics, public reasoning, and options for consideration by decision makers.

Public deliberation is on one end of a continuum of public consultation methods that is defined by the intensity of participants' engagement. Information obtained through public deliberation differs from that collected through public surveys, polls, and other public opinion research methods that obtain "top-of-mind" responses and reactions to public issues [4, 14]. Similarly, public deliberation methods differ from town hall meetings, which focus on informing the public and answering questions but do not require participants to examine and clarify their values.

The fundamental attributes that characterize public deliberation and distinguish it from other methods stem from its roots in theories of deliberative democracy. Democracy is predicated on the idea that an informed public shapes the policies and decisions that affect citizens' lives and well-being. The ideal of public deliberation, in use since ancient Greece, is grounded in philosophies of the social contract and bonds among individuals and institutions that shape political and social life. Such deliberation is a means to bolster democratic life, include underrepresented groups, and promote moral reasoning and mutual understanding.

Amy Gutmann and Dennis Thompson define deliberative democracy as "a form of government in which free and equal citizens (and their representatives) justify decisions in a process in which they give one another reasons that are mutually acceptable and generally accessible, with the aim of reaching conclusions that are binding in the present on all citizens but open to challenge in the future" [5].

Deliberative democracy is distinguished from a "minimalist democracy," which relies solely upon traditional voting and tallying of votes in service of electing a leader. As Robert E. Goodin describes,

> All [that a minimalist democracy] asks of citizens is to cast a ballot from time to time: in most places, if and only if they feel like it… [it] does not ask them to pay attention to public debates on the issues of the day. It does not ask them to get together with others to discuss the issues. It does not ask them to justify their voting decision to anyone else. Still less does it ask people to… persuade others that they should vote the same way [11].

Public deliberation has a number of outcomes, the relative importance of which depend on the goals of the sponsor. These range from changes in the individual participant's knowledge and civic-mindedness to increased public acceptance of policy decisions and improved societal decision making [2, 4, 5, 13, 15-17]. Deliberation can uncover, articulate, and foster shared values as well as diverging perspectives. Further, public deliberation aspires to give voice to underserved groups by providing a shared forum in which majority and minority perspectives are equally encouraged—a goal dependent on the sponsor's commitment and willingness to support and engage participants from diverse backgrounds. This goal ultimately reinforces deliberative outcomes involving both individual participants and social decision-making [4, 5, 13, 18, 19]. Finally, deliberation is an inherently transformative process—designed to alter not only participants who go through the process but also the sponsors and other decision-making entities that seek public input.

Use of Public Deliberation in Health care

The literature describes five general tasks for which deliberation has been used in health care applications. These include (1) developing policy guidance or recommendations, (2) setting priorities, (3) providing guidance on ethical or values-based dilemmas, (4) assessing risks, and (5) determining who should have decision-making authority. Although emphasis varies, these objectives are consistent in U.S. efforts as well as internationally.

In the U.S., recent deliberation projects have included efforts to explore and develop guidance on such topics as childhood obesity, health care reform, health insurance coverage for the uninsured, coverage priorities for Medicare and Medicaid plans, and state-level pandemic planning. For example, in June 2012, potential users of the California Health Benefits Exchange participated in deliberations in seven California locations regarding how to establish fair cost-sharing to address the health needs of a broad population [20]. The primary purpose of the deliberations was to learn how the public prioritizes health needs for affordable coverage in order to inform the design of the health benefits offered through the exchange [20]. Participants prioritized chronic illness and catastrophic losses for higher levels of coverage.

In 2008, Washington, D.C. residents participated in a deliberative exercise, prioritizing socioeconomic and health interventions as part of a hypothetical benefit package of social programs for the purposes of maximizing health. Participants gave priority to health insurance, housing, job training, and dental care [21]. Community members in Michigan were engaged in a deliberative process about pandemic planning to inform public health officials about the courses of action, such as closing schools and workplaces, that would be acceptable responses [22].

In several other countries, deliberative processes have become a more widespread and sometimes institutionalized approach to incorporating public input into health care decisions. In Canada, citizen engagement exercises have taken place in all provinces as well as at a national level and have been used to address a variety of concerns. In New Brunswick, a provincewide initiative engaged the public in establishing priorities for primary care, acute/supportive care, and systemwide concerns [23]. In Ontario, public consultation was used to prioritize services at a community hospital that was facing budget deficits, with citizens acting to identify core services to be maintained at the hospital as well as noncore services (including, for example, a diabetes complication prevention clinic, complex continuing care, and outpatient rehabilitation) that would be closed [24, 25]. In Quebec, a consultation forum met over a 3-year period to provide input about the performance of the health care system and to discuss specific social and ethical dilemmas submitted to it by the Health and Welfare Commissioner [26].

In the United Kingdom, the National Institute for Health and Clinical Excellence (NICE), an independent "arm's length body" funded by the Department of Health, founded a standing 30-member Citizens Council to assist with its work, which includes developing guidance for public health and health care, evaluating new technologies, and establishing quality standards, among other activities [27]. NICE's Citizens Council represents possibly the most formalized version of a deliberative panel in health care, with ongoing governmental support, in contrast to more local and ad hoc efforts. The explicit role of the Citizens Council is issuing social value judgments—judgments that "take account of the ethical principles, preferences, culture and aspirations that should underpin the nature and extent of the care provided by a health service" [14].

The first Citizens Council convened in 2002 to discuss NICE's top priority topic—identifying the factors NICE should consider when making decisions about clinical need of patients with a particular disease or condition. NICE asked the council to think specifically about the most important features of conditions and of patients (apart from their conditions) that should be considered and the weight to give the views of various stakeholders (e.g., patients, health care professionals, family and caregivers, or patient advocates) in determining clinical need. Some of the important features the participants generated included the severity of the pain caused by a condition, whether a condition was potentially fatal or contagious, the availability of alternative treatments, and the patient's age and ability to undergo treatment [28]. Since this initial report, the Citizens Council has deliberated and released reports on

15 topics. Citizens Council reports summarize the social values and principles that are involved in each topic they consider; NICE's advisory committees and guideline development groups are expected to adopt the public's principles in issuing guidance [14].

Deliberation and Health Care Research

Public deliberation in the area of health research is as yet quite limited. Deliberation has been used to (1) address principles for the conduct of research, (2) explore priorities for research, and (3) consider how research evidence should be applied in health decisions.

In addressing research *principles*, deliberative projects have focused on consent and the use of health information. For example, participants in southeastern Michigan deliberated on whether society should allow surrogate consent for research participation for persons with dementia [9]. Participants supported development of a societal policy covering surrogate consent. A citizens' jury in New Zealand deliberated on the use of personal health data for studying drug safety, concluding that identifiable health data could be used for research purposes without consent as long as relevant laws and ethical principles were followed [29]. Similarly, a citizens' panel in Denmark deliberated about using data derived from electronic health records (EHR) in research [30]. The panel recommended that patients be informed of the uses of their data and that a formal policy governing the use of EHR data be developed.

One of the first deliberative projects to address *priorities* for research took place in Bristol, England, where a 20-member citizens' jury identified broad areas important for research and then priority questions in each area [31, 32]. For example, in the area of prevention, high-priority research questions included whether preventive measures provided through local primary care (immunizations, mental health, lifestyle advice) were effective and how to maximize the impact of preventive health strategies for high-risk groups. A citizens' jury in Alberta, Canada identified 13 criteria for setting priorities for health technology assessment [33]. Among the criteria they identified were a technology's potential to benefit many people, to extend length of life while maintaining quality, and to improve quality of life.

Addressing the *application* of health research, deliberative initiatives have engaged the public to elucidate ethical and social values related to the introduction of new technologies into health care practice. A 14-member Citizens' Reference Panel on Health Technologies in Ontario developed a set of social values and ethical principles that should be taken into account in technology assessments of such health technologies as colorectal and breast cancer screening [34]. In the U.S., Gold and colleagues explored the acceptability of cost-effectiveness information as a supplement to evidence on medical effectiveness in determining the priority for Medicare's coverage of treatments [35]. Currently, the U.S. Agency for Healthcare Research and Quality (AHRQ) is sponsoring a large demonstration of public deliberation to obtain public input regarding the appropriate ways to use medical

evidence to guide health care practice. Findings from the demonstration are expected in 2013, and will be used to inform the comparative effectiveness research enterprise in AHRQ's Effective Healthcare (EHC) Program.

Conclusions

The role of health care research in health care decisions is complex. Evidence from medical research is often less than conclusive, and policies based on evidence may imply tradeoffs regarding the distribution of risks and benefits across society, tensions between short-term and long-term outcomes, and a range of effects on quality of life, among other considerations and consequences [15, 36]. Recommendations and guidelines designed to improve health care based on research may conflict with public perceptions of quality care and meet with suspicion [37]. Effective use of health care research depends on understanding and consideration of public values and perceptions, in addition to the development of scientifically valid research results.

Public deliberation aims to facilitate meaningful and inclusive public engagement in policy and social issues. As efforts to include patients in the design and execution of research studies continue to expand, public deliberation offers a means for the general public to become involved with the broader social context that determines the impact of research, from the identification of research priorities to the use of research results to shape health care policy and practice. As a result, deliberation offers a means to enhance the value of research in improving the health of the public.

References

1. Agency for Healthcare Research and Quality Effective Health Care Program. Resources for getting involved and involving others. http://www.effectivehealthcare.ahrq.gov/tools-and-resources/how-to-get-involved-in-the-effective-health-care-program/#getinvolved. Accessed December 13, 2012.
2. Chambers S. Deliberative democratic theory. *Ann Rev Polit Sci*. 2003;6(1):307-326.
3. Daniels N, Sabin J. Limits to healthcare: fair procedures, democratic deliberation, and the legitimacy problem for insurers. *Philos Public Aff*. 1997;26(4):303-350.
4. Fishkin JS. *When the People Speak: Deliberative Democracy and Public Consultation*. Oxford: Oxford University Press; 2009.
5. Gutmann A, Thompson D. *Why Deliberative Democracy?* Princeton, NJ: Princeton University Press; 2004: 7.
6. Young IM. *Inclusion and Democracy*. Oxford: Oxford University Press; 2000.
7. Abelson J, Forest PG, Eyles J, Smith P, Martin E, Gauvin FP. Deliberations about deliberative methods: issues in the design and evaluation of public participation processes. *Soc Sci Med*. 2003;57(2):239-251.

8. Burkhalter S, Gastil J, Kelshaw T. A conceptual definition and theoretical model of public deliberation in small face-to-face groups. *Commun Theory*. 2002;12(4):398-422.
9. De Vries R, Stanczyk A, Wall IF, et al. Assessing the quality of democratic deliberation: A case study of public deliberation on the ethics of surrogate consent for research. *Soc Sci Med*. 2010;70(12):1896-1903.
10. Fishkin J, Farrar C. Deliberative polling: From experiment to community resource. In: Gastil J and Levine P, eds. *The Deliberative Democracy Handbook: Strategies for Effective Civic Engagement in the Twenty-First Century*. San Francisco: Jossey-Bass; 2005:68-79.
11. Goodin RE. *Innovating Democracy: Democratic Theory and Practice after the Deliberative Turn*. Oxford: Oxford University Press; 2008: 1.
12. Gracia D. Ethical case deliberation and decision making. *Med Health Care Philos*. 2003;6(3):227-233.
13. Jacobs LR, Cook FL, delli Carpini MX. *Talking Together: Public Deliberation and Political Participation*. Chicago: University of Chicago Press; 2009.
14. Rawlins MD. Pharmacopolitics and deliberative democracy. *Clin Med*. 2005;5(5):471-475.
15. Davies C, Wetherell M, Barnett E, Seymour-Smith S. *Opening the Box: Evaluating the Citizens Council of NICE*. Milton Keynes: Open University; 2005. http://www.nice.org.uk/media/A0D/35/Final_evaluation_document_-_as_published_(18-3-05).pdf. Accessed December 13, 2012.
16. Arvai JL. Using risk communication to disclose the outcome of a participatory decision-making process: effects on the perceived acceptability of risk-policy decisions. *Risk Anal*. 2003;23(2):281-289.
17. Button M, Ryfe DM. What can we learn from the practice of deliberative democracy? In: Gastil J, Levine P, eds. *The Deliberative Democracy Handbook: Strategies for Effective Civic Engagement in the Twenty-First Century*. San Francisco: Jossey-Bass; 2005:20-34.
18. Smith G, Wales C. Citizens' juries and deliberative democracy. *Political Studies*. 2000;48(1):51-65.
19. Kohn M. Language, power, and persuasion: Toward a critique of deliberative democracy. *Constellations*. 2000;7(3):408-429.
20. Ginsburg M, Glasmire K, Foster T. Sharing in the costs of care: perspectives from potential health plan users of the California Health Benefit Exchange. Center for Healthcare Decisions; 2012. http://chcd.org/docs/hbex_report_6.15.12.pdf. Accessed December 13, 2012.
21. Pesce JE, Kpaduwa CS, Danis M. Deliberation to enhance awareness of and prioritize socioeconomic interventions for health. *Soc Sci Med*. 2011;72(5):789-797.
22. Baum NM, Jacobson PD, Goold SD. Listen to the people: public deliberation about social distancing measures in a pandemic. *Am J Bioeth*. 2009;9(11):4-14.

23. New Brunswick Health Council, Pollack D, Mackinnon MP. Case 7: Our health. our perspectives. our solutions: establishing a common health vision. In: *CIHR's Citizen Engagement in Health Casebook*. Ontario: Canadian Institutes of Health Research; 2012: 45-50. http://www.cihr-irsc.gc.ca/e/45358.html#a9. Accessed December 13, 2012.
24. Biron R, Gillard J. Shared challenge, shared solution: Northumberland Hills Hospital's collaborative budget strategy. In: *CIHR's Citizen Engagement in Health Casebook*. Ontario: Canadian Institutes of Health Research; 2012: 37-43. http://www.cihr-irsc.gc.ca/e/45358.html#a8. Accessed December 13, 2012.
25. Northumberland Hills Hospital. Citizens' Advisory Panel on Health Service Prioritization Final Report. http://www.nhh.ca/SharedChallengeSharedSolution/FinalReport.aspx. Accessed December 13, 2012.
26. Gauvin FP, Martin E, Abelson J. Quebec health and welfare commissioner's consultation forum. *CIHR's Citizen Engagement in Health Casebook*. Ontario: Canadian Institutes of Health Research; 2012: 69-73. http://www.cihr-irsc.gc.ca/e/45358.html#a6. Accessed December 13, 2012.
27. National Institute for Health and Clinical Excellence (NICE). What we do. http://www.nice.org.uk/aboutnice/whatwedo/what_we_do.jsp. Accessed December 13, 2012.
28. NICE Citizens Council. Report of the first meeting of the nice citizens council: determining "clinical need." http://www.nice.org.uk/media/065/B4/FINALNICEFirstMeeting_FINALReport.pdf. Accessed December 13, 2012.
29. Parkin L, Paul C. Public good, personal privacy: a citizens' deliberation about using medical information for pharmacoepidemiological research. *J Epidemiol Commun Health*. 2011;65(2):150-156.
30. Zurita L, Nohr C. Patient demands and the development of EHR systems. *Stud Health Technol Inform*. 2003;95:880-885.
31. Gooberman-Hill R, Horwood J, Calnan M. Citizens' juries in planning research priorities: process, engagement and outcome. *Health Expect.* 2008;11(3):272-281.
32. Bristol Citizens' Jury. *Towards a More Caring City*. Bristol: MRC Health Services Research Collaboration, 2006.
33. Menon D, Stafinski T. Setting priorities for health technology assessment: Public participation using a citizens' jury. *Health Expect*. 2008;11(3):282-293.
34. Abelson J, Wagner F, Levin L, et al. Consulting Ontario citizens to inform the evaluation of health technologies: the citizens' reference panel on health technologies. In: *CIHR's Citizen Engagement in Health Casebook*. Ontario: Canadian Institutes of Health Research; 2012: 69-73. http://www.cihr-irsc.gc.ca/e/45358.html#a16. Accessed December 13, 2012.
35. Gold MR, Franks P, Siegelberg T, Sofaer S. Does providing cost-effectiveness information change coverage priorities for citizens acting as social decision makers? *Health Policy*. 2007;83(1):65-72.

36. Culyer AJ. NICE's use of cost effectiveness as an exemplar of a deliberative process. *Health Econ Policy Law*. 2006;1(Pt 3):299-318.
37. Carman KL, Maurer M, Yegian JM, et al. Evidence that consumers are skeptical about evidence-based healthcare. *Health Aff (Millwood)*. 2010;29(7):1400-1406.

Joanna E. Siegel, ScD, is a senior scientist in the Agency for Healthcare Research and Quality's Center for Outcomes and Evidence. She coordinates the Community Forum Project, which includes a demonstration that is evaluating use of deliberative methods for providing public input on the use of medical evidence in health care decision making, and projects to enhance patient and other stakeholder involvement in AHRQ's Effective Health Care program.

Jessica Waddell Heeringa, MPH, is a research analyst at Mathematica Policy Research. She co-wrote this manuscript while working at the American Institutes for Research on the Agency for Healthcare Research and Quality's Community Forum. Her research interests include health reform implementation, mental health care, and health disparities.

Kristin L. Carman, PhD, co-directs the Health Policy and Research group at the American Institutes for Research in Washington, DC. She leads the Community Forum, an initiative of the Agency for Healthcare Research and Quality's Effective Health Care program that will evaluate and develop approaches to expand the participation of the public and various stakeholder groups in improving the effectiveness of health care.

Acknowledgment
The authors would like to acknowledge the contributions of Susan K.R. Heil, PhD, to the literature review that informed this piece.

Related in VM
Power, Politics, and Health Spending Priorities, November 2012

Virtual Mentor
American Medical Association Journal of Ethics
January 2013, Volume 15, Number 1: 65-70.

MEDICINE AND SOCIETY

Paradigms, Coherence, and the Fog of Evidence

Dien Ho, PhD

It is wrong always, everywhere, and for anyone, to believe anything upon insufficient evidence.
William Clifford, "The Ethics of Belief"

In the great boarding-house of nature, the cakes and the butter and the syrup seldom come out so even and leave the plates so clean. Indeed, we should view them with scientific suspicion if they did.
William James, "The Will to Believe"

The idea that the practice of medicine should be based on evidence strikes most laypersons as trivially true. An expert's claims differ from the opinions of a nonexpert precisely because the former enjoy a certain evidential support that the latter lack. In this respect, the emergence of evidence-based medicine (EBM) circa 1992 as a "new paradigm" implies a worrisome question to most people: If medical professionals are now encouraged to make claims (diagnoses, prognoses, therapeutic recommendations, etc.) on the basis of evidence, what were they doing before 1992 [1]? Of course, upon a closer examination, one learns that EBM is less of a revolution than an urging that medical practitioners guide their clinical judgments on the basis of the best available clinical research—most often, the results of randomized clinical trials (RCTs) representing the "gold standard" [2].

Throughout the discussion on the nature and merit of EBM, the questions of what constitutes evidence and how it relates to how we should act and think remain largely unexplored. Indeed, both advocates and opponents of EBM have assumed that we have a relatively unproblematic understanding of evidence. Critics of EBM have argued that EBM places too great an emphasis on data derived from RCTs while ignoring non-RCT evidence that can be clinically useful [3]. But the disagreement here concerns the scope of clinically relevant evidence and not the nature of evidence *per se*. From a philosophical point of view, however, evidence and its logical relationship to theories and judgments represent one of the most perplexing puzzles in the philosophy of science.

An Abbreviated History of Evidential Reasoning

David Hume's argument against the rationality of inductive reasoning arguably marks the start of a critical examination of evidential reasoning—that is, forming beliefs on the basis of evidence—in the modern era [4]. Hume argued that empirical generalizations on the basis of past observed evidence must rely on the principle that

nature behaves in a uniform fashion; that is, all else being equal, future events will resemble past instances of the similar sort. The problem, however, is that if we attempt to justify this principle of uniformity of nature by appealing to evidence of its past success, we risk justifying induction inductively. To put it another way, if one questions the rationality of inductive reasoning, appealing to inductive reasoning to answer that worry can hardly bring any reassurance. It is analogous to allaying one's concern of someone's trustworthiness by asking the person if he can be trusted.

Hume left empiricists and admirers of science with a challenge: Make sense of the apparent success and superiority of evidential reasoning in the face of the argument that it is not rationally justifiable.

Impressed by advances made in mathematics, logic, and physics in the early twentieth century, logical positivists such as Rudolf Carnap and Carl Hempel attempted to formalize the logic of confirmation and the logic of science in general. The implicit assumption behind the logical positivists' project was that observations supply theory-neutral evidence that can adjudicate scientific disagreements. The burst of scientific progress in the early twentieth century, the positivists believed, was the product of rigorous adherence to the scientific method. Rather than addressing Hume's fundamental challenge to the rationality of induction, the positivists focused on explicating the logic of induction. They argued that the success of modern science should provide *prima facie* justification for evidential reasoning.

The logical positivists were followed by theorists like Karl Popper, who agreed that there is a logic to the scientific method but argued that this logic was not about confirming theories by collecting confirming evidence—rather it was about attempting to falsify theories. In Popper's view, what distinguished Einstein's general relativity from Marx's theory of history was not that the former enjoyed evidential support while the latter did not. Popper argued that proponents of Marx's theory of history could find evidence wherever they looked; supporting evidence, it turns out, can be acquired too easily. The real separation between these two theories was that Marx's theory of history was not *falsifiable;* there was no possible evidence one could find to refute the theory. Therefore, it was unscientific. Popper's view of what constituted scientific evidence depended on a clear concept of falsification, on the ability to logically refute a hypothesis on the basis of a contrary piece of evidence.

Most philosophers today believe that evidence provides support for theories. One might define confirming evidence as follows:

$$P(H/e+) > P(H)$$

That is, the probability of a hypothesis H being true given evidence $e+$ is greater than the probability of the hypothesis alone being true.

Moreover, a piece of evidence can also be used to disconfirm a hypothesis. Contrary evidence in this case would lower the probability that the hypothesis is true. More precisely, if *e*- is a piece of disconfirming evidence, the relationship can be represented as:

$$P(H/e\text{-}) < P(H)$$

These two conditions appear to be necessary for any full analysis of the concept of evidence. However, the crucial question for those who are interested in the logic of confirmation (and disconfirmation) regard the conditions under which these definitions are true.

The optimistic belief that philosophers would eventually uncover the logic of evidential reasoning began to wane in the second half of the twentieth century. As reasons for this change, a number of philosophers have offered a variety of arguments challenging the possibility of an objective logical relationship between evidence and theories. Below are two well-known arguments.

Thomas Kuhn's Paradigms

In *The Structure of Scientific Revolutions*, Kuhn combined a rigorous examination of the history of scientific revolutions with a critical analysis of the logic of science. According to Kuhn, normal science consists of attempts to solve specific problems in accordance with the accepted paradigm of the time. A paradigm specifies among other things what constitutes well-formulated puzzles, acceptable solutions to those puzzles, and properly ignorable abnormalities. Puzzle-solving exemplars identified by the paradigm provide the models that practitioners of science should emulate. Moreover, as scientific revolutions replace one paradigm with another, the criteria that practitioners use to evaluate the support an observation offers for a theory also change.

Take, for example, Newton's concept of gravitational attraction, which entails action at a distance. Lack of an explanation that involved direct physical contact would have rendered it a nonstarter in a pre-Newtonian paradigm. But as Newtonian mechanics became the dominant paradigm, explaining action at a distance ceased to be a puzzle that demanded attention. The failure to explain away action at a distance no longer represented a defect in a proposed solution to a research problem or counted as a piece of evidence against the plausibility of the solution. What constituted acceptable evidence changed based on the new paradigm. Other examples of such "paradigm shifts" include Copernicus' postulation that the sun, rather than the Earth, is at the center of our solar system and Einstein's theory of relativity.

Most Western medical professionals subscribe to roughly the same paradigm: Diseases and disabilities stem from morphological, chemical, or genetic causes. Furthermore, there is a general agreement in terms of what qualifies as evidence (e.g., RCTs, cell biology, organic chemistry, and so on). The supportive strength of a

given piece of evidence is also largely uncontroversial. RCTs, for instance, are generally thought to provide more support than other types of studies designed to answer the same question.

Nevertheless, across different paradigms, the evidential support of a particular observation can vary significantly. A Chinese physician who explains diseases and disabilities in terms of improper flow of *qi* would not look at a tumor mass as evidence of cancer. Indeed, the very description of the observation would be different. There is no way of defining *qi* without a fairly robust acceptance of an entirely different medical approach. In other words, there is no way to incorporate *qi* into Western medical paradigms without a fundamental change in the accepted paradigm. I am not endorsing here the soundness of Chinese medicine, to be sure. However, this example demonstrates that the idea of evidence serving as neutral arbiter of choosing theories is simply incompatible with Kuhn's view that evidence cannot be evaluated in a paradigm-independent manner. The appearance of irrationality or quackery can only be measured from the point of view of one's accepted paradigm. There is, in other words, no appeal to evidence that does not rely on an accepted paradigm.

Willard Van Orman Quine and the Role of Psychology

No other philosopher played as important of a role in the Anglo-American analytic philosophy tradition in the latter half of the twentieth century as did Quine. In "Two Dogmas of Empiricism," Quine argued forcefully that logical truths and empirical truths differ only in degree [5]. The picture that Quine paints is essentially this: Our understanding of the world is based on an interconnected web of beliefs. Sitting in the core of this web are firmly held beliefs such as the law of noncontradiction (i.e., contradictory claims are never true), some fundamental claims of physics (e.g., gravity), and so on. As one extends to the periphery of the web, beliefs become less significant in the sense that a rejection of one of these beliefs requires only a minor revision in the web to restore logical consistency.

Suppose one observes that a patient receiving benzodiazepine to treat anxiety has instead become *more* anxious. How would one reconcile that against one's expectation? According to Quine, when there is a disturbance to our web of beliefs, we revise our web in the least cognitively taxing manner. In other words, we revise our peripheral beliefs before we revise our core beliefs.

In the case of the unexpected effect of benzodiazepine, one might choose to abandon one's belief that the pharmacokinetics of the drug is fully known. Alternatively, one *could* revise a core belief to accommodate the observation. For example, one might conclude that anxiety (and perhaps other psychobehavioral disease) is not caused by biochemical processes affected by benzodiazepine—but this latter revision strikes us as implausible. The cognitive price it would require, its effect on so many other beliefs, would be much higher than merely believing that the drug's function is not fully known. Nevertheless, Quine insists that such a revision would not be impossible or necessarily incorrect.

In the history of science, there have been numerous occasions when core beliefs have been revised (e.g., the Copernican revolution). Why and when individuals and the community *decide* to make the deep revision is a matter of sociology and psychology—is an individual or a society ready to make the cognitive leap it takes to believe something new? It is perhaps here that Kuhn's insight into scientific revolutions becomes relevant. The important point is that the relationship between evidence and theory may hinge more on psychology than on logic or the pursuit of objective truth.

Why Defining Evidence Matters for Clinicians

I have selected these arguments from the tradition of analytic philosophy to raise some doubts that anyone can rely on an unproblematic concept of evidential support. Medical professionals ought to appreciate the complexity of the concept of evidence as outlined by philosophers in the past 50 years.

It is important to remember these arguments when acting on various types of evidence in clinical settings. Insisting like William Clifford, who is quoted at the beginning of this article, that no judgment be made without sufficient evidence constitutes the adoption of a principle that relies on deeply problematic concepts [6]. Rather, clinicians would be wise to remember that evidence rarely comes out so even or clean.

References

1. For a critical examination of EBM's use of "paradigm shift" see Couto J. Evidence-based medicine: a Kuhnian perspective of a transvestite non-theory. *J Eval Clin Pract*. 1998;4(4):267-275.
2. One of the earliest articulations of EBM can be found in Evidence-Based Medicine Working Group. Evidence-based medicine: a new approach to teaching the practice of medicine. *JAMA*. 1992;268(17):2420-2425.
3. Feinstein AR, Horwitz RI. Problems in the "evidence" of "evidence-based medicine." *Am J Med*. 1997;103(6):529-535.
4. Hume D. *An Enquiry Concerning Human Understanding*. Beauchamp T, ed. Oxford: Oxford University Press; 2006: sec IV.
5. Quine W. *From a Logical Point of View: Nine Logico-Philosophical Essays.* Boston: Harvard University Press; 1980.
6. Clifford W. The ethics of belief. In: Madigan T, ed. *The Ethics of Beliefs and Other Essays*. Amherst, NY: Prometheus; 1999.

Dien Ho, PhD, is an associate professor of philosophy and health care ethics at Massachusetts College of Pharmacy and Health Sciences. His research focuses primarily on the ethics of organ transplantation, reproductive autonomy, pharmacist ethics, and theoretical reasoning.

Acknowledgement

I would like to thank Susan Gorman, Ken Richman, and the editors of *Virtual Mentor* for their helpful comments.

Related in VM
Antidepressants and the FDA's Black-Box Warning: Determining a Rational Public Policy in the Absence of Sufficient Evidence, June 2012

Virtual Mentor
American Medical Association Journal of Ethics
January 2013, Volume 15, Number 1: 71-76.

HISTORY OF MEDICINE

Evidence-Based Medicine: A Short History of a Modern Medical Movement

Ariel L. Zimerman, MD, PhD

On a cold morning in October 1993, Gordon Guyatt, a young faculty member at McMaster Medical School in Hamilton, Ontario, found a brochure published by the American College of Physicians (ACP) in his mailbox bearing this title: *In This Era of Evidence-Based Medicine!* (personal communication). For Guyatt, who had coined the term nearly 3 years earlier in a short editorial for the *ACP Journal Club*, the copywriter's blunt assertion proved not to be an exaggeration [1]. Over a short period, evidence-based medicine, defined as "the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients" [2], had evolved into an emblem for an entire generation, becoming synonymous with the practices of quantification and statistics that pervaded the medical milieu at the end of the twentieth century. Indeed, an era of evidence-based medicine (EBM) had been ushered in.

It is difficult to exaggerate the impact of EBM on the medical world. A PubMed bibliometric search for the string "evidence-based medicine" reflects the term's meteoric rise in popularity (see figure 1). In 1992, only two article titles included the phrase. These were followed by a virtual avalanche of publications; in just 5 years, by 1997, more than 1,000 articles had used the new phrase. A similar picture emerges when considering medical textbooks, dedicated journals, and web sites. A survey in 2004 found 24 dedicated textbooks, nine academic journals, four computer programs, and 62 Internet portals all dedicated to the teaching and development of EBM [3].

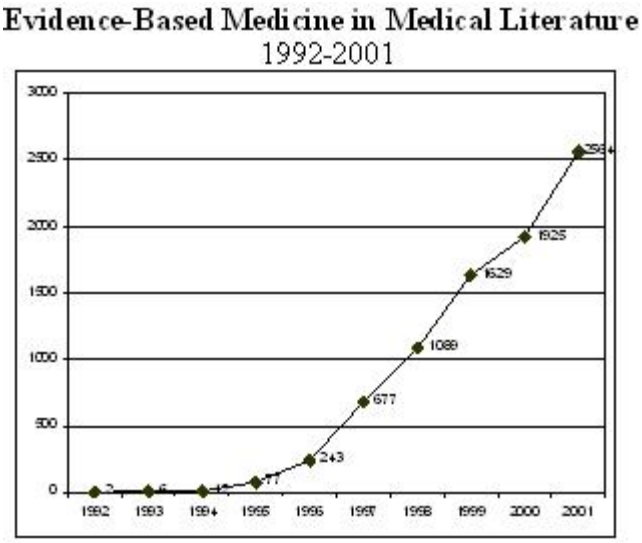

Figure 1. The first decade of "evidence-based medicine" in publication. The graph displays the number of publications using the term per year retrieved by PubMed, beginning with its first major publication in *JAMA* in November 1992.

Despite the wide use of evidence-based methods in medicine and the interest it generated in the fields of clinical methodology, medical sociology, bioethics, and public health, an account of the history of the creation of the methods of EBM is lacking. Here, the development of EBM until its appearance in *JAMA* in 1992 is explored.

North American Clinical Epidemiology

Evidence-based medicine and modern epidemiology share common roots. The history of modern epidemiology and its methods of quantification, surveillance, and control have been traced back to social processes in eighteenth- and nineteenth-century Europe and the introduction of statistics and probability methods. Toward the middle of the twentieth century doctors began to apply these tools to the evaluation of clinical treatment of individual patients mainly in North America and England.

In North America, this took place within the new field of clinical epidemiology, the name of which was coined by John R. Paul in 1938 while working at the Yale School of Medicine. For Paul, clinical epidemiology comprised a multifaceted observation of disease including its social and environmental factors. The primary focus was not entire populations but rather the study of individual patients and their close entourage ("medical ecology") [4]. His concepts were further developed by Alvan R. Feinstein, one of his followers at Yale, who later became the central figure of North American clinical epidemiology. Having studied mathematics prior to studying medicine, Feinstein introduced the use of statistical research methods and Boolean logic into the quantification of clinical practices and study of the medical decision-making process.

Two developments at the end of the 1960s were instrumental for creating the crucible in which evidence-based medicine took shape. The first was the development of new methods in medical education in North America, and the second, the move toward universal coverage of medical costs in the Canadian health system. The reorganization of the Canadian Health System in 1964 by the newly elected liberal party of L. B. Pearson led to the establishment of four new medical schools committed to new ways of integrating public health into medicine [5]. One of these, McMaster University, was founded in 1968. The new school introduced an integrative curriculum it called "problem-based learning" that combined the study of basic sciences and clinical medicine using clinical problems in a tutorship system.

David Sackett and the Department of Clinical Epidemiology and Biostatistics at McMaster

The clinical epidemiology methods developed in North America became a central element in the McMaster curriculum. The new medical school established the world's first department of clinical epidemiology and biostatistics, which was directed by David Sackett. Alvan Feinstein, who proposed the creation of the department, was invited as a visiting professor for the first 2 years of the program. The combination of Sackett's dominant role in preparing the curriculum and the new

department and Feinstein's presence ensured the inclusion of the most novel ideas of clinical epidemiology into McMaster's problem-based learning method. The influence worked in both directions: the McMaster version of clinical epidemiology began to reflect the methods of its problem-based learning curriculum, including an interest in practical clinical problem solving and the analysis of medical decision making.

In 1978, Sackett presented a compilation of the different strategies developed at McMaster in a short course entitled "Critical Appraisal" and later "Critical Appraisal of the Literature." In accordance with the problem-based learning methodology, the courses employed specific clinical problems as a platform for inquiry and discussion in small group tutorial sessions. Sackett's short course was described in a way that easily fits today's EBM:

> using problem solving and small group format, these courses consider the critical assessment of clinical information pertaining to the selection and interpretation of diagnostic tests, the study of etiology and causation, the interpretation of investigation of the clinical course and natural history of human disease, the assessment of therapeutic claims and the interpretation of studies of the quality of clinical care [6].

In 1981, the critical appraisal courses were followed by the first publication of the methods behind them in a seminal nine-article series in the *Canadian Medical Association Journal* (*CMAJ*), later known as the "Readers' Guides" and the opening of McMaster's summer workshop, which would be instrumental for future dissemination of the methods [7]. Over the next 15 years, Sackett and the department undertook five publications of the critical appraisal method in journal series and books under different approaches and names until its final publication under the name "evidence-based medicine" by Gordon Guyatt in 1992 [8].

Table 1. Major publications of the McMaster methods by David Sackett and colleagues by year.

Year	Content	Format	Location
1978	Critical Appraisal Courses	10-week course	McMaster Medical School
1981	"Clinical Epidemiology Rounds" series	article series	*Canadian Medical Association Journal* (*CMAJ*)
1985	*Clinical Epidemiology*	book	
1986	"How to Keep Up with the Medical Literature" series	article series	*Annals of Internal Medicine*
1986, 1990	n-of-1 trials	articles	*New England Journal of Medicine* (*NEJM*) and *Annals of Internal Medicine*
1992	"The User's Guides" series	article series	*JAMA*

The Publication of *Evidence-Based Medicine*: The *JAMA* Users' Guides

The appropriate forum for republication of the McMaster methods appeared in 1990, when Drummond Rennie, a *JAMA* deputy editor, approached David Sackett and the McMaster clinical epidemiology and biostatistics department. Sackett's and Rennie's plan was to publish an updated version of the 1980 *CMAJ* "Readers' Guides" divided into two series. The first, entitled "The Rational Clinical Examination," would be edited by Sackett and would deal with the evaluation of clinical measurements. The second series, entitled "The Users' Guides," would be directed by Gordon Guyatt, a young faculty member from Sackett's department, and would update the critical appraisal methods with greater emphasis on applicability.

Around the same time that Rennie approached the McMaster Group, Gordon Guyatt organized a new medical residency program at McMaster based on the extensive application of the critical appraisal methods, christening it "evidence-based medicine." Guyatt, who in his residency years had been co-founder of the Medical Reform Group, a Canadian medical activist group composed of young doctors and nurses based in Toronto, brought a radical approach to the McMaster critical appraisal methods. The first article of the *JAMA* series appeared on November 4, 1992, using the new name "evidence-based medicine" and language closer to a political manifesto, called for a far-reaching change in the practice of medicine—a "*paradigm shift*"—in order to turn it into an objective and scientific enterprise. It opens accordingly: "A new paradigm for medical practice is emerging. Evidence-based medicine de-emphasizes intuition, unsystematic clinical experience, and pathophysiologic rationale as sufficient grounds for clinical decision making and stresses the examination of evidence from clinical research" [9].

Working with *JAMA* editor Drummond Rennie, Guyatt devised a publication strategy to ensure the approach would have the greatest possible impact on the literature [10, 11]. The first article was written by a new anonymous Evidence-Based Medicine Working Group, giving it the authority of a consensus paper. The working group, Drummond Rennie, and *JAMA* remained the main advocates of the EBM for the first critical years: out of 22 articles on EBM published in the first 3 years, 12 were published by *JAMA*, reflecting Rennie's and *JAMA*'s remarkable commitment to the new methods. The series was initially planned to comprise 10 articles over the span of 3 years; in the end, they were continued for 8 more years and 32 articles. The result was a remarkably complete presentation of the method in a top-tier medical journal, comprising the latest developments in clinical epidemiology encapsulated under the new term "evidence-based medicine."

The new series also become an enterprise of the McMaster Department of Clinical Epidemiology. The influential summer workshop had been renamed using the new term, and McMaster faculty members began publishing articles using that term in other journals.

In 1995, the number of articles with the words "evidence-based medicine" in the title authored by researchers outside the McMaster circle surpassed those by the

McMaster faculty; out of 77 articles published that year, only 23 were related to McMaster members. The rise in articles authored by researchers unconnected to McMaster signaled that the term and the approach had been integrated into the medical discourse.

Conclusion

When the first article on evidence-based medicine was published in November 1992, the methods were not new; they were nearly a quarter-century old. Like its earlier iteration in 1978, the 1992 version of evidence-based medicine was developed and presented in the immediate context of medical education at McMaster. This intimate relation between medical education and medical methodological reforms should be no surprise; as exemplified by Abraham Flexner's report and the changes it conveyed to the North American medical world, reforms in medical education and medical practice in North America have been closely related.

But why did these methods become so widely accepted in the 1990s? There appear to be several reasons to explain the meteoric rise of evidence-based medicine. First, the name itself was indeed a good choice; it was catchy, and it conveyed an intuitive message about the nature of the method. Most physicians did not need to read an entire article series to understand more or less what the name denoted. Second, the support of the McMaster department and *JAMA*'s Drummond Rennie were critical for advancing the methods and their introduction to medical discourse. Finally, the social and cultural milieu of North American medicine in the early 1990s, into which EBM made its debut, was on the whole ripe for the new methods. Quantification practices, the use of statistics and epidemiology, the introduction of computers and online digital databases, and new clinical research methodology saturated the medical environment of the time.

Thus, the universal scope that EBM preached, its encyclopedic inclusion of quantification techniques, and its cadre of unconditional supporters served as a crystallization point and name for the practices of statistics, epidemiology, bioinformatics and clinical research that had already saturated the medical milieu at the turn of the century. Moreover, EBM provided a rejoinder not only to questions regarding the nature of medical knowledge, the evaluation of medical literature, and the use of the new information technologies in the medical field—but also to broader questions of medical authority, relations inside the medical profession and relations between the medical profession and society. As these questions and their answers continue to be examined, it is insightful to understand how EBM, as both a method and a medical movement, entered this discourse.

References

1. Guyatt GH. Evidence-based medicine. *ACP J Club*. 1991;114(suppl 2):A-16.
2. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ*. 1996;312(7023):71.
3. Haynes B. Advances in evidence-based information resources for clinical practice. *ACP J Club*. 2000;132(1):A11-A14.
4. Paul JR. President's address clinical epidemiology. *J Clin Invest*. 1938;17(5):539-541.
5. Spaulding WB, Cochran J. *Revitalizing Medical Education: McMaster Medical School's Early Years*. Hamilton, ON: BC Decker; 1991.
6. Department of Clinical Epidemiology & Biostatistics Annual Report, 1979. In the archives of the Hamilton Health Sciences Corporation and the Faculty of Health Sciences of McMaster University.
7. How to read clinical journals: I. why to read them and start to read critically. *Can Med Assoc J*. 1981;124(5):555-558.
8. Evidence-Based Medicine Working Group. Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA*. 1992;268(17):2420-2425.
9. Evidence-Based Medicine Working Group, 2420.
10. Personal interview with Drummond Rennie, September 15, 2005.
11. Personal interview with Gordon Guyatt, July 30, 2004.

Ariel L. Zimerman, MD, PhD, recently finished his doctorate on the history of evidence-based medicine in the Graduate Program in Science, Technology and Society at Bar-Ilan University in Ramat-Gan, Israel. His research interests are the history of medical epistemology, the introduction of quantification practices to clinical medicine, and the history of medicine in the late twentieth century. He is also a practicing physician and board certified in obstetriccs and gynecology.

Related in VM

The Origins of Evidence-Based Medicine—A Personal Perspective, January 2011

Virtual Mentor
American Medical Association Journal of Ethics
January 2013, Volume 15, Number 1: 77-81.

MEDICAL NARRATIVE
The Median Isn't the Message
Stephen Jay Gould, PhD

Editors' note. This personal essay is reprinted from Bully for Brontosaurus: Reflections in Natural History *by Stephen Jay Gould (c) 1991 by Stephen Jay Gould. Used with permission of the publisher, W. W. Norton & Company, Inc.*

My life has recently intersected, in a most personal way, two of Mark Twain's famous quips. One I shall defer to the end of this essay. The other (sometimes attributed to Disraeli) identifies three species of mendacity, each worse than the one before—lies, damned lies, and statistics.

Consider the standard example of stretching truth with numbers—a case quite relevant to my story. Statistics recognizes different measures of an "average," or central tendency. The mean represents our usual concept of an overall average—add up the items and divide them by the number of sharers (100 candy bars collected for five kids next Halloween will yield twenty for each in a fair world). The median, a different measure of central tendency, is the halfway point. If I line up five kids by height, the median child is shorter than two and taller than the other two (who might have trouble getting their mean share of the candy). A politician in power might say with pride, "The mean income of our citizens is $15,000 per year." The leader of the opposition might retort, "But half our citizens make less than $10,000 per year." Both are right, but neither cites a statistic with impassive objectivity. The first invokes a mean, the second a median. (Means are higher than medians in such cases because one millionaire may outweigh hundreds of poor people in setting a mean, but can balance only one mendicant in calculating a median.)

The larger issue that creates a common distrust or contempt for statistics is more troubling. Many people make an unfortunate and invalid separation between heart and mind, or feeling and intellect. In some contemporary traditions, abetted by attitudes stereotypically centered upon Southern California, feelings are exalted as more "real" and the only proper basis for action, while intellect gets short shrift as a hang-up of outmoded elitism. Statistics, in this absurd dichotomy, often becomes the symbol of the enemy. As Hilaire Belloc wrote, "Statistics are the triumph of the quantitative method, and the quantitative method is the victory of sterility and death."

This is a personal story of statistics, properly interpreted, as profoundly nurturant and life-giving. It declares holy war on the downgrading of intellect by telling a small

story to illustrate the utility of dry, academic knowledge about science. Heart and head are focal points of one body, one personality.

In July 1982, I learned that I was suffering from abdominal mesothelioma, a rare and serious cancer usually associated with exposure to asbestos. When I revived after surgery, I asked my first question of my doctor and chemotherapist: "What is the best technical literature about mesothelioma?" She replied, with a touch of diplomacy (the only departure she has ever made from direct frankness), that the medical literature contained nothing really worth reading.

Of course, trying to keep an intellectual away from literature works about as well as recommending chastity to Homo sapiens, the sexiest primate of all. As soon as I could walk, I made a beeline for Harvard's Countway medical library and punched mesothelioma into the computer's bibliographic search program. An hour later, surrounded by the latest literature on abdominal mesothelioma, I realized with a gulp why my doctor had offered that humane advice. The literature couldn't have been more brutally clear: Mesothelioma is incurable, with a median mortality of only eight months after discovery. I sat stunned for about fifteen minutes, then smiled and said to myself: So that's why they didn't give me anything to read. Then my mind started to work again, thank goodness.

If a little learning could ever be a dangerous thing, I had encountered a classic example. Attitude clearly matters in fighting cancer. We don't know why (from my old-style materialistic perspective, I suspect that mental states feed back upon the immune system). But match people with the same cancer for age, class, health, and socio-economic status, and, in general, those with positive attitudes, with a strong will and purpose for living, with commitment to struggle, and with an active response to aiding their own treatment and not just a passive acceptance of anything doctors say tend to live longer. A few months later I asked Sir Peter Medawar, my personal scientific guru and a Nobelist in immunology, what the best prescription for success against cancer might be. "A sanguine personality," he replied. Fortunately (since one can't reconstruct oneself at short notice and for a definite purpose), I am, if anything, even-tempered and confident in just this manner.

Hence the dilemma for humane doctors: Since attitude matters so critically, should such a somber conclusion be advertised, especially since few people have sufficient understanding of statistics to evaluate what the statements really mean? From years of experience with the small-scale evolution of Bahamian land snails treated quantitatively, I have developed this technical knowledge—and I am convinced that it played a major role in saving my life. Knowledge is indeed power, as Francis Bacon proclaimed.

The problem may be briefly stated: What does "median mortality of eight months" signify in our vernacular? I suspect that most people, without training in statistics, would read such a statement as "I will probably be dead in eight months"—the very

conclusion that must be avoided, both because this formulation is false, and because attitude matters so much.

I was not, of course, overjoyed, but I didn't read the statement in this vernacular way either. My technical training enjoined a different perspective on "eight months median mortality." The point may seem subtle, but the consequences can be profound. Moreover, this perspective embodies the distinctive way of thinking in my own field of evolutionary biology and natural history.

We still carry the historical baggage of a Platonic heritage that seeks sharp essences and definite boundaries. (Thus we hope to find an un-ambiguous "beginning of life" or "definition of death," although nature often comes to us as irreducible continua.) This Platonic heritage, with its emphasis on clear distinctions and separated immutable entities, leads us to view statistical measures of central tendency wrongly, indeed opposite to the appropriate interpretation in our actual world of variation, shadings, and continua. In short, we view means and medians as hard "realities," and the variation that permits their calculation as a set of transient and imperfect measurements of this hidden essence. If the median is the reality and variation around the median just a device for calculation, then "I will probably be dead in eight months" may pass as a reasonable interpretation.

But all evolutionary biologists know that variation itself is nature's only irreducible essence. Variation is the hard reality, not a set of imperfect measures for a central tendency. Means and medians are the abstractions. Therefore, I looked at the mesothelioma statistics quite differently —and not only because I am an optimist who tends to see the doughnut instead of the hole, but primarily because I know that variation itself is the reality. I had to place myself amidst the variation.

When I learned about the eight-month median, my first intellectual reaction was: Fine, half the people will live longer; now what are my chances of being in that half? I read for a furious and nervous hour and concluded, with relief: damned good. I possessed every one of the characteristics conferring a probability of longer life: I was young; my disease had been recognized in a relatively early stage; I would receive the nation's best medical treatment; I had the world to live for; I knew how to read the data properly and not despair.

Another technical point then added even more solace. I immediately recognized that the distribution of variation about the eight-month median would almost surely be what statisticians call "right skewed." (In a symmetrical distribution, the profile of variation to the left of the central tendency is a mirror image of variation to the right. Skewed distributions are asymmetrical, with variation stretching out more in one direction than the other—left skewed if extended to the left, right skewed if stretched out to the right.) The distribution of variation had to be right skewed, I reasoned. After all, the left of the distribution contains an irrevocable lower boundary of zero (since mesothelioma can only be identified at death or before). Thus, little space exists for the distribution's lower (or left) half—it must be scrunched up between

zero and eight months. But the upper (or right) half can extend out for years and years, even if nobody ultimately survives. The distribution must be right skewed, and I needed to know how long the extended tail ran—for I had already concluded that my favorable profile made me a good candidate for the right half of the curve.

The distribution was, indeed, strongly right skewed, with a long tail (however small) that extended for several years above the eight-month median. I saw no reason why I shouldn't be in that small tail, and I breathed a very long sigh of relief. My technical knowledge had helped. I had read the graph correctly. I had asked the right question and found the answers. I had obtained, in all probability, that most precious of all possible gifts in the circumstances—substantial time. I didn't have to stop and immediately follow Isaiah's injunction to Hezekiah—set thine house in order: for thou shalt die, and not live. I would have time to think, to plan, and to fight.

One final point about statistical distributions. They apply only to a prescribed set of circumstances—in this case to survival with mesothelioma under conventional modes of treatment. If circumstances change, the distribution may alter. I was placed on an experimental protocol of treatment and, if fortune holds, will be in the first cohort of a new distribution with high median and a right tail extending to death by natural causes at advanced old age.*

It has become, in my view, a bit too trendy to regard the acceptance of death as something tantamount to intrinsic dignity. Of course I agree with the preacher of Ecclesiastes that there is a time to love and a time to die—and when my skein runs out I hope to face the end calmly and in my own way. For most situations, however, I prefer the more martial view that death is the ultimate enemy—and I find nothing reproachable in those who rage mightily against the dying of the light.

The swords of battle are numerous, and none more effective than humor. My death was announced at a meeting of my colleagues in Scotland, and I almost experienced the delicious pleasure of reading my obituary penned by one of my best friends (the so-and-so got suspicious and checked; he too is a statistician, and didn't expect to find me so far out on the left tail). Still, the incident provided my first good laugh after the diagnosis. Just think, I almost got to repeat Mark Twain's most famous line of all: The reports of my death are greatly exaggerated.**

*So far so good.

**Since writing this, my death has actually been reported in two European magazines, five years apart. Fama volat (and lasts a long time). I squawked very loudly both times and demanded a retraction; guess I just don't have Mr. Clemens's savoir faire.

Stephen Jay Gould, PhD, (1941-2002) was a evolutionary biologist and natural historian, professor at Harvard University and New York University, and acclaimed author of essays and popular science books.

Related in VM
The Message Isn't as Mean as We May Think, January 2013

Virtual Mentor
American Medical Association Journal of Ethics
January 2013, Volume 15, Number 1: 82-85.

MEDICAL NARRATIVE

The Message Isn't as Mean as We May Think

Thomas W. LeBlanc, MD, MA

As medical knowledge grows in scope and complexity, so does the challenge of applying it to individual patients. This is perhaps the single most difficult task for twenty-first-century clinicians, and yet it remains among the most important to patient outcomes.

"How much time do I have, doc?" This is a question I hear almost daily as an oncologist. Although I know what the literature says about the 5-year survival for a particular disease, I struggle with how to discuss statistical predictors with my patients; it is often unclear how statistical data applies to a particular person and situation, or how a patient might use and interpret it. Along with clinical experience, statistical information is a key type of "evidence" that facilitates clinical decision making, but we are far from certain in our predictions of individual patients' outcomes. Stephen Jay Gould's remarkable story, retold in this issue of Virtual Mentor, is ultimately about this challenge. How does one "prognosticate," as physicians call this practice, and how does one do so accurately? How does prognostication help patients with their decision making and improve their lived experience?

Truthfully, there are two distinct challenges here: One is prognostication itself, and the other is communication of these predictions. As difficult it is to get the former correct, physicians are perhaps worse at the latter. For example, older patients with acute myeloid leukemia (AML) tend to grossly overestimate their likelihood of cure, despite extensive discussions with their physicians about survival. In one study, 74 percent of patients expected cure rates above 50 percent, while 89 percent of physician estimates were at 10 percent or less [1]. What remains unclear, however, is why this "prognostic discordance" between patients and their doctors persists. Is it a problem of patients' "numeracy," their ability to understand statistical and numerical information and then contextualize it with regards to their diagnosis? Or is it a problem regarding the ways in which doctors communicate this information to patients? Perhaps is it both.

To date, physicians are still learning how best to provide prognostic information to patients, at least in terms of how specific communication behaviors translate into patient understanding. Published literature does, however, offer some hints about what factors might be important, along with evidence of several shortcomings in current practice. For example, an analysis of recorded conversations about prognosis found a positive correlation between the number of pessimistic statements made by

oncologists and the likelihood of patients exhibiting prognostic concordance with them [2].

For prognoses to be understood and be useful, though, they must actually be rendered. Unfortunately, several studies suggest that physicians provide prognostic estimates less frequently than one might hope, particularly in cases of terminal disease. This has been shown in a large population-based study in Europe [3], a large survey study of U.S. oncologists [4], and a study of physicians at several Chicago hospices [5]. The latter is particularly surprising, since prognoses in the hospice setting are poor by definition; one cannot even enroll without a reasonable likelihood of dying within 6 months.

When physicians do attempt to convey prognostic information, poor numeracy and lack of statistical understanding stand as well-recognized barriers to comprehension for patients and families. One study about prognostic understanding in patients with early-stage breast cancer showed that 73 percent lacked adequate understanding of the term "median survival" and that 33 percent believed their oncologist could actually predict their individual outcome [6]. There were also significant differences in patients' preferences for information; 43 percent preferred positively framed messaging ("chance of cure") whereas 33 percent preferred negatively framed phraseology ("chance of relapse"). How should an oncologist communicate prognosis then, if many patients cannot understand our standard ways of conceptualizing it numerically, and when each person has different preferences and needs with regards to framing?

I wish I had the answers to these questions; it would make me a better oncologist and help me better address the needs of my patients.

In oncology, in particular, there is a palpable tendency to think that the role of a physician is to "maintain hope." However, many oncologists equate brutal prognostic honesty with the notion of "taking away hope." Hope is essential, no doubt. Patients' outlook and attitude is of utmost importance when facing a cancer diagnosis, as Gould so eloquently demonstrates for us in his essay. Thus, it comes as no surprise that Gould's oncologist was somewhat evasive in discussing the literature on mesothelioma.

Hope is not the only important consideration for patients, though, and the emphasis on maintaining hope raises several important questions. Does a positive outlook allow patients to live better or easier? Does it somehow result in different or better decisions? Or does it actually make things worse, leading patients to choose aggressive therapies that significantly impair quality of life, even when these treatments are overwhelmingly unlikely to result in significant benefit? Does it affect patients' decisions about transitions to palliative care? Might this result in more aggressive care at the end of life, including hospitalization, intubation, and resuscitation? I worry that prognostic avoidance as a strategy to "maintain hope" yields many of these negative results. As a palliative care fellow, I saw this happen

countless times; a well-intentioned avoidance of giving bad news can lead to much misery at the end of life.

James Tulsky, a palliative care physician who studies patient-doctor communication, encourages us to think differently about hope [7]. Our tendency is to think of hope in terms of an expectation about a specific desired outcome, such as cure. This definition creates the tension we feel between prognostic disclosure and the maintenance of hope in oncology practice. Tulsky suggests that both patients and doctors should reimagine hope in terms of "trust" and "reliance," an older definition of "hope" that can still be found in many dictionaries [8]. This conceptualization realigns the interests of the patient and oncologist such that frank prognostic disclosure is not only acceptable but also likely to facilitate more informed decision making and perhaps even better planning for the future. Indeed, this is what patients say they want.

In one survey study of patients with incurable cancer, 98 percent said they wanted their doctor to be realistic [9]. And when asked about hope, they associated hopefulness with 3 specific physician behaviors: offering the most up-to-date treatment, being very knowledgeable about their diagnosis, and providing reassurance that pain will be controlled regardless of outcome. Notice that none of these involves an expectation of cure or a desire for "false hope." These behaviors very much mirror Tulsky's conception of hope, relating to issues of trust and reliance in the patient-doctor relationship.

In this sense, the prognostic message is not as awful as we might think. Frank prognostic disclosure hardly squelches hope, if hope is about the journey and the process rather than any particular expected outcome, and it may be better for helping patients make informed choices consistent with their values and preferences.

As Gould warns, "the median isn't the message." I wholeheartedly agree. But the message isn't as mean as we may think.

References

1. Sekeres MA, Stone RM, Zahrieh D, et al., Decision-making and quality of life in older adults with acute myeloid leukemia or advanced myelodysplastic syndrome. *Leukemia.* 2004;18(4):809-816.
2. Robinson TM, Alexander SC, Hays M, et al. Patient-oncologist communication in advanced cancer: predictors of patient perception of prognosis. *Support Care Cancer*. 2008;16(9):1049-1057.
3. Costantini M, Morasso G, Montella M, et al. Diagnosis and prognosis disclosure among cancer patients. Results from an Italian mortality follow-back survey. *Ann Oncol*. 2006;17(5):853-859.
4. Daugherty CK, Hlubocky FJ. What are terminally ill cancer patients told about their expected deaths? A study of cancer physicians' self-reports of prognosis disclosure. *J Clin Oncol*. 2008;26(36): 5988-5993.

5. Lamont EB, Christakis NA. Prognostic disclosure to patients with cancer near the end of life. *Ann Intern Med*. 2001;134(12):1096-1105.
6. Lobb EA, Butow PN, Kenny DT, et al. Communicating prognosis in early breast cancer: do women understand the language used? *Med J Austr*. 1999;171(6):290-294.
7. Tulsky JA. Hope and hubris. *J Pall Med*. 2002;5(3):339-341.
8. Hope. *Merriam-Webster Dictionary*, online edition. http://www.merriam-webster.com/dictionary/hope. Accessed October 29, 2012.
9. Hagerty RG, Butow PN, Ellis PM, et al. Communicating with realism and hope: incurable cancer patients' views on the disclosure of prognosis. *J Clin Oncol*. 2005;23(6):1278-1288.

Thomas W. LeBlanc, MD, MA, is a senior fellow in medical oncology at Duke University in Durham, North Carolina. He is fellowship-trained in hospice and palliative medicine, and his clinical practice focuses on the care of patients with hematologic malignancies.

Related in VM

The Median Isn't the Message, January 2013

When Patients Seem Overly Optimistic, July 2012

Hoping for the Best, Preparing for the Worst, August 2005

Virtual Mentor
American Medical Association Journal of Ethics
January 2013, Volume 15, Number 1: 86-89.

OP-ED

A Call to Integrate Ethics and Evidence-Based Medicine

Ross E.G. Upshur, MD, MSc

It is generally accepted that medicine, and indeed all health care, should be based on or informed by evidence. Yet this truism belies the complexities and nuances involved in understanding what we mean by evidence and how it serves as a base to medicine. How evidence and ethics interrelate is an often neglected and overlooked dimension of evidence-based approaches to health care. While the laudable aim of evidence-based medicine (EBM) is to avoid biased and arbitrary decisions in medical care, the more important task is to manage the inherent uncertainty that is constitutive of the practice of medicine—a task that calls for considering how ethics should be integrated with evidence.

What Is Evidence-Based Medicine?

It seems obvious to expect that medicine be based on evidence; otherwise it would rely on caprice, whim, or arbitrary authority. Yet EBM established itself as the dominant approach to clinical medicine only in the late twentieth century. It has been rapidly taken up in all clinical fields and has been regarded as revolutionary by some and, by others, as the unrivaled standard by which medicine is to be practiced [1]. The descriptor "evidence-based" is now ubiquitous, but there are multiple claims to approaches being evidence-based, and these show considerable heterogeneity [2].

EBM emerged in the early 1990s as an approach to clinical medicine that sought to orient clinician decision making away from reliance on experience, authority, or pathophysiological inference to reliance on rigorously designed, clinically based, research [3]. The practice of EBM consists of five sequential steps [4].

1. Asking a focused question,
2. Finding the research evidence relevant to that question through a systematic search of the literature,
3. Critically appraising the results of the search for its validity and applicability to the question,
4. Applying the results of this search in practice and integrating it with patient preferences and values, and
5. Evaluating the impact of this decision in the care of the patient and in terms of performance.

Supporting this process is a graded hierarchy of evidence that is supposed to link to the strength of clinical recommendations. This hierarchy gives preference to systematic reviews, meta-analysis and randomized controlled trials for most therapeutic decisions.

What Is "Evidence"?
Proponents of EBM seldom define what they mean by "evidence." The closest approximation to a definition is provided by Brian Haynes. He writes:

> …it is hardly surprising that the term evidence-based medicine is confusing to many, who do not appreciate that its evidence is narrowly defined as having to do with systematic observations from certain types of research. The very name has been an impediment to getting across its main objective, namely, that healthcare research is nowadays producing important results that, if applied, can benefit patients more than the treatments that clinicians are experienced in recommending. Using the technical definition of EBM, evidence from healthcare research is a modern, never-before-available complement to traditional medicine. Perhaps a better name would be "certain-types-of-high-quality-and-clinically-relevant-evidence-from-health-care-research-in-support-of-health-care-decision-making"...an accurate but mind-numbing descriptor [5].

It is important to be clear on just what emerges from "systematic observations from certain types of research" and what it means to base treatment decisions on this type of evidence. Most research studies provide estimates of measures of effect expressed in statistical terms with ranges of uncertainty associated with these estimates. They are claims to truth, but not necessarily true, nor inevitably applicable in any individual case. Evidence has probative status, but can be overturned, displaced, or superseded in light of new findings. The probabilistic demonstration of an effect does not translate into the necessity of this effect occurring in any particular case at hand.

In essence, evidence—narrowly defined or not—is a provisional departure point in the consideration of whether or not a particular course of action is to be taken in any clinical context.

There are other properties of evidence worth considering. Access to evidence is not equally available to all specialties in medicine and health care. Ethical considerations dictate that certain types of evidence never be available (an idea most famously captured by R.A. Fisher's assertion that no randomized controlled trial could be performed to assess the harm due to cigarettes) [6]. There will always necessarily be significant swaths of uncertainty and large "grey zones" in practice [7, 8].

Integrating Evidence and Ethics
In my interpretation, evidence-based approaches are one manner by which to manage uncertainty. Uncertainty has various dimensions. One dimension stems from lack of knowledge when knowledge is available, and EBM seeks to inculcate lifelong learning strategies to reduce this. The more significant sense of uncertainty relates to incomplete knowledge. Research attempts to fill or reduce this incompleteness—however, this is an eternal undertaking.

At the time when clinical decisions are required, appropriate evidence may not be available or there may not be agreement on the interpretation of existing evidence. These points illustrate why integrating ethics and evidence-based approaches is essential.

That EBM has significant ethical implications has been well described in the literature. I direct readers to two particularly insightful accounts, one by Ian Kerridge and colleagues and another by Mona Gupta [9, 10]. Kerridge et al. point out that, in the practice of EBM, certain outcomes that are poorly measured or cannot be measured are either neglected or lessened in significance. Measurement may crowd out meaning. They also note that EBM is poorly adapted to explicating and mediating conflicting values and interests in the creation and dissemination of evidence. Gupta carefully examines the ethical assumptions of EBM and notes that EBM assumes following the five steps outlined above is the optimal method to securing the best course of action for a particular decision. It presumes that there exists a moral imperative to practice these five steps. However, as Gupta demonstrates, there are good reasons to question this moral imperative.

EBM may give the false impression of its own value neutrality, an interpretation reinforced by a critical reading of the steps of EBM. The values of those posing a clinical question and the manner in which these values can influence the interpretation of the evidence are not considered. In a value-neutral view of EBM, the research literature is regarded as a set of accumulated "facts." This is very likely a naive conceptualization, particularly in light of a wide range of influences on the published literature, some of which (e.g., ghostwriting) can easily escape the most sophisticated critical appraisal tools. In much discussion of EBM, values appear as something only relevant to patients (alongside "preferences") to be integrated with a seemingly value-free clinical judgment informed by evidence. That the facts themselves contain claims about desired human states of affairs seems to have escaped notice. Indeed, the distinction between facts and values is often "collapsed" in clinical research [11].

Thus, the relationship between ethics and EBM is by no means straightforward or unproblematic. It deserves attention for the educators who have avidly introduced EBM in medical curricula at the undergraduate and postgraduate level, but seldom include discussions of its ethical dimensions. In my view, there is much to be gained by countering this divorce and taking steps for greater integration. This would entail a more explicit focus on the implications of uncertainty for clinical practice and time spent exploring reflexivity with respect to the values of the clinician in the steps of EBM. A promising direction that may resonate with practitioners is emerging from virtue theory, in which virtues such as curiosity, courage, honesty, and humility leading to prudence and practical wisdom are seen as promoting excellence in practice [12]. Within this framework, attention to evidence, however conceived, is linked to commitment to care. Rather than being seen as distinct spheres, ethics and evidence become part of an integrated whole.

References

1. Reilly BM. The essence of EBM. *BMJ*. 2004;329(7473):991-992.
2. Upshur RE. Evidence-based medicine, reasoned medicine or both? *J Eval Clin Pract*. 2006;12(4):420-422.
3. Evidence-Based Medicine Working Group. Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA*. 1992;268(17):2420-2425.
4. Centre for Evidence Based Medicine web site. http://www.cebm.net/index.aspx?o=1023. Accessed October 14, 2012.
5. Haynes RB. What kind of evidence is it that evidence-based medicine advocates want health care providers and consumers to pay attention to? *BMC Health Services Res*. 2002;2:3. http://www.biomedcentral.com/1472-6963/2/3. Accessed October 14, 2012.
6. Upshur RE. Seven characteristics of medical evidence. *J Eval Clin Pract*. 2000;6(2):93-97.
7. Naylor CD. Grey zones of clinical practice: some limits to evidence-based medicine. *Lancet*. 1995;345(8953):840-842.
8. Mike V. Outcomes research and the quality of health care: the beacon of an ethics of evidence. *Eval Health Prof*. 1999;22(1):3-32.
9. Kerridge I, Lowe M, Henry D. Ethics and evidence based medicine. *BMJ*. 1998;316(7138):1151-1153.
10. Gupta M. A critical appraisal of evidence-based medicine: some ethical considerations. *J Eval Clin Pract*. 2003;9(2):111-121.
11. Putnam H. *The Collapse of the Fact/Value Dichotomy and Other Essays*. Boston: Harvard University Press, 2002.
12. Marcum JA. The epistemically virtuous clinician. *Theor Med Bioeth*. 2009;30(2):249-265.

Ross E.G. Upshur, MD, MSc, is Canada Research Chair in Primary Care Research and a professor in the Department of Family and Community Medicine, the Dalla Lana School of Public Health, the Institute for the History and Philosophy of Science and Technology, the Institute for Clinical Evaluative Sciences, and the Centre for the Environment at the University of Toronto in Ontario. He is a primary care physician at Sunnybrook Health Sciences Centre. His research interests encompass the philosophy of medicine (both ethics and epistemology), primary care and chronic disease management (particularly in aging populations), and public health control of communicable diseases.

Related in VM

Teaching Critical Appraisal of Medical Evidence, January 2013
Evidence-Based Guidelines and Quality Measures in the Care of Older Adults, January 2013
The Limitations of Evidence-Based Medicine—Applying Population-Based Recommendations to Individual Patients, January 2011

Virtual Mentor
American Medical Association Journal of Ethics
January 2013, Volume 15, Number 1: 90-93.

Suggested Readings and Resources

Philosophy of EBM

Dickersin K, Straus SE, Bero LA. Evidence based medicine: increasing, not dictating, choice. *BMJ*. 2007;334(suppl 1):s10-s10.

Eddy DM. Evidence-based medicine: a unified approach. *Health Aff(Millwood)*. 2005;24(1):9-17.

Evidence-Based Medicine Working Group. Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA*. 1992;268(17):2420-2425.

Feinstein AR, Horwitz RI. Problems in the "evidence" of "evidence-based medicine".
Am J Med. 1997;103(6):529-535.

Goodman KW. Comment on M.R. Tonelli, "The challenge of evidence in clinical medicine." *J Eval Clin Pract*. 2010;16(2):390-391.

Goodman KW. *Ethics and Evidence Based Medicine: Fallibility and Responsibility in Clinical Science*. Cambridge: Cambridge University Press; 2003.

Guyatt GH. Evidence-based medicine. *ACP J Club*. 1991;1(2): A-16.

Guyatt G, Cook D, Haynes B. Evidence based medicine has come a long way. *BMJ*. 2004;329(7473):990-991.

Guyatt GH, Rennie D. *Users' Guide to the Medical Literature*. Chicago, IL: American Medical Association; 2001.

Hartzband P, Groopman J. The new language of medicine. *New Engl J Med*. 2011;365(15):1372-1373.

Haynes RB. What kind of evidence is it that Evidence-Based Medicine advocates want health care providers and consumers to pay attention to? *BMC Health Serv Res*. 2002;2(1):3.

Haynes RB, Devereaux PJ, Guyatt GH. Clinical expertise in the era of evidence-based medicine and patient choice. *ACP J Club*. 2002;136(2):A11-14.

Henry SG, Zaner RM, Dittus RS. Viewpoint: moving beyond evidence-based medicine. *Acad Med*. 2007;82(3):292-297.

Montori VM, Guyatt GH. Progress in evidence-based medicine. *JAMA*. 2008;300(15):1814-1816.

Straus S, Haynes B, Glasziou P, Dickersin K, Guyatt G. Misunderstandings, misperceptions, and mistakes. *Evid Based Med*. 2007;12(1):2-3.

Tonelli MR. The challenge of evidence in clinical medicine. *J Eval Clin Pract*. 2010;16(2):384-389.

Tonelli MR. The philosophical limits of evidence-based medicine. *Acad Med*. 1998;73(12):1234-1240.

Quill TE, Holloway RG. Evidence, preferences, recommendations—finding the right balance in patient care. *New Engl J Med*. 2012;366(18):1653-1655.

Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ*. 1996;312(7023):71-72.

Upshur R. Looking for rules in a world of exceptions: reflections on evidence-based practice. *Perspect Biol Med*. 2005;48(4):477-489.

Teaching Evidence-Based Medicine

Guyatt GH, Meade MO, Jaeschke RZ, Cook DJ, Haynes RB. Practitioners of evidence based care. Not all clinicians need to appraise evidence from scratch but all need some skills. *BMJ*. 2000;320(7240):954-955.

Hatala R, Keitz SA, Wilson MC, Guyatt G. Beyond journal clubs. Moving toward an integrated evidence-based medicine curriculum. *J Gen Intern Med*. 2006;21(5):538-541.

JAMAevidence web site. http://www .jamaevidence.com. Accessed December 17, 2012.

Rao G, Kanter SL. Physician numeracy as the basis for an evidence-based medicine curriculum. *Acad Med*. 2010;85(11):1794-1749.

Schwartz MD, Dowell D, Aperi J, Kalet AL. Improving journal club presentations, or, I can present that paper in under 10 minutes. *Evid Based Med*. 2007;12(3):66-68.

Straus SE, Richardson SW, Glasziou P, Haynes RB. *Evidence-Based Medicine: How To Practice and Teach EBM*. 4th ed. New York: Churchill Livingstone, 2010.

N-of-1 Trials
Gabler NB, Duan N, Vohra S, Kravitz RL. N-of-1 trials in the medical literature: a systematic review. *Med Care*. 2011;49(8):761-768.

Guyatt GH, Keller JL, Jaeschke R, Rosenbloom D, Adachi JD, Newhouse MT. The n-of-1 randomized controlled trial: clinical usefulness. Our three-year experience. *Ann Intern Med*. 1990;112(4):293-299.

Guyatt GH, Sackett D, Adachi J, et al. A clinician's guide for conducting randomized trials in individual patients. *CMAJ*. 1988;139(6):497-503.

Guyatt GH, Sackett D, Taylor DW, Ghong J, Roberts R, Pugsley S. Determining optimal therapy—randomized trials in individual patients. *N Engl J Medicine*. 1986;314(14):889-892.

Kravitz RL, Duan N, Niedzinski EJ, Hay MC, Subramanian SK, Weisner TS. What ever happened to n-of-1 trials? Insiders' perspectives and a look to the future. *Milbank Q*. 2008;86(4):533-555.

Nikles CJ, Mitchell GK, Del Mar CB, Clavarino A, McNairn N. An n-of-1 trial service in clinical practice: testing the effectiveness of stimulants for attention-deficit/hyperactivity disorder. *Pediatrics*. 2006;117(6):2040-2046.

Scuffham PA, Nikles J, Mitchell GK, et al. Using N-of-1 trials to improve patient management and save costs. *J Gen Intern Med*. 2010;25(9):906-913.

Vohra S, Shamseer L, Sampson M. Efficacy research and unanswered clinical questions. *JAMA*. 2011;306(7):709.

Risk/Benefit Communication
Elwyn G, Frosch D, Thomson R, et al. Shared decision making: a model for clinical practice. *J Gen Intern Med*. 2012;27(10):1361-1367.

Fagerlin A, Zikmund-Fisher BJ, Ubel PA. Helping patients decide: ten steps to better risk communication. *J Natl Cancer Inst*. 2011;103(19):1436-1443.

Fischhoff B, Brewer NT, Downs JS, eds. *Communicating Risks and Benefits: An Evidence Based User's Guide*. Silver Spring, MD: Food and Drug Administration; 2012.
http://www.fda.gov/downloads/AboutFDA/ReportsManualsForms/Reports/UCM268069.pdf. Accessed December 17, 2012.

Lloyd FJ, Reyna VF. Clinical gist and medical education. *JAMA*. 2009;302(12):1332-1333.

Paling J. Strategies to help patients understand risks. *BMJ*. 2003;327(7417):745-748.

EBM and Clinical Guidelines
Boyd CM, Darer J, Boult C, Fried LP, Boult L, Wu AW. Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. *JAMA*. 2005;294(6):716-724.

Hurwitz B. How does evidence based guidance influence determinations of medical negligence? *BMJ*. 2004;329(7473):1024-1028.

Muir Gray JA. Evidence based policy making. *BMJ*. 2004;329(7473):988-989.

Rosoff AJ. The role of clinical practice guidelines in healthcare reform: an update. *Ann Health Law*. 2012;21:21-281.

Sniderman AD, Furberg CD. Why guideline-making requires reform. *JAMA*. 2009;301(4):429-431.

Walters CL, Davidowitz NO, Heineken PA, Covinsky KE. Pitfalls of converting practice guidelines into quality measures: lessons learned from a VA performane measure. *JAMA*. 2004;291(20):2466-2470.

Health Advocacy Organizations and EBM
Rothman SM. Health advocacy organizations and evidence-based medicine. *JAMA*. 2011;305(24):2569-2570.

Issues with the "Evidence Base"
Ioannidis JP. An epidemic of false claims. *Sci Am*. 2011;304(6):16-16.

Ioannidis JP. Why most published research findings are false. *PLoS Medicine*. 2005;2(8):e124.

Prasad V, Cifu A, Ioannidis JP. Reversals of established medical practices: evidence to abandon ship. *JAMA*. 2012;307(1):37-38.

Rennie D. When evidence isn't: trials, drug companies and the FDA. *J Law Policy*. 2007;15(3):991.

Virtual Mentor
American Medical Association Journal of Ethics
January 2013, Volume 15, Number 1: 94-97.

About the Contributors

Theme Issue Editor
Matthew Rysavy is a medical student at the University of Iowa's Carver College of Medicine and a PhD candidate in epidemiology at the university's College of Public Health in Iowa City. His interests include clinical epidemiology and medical education. He has been active in the design and implementation of the evidence-based medicine curriculum at the Carver College of Medicine.

Contributors
Valarie Blake, JD, MA, is a senior research associate for the American Medical Association Council on Ethical and Judicial Affairs in Chicago. Ms. Blake completed the Cleveland Fellowship in Advanced Bioethics, received her law degree with a certificate in health law and concentrations in bioethics and global health from the University of Pittsburgh School of Law, and obtained a master's degree in bioethics from Case Western Reserve University. Her research focuses on ethical and legal issues in assisted reproductive technology and reproductive tissue transplants, as well as regulatory issues in research ethics.

Kristin L. Carman, PhD, co-directs the Health Policy and Research group at the American Institutes for Research in Washington, DC. She leads the Community Forum, an initiative of the Agency for Healthcare Research and Quality's Effective Health Care program that will evaluate and develop approaches to expand the participation of the public and various stakeholder groups in improving the effectiveness of health care.

Martha Carvour, MD, PhD, is a first-year resident in internal medicine at the University of Texas Southwestern Physician Scientist Training Program in Dallas. Prior to residency, Dr. Carvour completed her MD and PhD (epidemiology) degrees at the University of Iowa, where she first became involved in the design, development, and delivery of an evidence-based medicine curriculum.

Paul J. Christine, MPH, is an MD-PhD student at the University of Michigan Medical School and School of Public Health in Ann Arbor. He is pursuing a doctoral degree in epidemiology and is broadly interested in the relationship between population health and clinical practice.

William Dale, MD, PhD, is the section chief of Geriatrics and Palliative Medicine at the University of Chicago, where he established the Specialized Oncology Care and Research in the Elderly (SOCARE) clinic. He has published more than 50 articles on medical decision making, geriatric oncology, and quality of life for older adults,

particularly older adults with cancer. Dr. Dale is a board-certified geriatrician with a PhD in health policy.

Stephen Jay Gould, PhD, (1941-2002) was a evolutionary biologist and natural historian, professor at Harvard University and New York University, and acclaimed author of essays and popular science books.

Jodi Halpern, MD, PhD, is an associate professor of bioethics and medical humanities in the Joint Medical Program and the School of Public Health at the University of California, Berkeley. A psychiatrist with a doctorate in philosophy, she is the author of *From Detached Concern to Empathy: Humanizing Medical Practice* (paperback 2011, Oxford University Press).

Jessica Waddell Heeringa, MPH, is a research analyst at Mathematica Policy Research. She co-wrote this manuscript while working at the American Institutes for Research on the Agency for Healthcare Research and Quality's Community Forum. Her research interests include health reform implementation, mental health care, and health disparities.

Dien Ho, PhD, is an associate professor of philosophy and health care ethics at Massachusetts College of Pharmacy and Health Sciences. His research focuses primarily on the ethics of organ transplantation, reproductive autonomy, pharmacist ethics, and theoretical reasoning.

Chetan Huded, MD, is a resident physician in internal medicine at Northwestern Memorial Hospital in Chicago. He was last year's winner of the Rambach Award, given to the most meritorious resident of the intern class. He is interested in cardiology.

Lauris C. Kaldjian, MD, PhD, is director of the Program in Bioethics and Humanities and a professor in the Department of Internal Medicine at the University of Iowa Carver College of Medicine in Iowa City. Dr. Kaldjian practices general internal medicine, and his research has focused on goals of care and patients' end-of-life treatment preferences, physician disclosure of medical errors, ethics education, and the role of philosophical and religious beliefs in clinical decision making.

Richard L. Kravitz, MD, MSPH, is professor and co-vice chair (research) in the Department of Internal Medicine at University of California, Davis. He is also co-editor in chief of the *Journal of General Internal Medicine*.

Thomas W. LeBlanc, MD, MA, is a senior fellow in medical oncology at Duke University in Durham, North Carolina. He is fellowship-trained in hospice and palliative medicine, and his clinical practice focuses on the care of patients with hematologic malignancies.

Vinay Prasad, MD, is chief fellow in medical oncology at the National Cancer Institute in Bethesda, Maryland. Dr. Prasad coined the term "medical reversal" and has published on the subject in the *Journal of the American Medical Association* and *Archives of Internal Medicine*, among others. He is interested in the adoption of rational methods in medical practice.

Salima Punja, BSc, is a PhD candidate in the Department of Medicine at the University of Alberta in Edmonton.

Erika Ramsdale, MD, is in the final year of her hematology/oncology fellowship at the University of Chicago. She has also completed fellowships in geriatric medicine and clinical medical ethics and is planning to pursue a career in geriatric oncology. Her research focuses on clinical decision making in older adults with cancer, particularly those with complex comorbidities or frailty.

Valerie F. Reyna, PhD, is a professor at Cornell University and Weill Cornell Medical College and co-director of Cornell's University Magnetic Resonance Imaging Facility and its Center for Behavioral Economics and Decision Research. Her recent work has focused on numeracy, medical decision making, risk communication, risk taking, neurobiological models of decision making, and neurocognitive impairment.

Jill Rosno, MD, is a resident physician in internal medicine at Northwestern Memorial Hospital in Chicago. A graduate of Dartmouth Medical School, she is interested in general internal medicine.

Joanna E. Siegel, ScD, is a senior scientist in the Agency for Healthcare Research and Quality's Center for Outcomes and Evidence. She coordinates the Community Forum Project, which includes a demonstration that is evaluating use of deliberative methods for providing public input on the use of medical evidence in health care decision making, and projects to enhance patient and other stakeholder involvement in AHRQ's Effective Health Care program.

Ross E.G. Upshur, MD, MSc, is Canada Research Chair in Primary Care Research and a professor in the Department of Family and Community Medicine, the Dalla Lana School of Public Health, the Institute for the History and Philosophy of Science and Technology, the Institute for Clinical Evaluative Sciences, and the Centre for the Environment at the University of Toronto in Ontario. He is a primary care physician at Sunnybrook Health Sciences Centre. His research interests encompass the philosophy of medicine (both ethics and epistemology), primary care and chronic disease management (particularly in aging populations), and public health control of communicable diseases.

Sunita Vohra, MD, MSc, a pediatrician and clinician scientist with a master's degree in clinical epidemiology and fellowship training in clinical pharmacology, is the leader of a Canadian Institutes of Health Research project investigating n-of-1 trials.

A professor in the Faculty of Medicine and School of Public Health at the University of Alberta in Edmonton, Dr. Vohra is the founding director of Canada's first academic pediatric integrative medicine program, the Complementary and Alternative Research and Education (CARE) program at the Stollery Children's Hospital. Dr. Vohra is also the program director for Canada's first fellowship program in pediatric integrative medicine and the founding director of the Canadian Pediatric CAM Network (PedCAM).

Evan A. Wilhelms is a PhD student in the Department of Human Development at Cornell University in Ithaca, New York. His research interests have focused on models of decision making in cognitive development and behavioral economics.

Ariel L. Zimerman, MD, PhD, recently finished his doctorate on the history of evidence-based medicine in the Graduate Program in Science, Technology and Society at Bar-Ilan University in Ramat-Gan, Israel. His research interests are the history of medical epistemology, the introduction of quantification practices to clinical medicine, and the history of medicine in the late twentieth century. He is also a practicing physician and board certified in obstetriccs and gynecology.