COMMENTS

We welcome comments by our readers reflecting agreement or disagreement with the material published in this section and, at the discretion of the Editor-in-Chief, will publish such comments. We would like to emphasize that material published anywhere in EBM, including this section, does not necessarily reflect the opinions of the Society or the Editor-in-Chief.

Medical Education, Evidence-Based Medicine, and the Disqualification of Physician-Scientists

BURTON E. SOBEL[1] AND MARK A. LEVINE

Fletcher Allen Health Care, Medicine Health Care Service, 111 Colchester Avenue, Burlington, Vermont 05401

The social transformation of American medicine (1), now virtually complete, has been driven by prodigious forces, including politics (the pharmaceutical industry as public enemy number one...ergo vote for me to save you money); economics (no society can sustain total health care costs in excess of 12% of its gross domestic product [2]...ergo HMOs, single payers, gatekeepers, docs in a box, proliferation of nurse practitioners, and other cost-cutting expedients); and technology ("we must make sure that no life-saving discovery is locked up in the laboratory [3]...ergo electron beam CT, MRI, and percutaneous coronary interventions in every community hospital). Revolutions such as this one are driven by ideology. They are shaped by words. As the March Hare said to Alice "You should say what you mean," to which she artfully responded, "I do, at least I mean what I say - that's the same thing you know" (4). But, of course, the Hatter was right. It isn't the same at all, anymore than "I see what I eat" is the same thing as "I eat what I see."

Some of the architects of the transformation of American medicine have used, and misused, words egregiously. In doing so, they have promulgated distortions, altered medical education as well as medical practice, and disqualified physician-scientists. One salient example is "evidenced-based medicine." Are we to forget Flexner, who in 1910, set modern medical education on a scientifically sound footing and who demanded that what was taught be predicated on hypothesis, experiment, analysis, refinement, refutation or validation of hypothesis, and, in aggregate, scientific method and scrutiny (5)? Are we to forget Hippocrates himself, the father of "wholly objective description of experience" who made "a clear distinction between the priesthood and the profession, and ... chose the latter" (6)? In the context of their tradition, "evidence-based medicine" is hardly a novel idea.

The current use of the term "evidence-based medicine" is not without redeeming features. The term was originally spawned to emphasize the need for judicious use of current, objective information in making decisions about the care of individual patients (7). It was coined to encourage proficiency in judgments by individual clinicians based not only on "experience," but also on experience informed by results acquired in systematic research. However, the paradigm exemplified by evidence-based medicine de-emphasizes the resort to traditional scientific reasoning predicated on an understanding of physiology and pathophysiology. Its use reflected the belief that knowledge of basic mechanisms underlying disease and pathophysiological principles, though necessary, are insufficient as guides for clinical practice (8). The founding fathers of evidence-based medicine acknowledge that the more the "evidence" is inadequate, the greater is the need to rely on knowledge of pathophysiological principles in clinical problem solving. Yet, they value knowledge of pathophysiology primarily as a means for generalizing results from clinical studies of highly selected populations to clinical practice as a whole. Thus, they subscribe to the notion that insights based on knowledge of pathophysiology are of secondary value com-

[1] To whom requests for reprints should be addressed at University of Vermont, Colchester Research Facility, 208 South Park Drive, Colchester, VT 05446. E-mail: burton.sobel@vtmednet.org

0037-9727/01/2268-0000$15.00

pared with those derived from observational and epidemiologic studies and clinical trials.

An unfortunate and perhaps unintended consequence of evidence-based medicine is excessive reliance on a tool box of processes such as appraisal skills, meta-analyses, and practice guidelines as a reaction formation against the "over-reliance" on "irrelevant" science, pathophysiological principles, and results of mechanistic research. The consequence, minimization of the value of critical thinking, is not only anti-intellectual, it is also prone to lead to wrong decisions when the specific characteristics of a given patient or set of clinical circumstances have profound diagnostic or therapeutic implications reflecting cause/consequence and mechanistic considerations that are overlooked by "worshipping at the shrine of the clinical trial" (9).

Another example of pejorative language that undermines recognition of the value of science in medicine is the term "outcomes research." Are we to believe that x-ray crystallography was not based on outcomes (x-ray scatter) in experiments confirming the helical structure of DNA? Are we to believe that Minot and Murphy did not assess outcome after their pioneering administration of liver extracts to correct the then unknown intrinsic factor deficiency responsible for pernicious anemia (10)? Are we to believe that Banting and Best did not assess outcome after their courageous administration of pancreatic extracts to reverse life-threatening diabetic ketoacidosis (11)?

The "new outcomes" are quantified with new "tools" such as meta-analysis. Despite the utility of this approach when applied rigorously with adequate documentation of homogeneity, it represents a 180-degree turn from a previously held biostatistical tenet. Classical biostatisticians have implied that "meta-analysis is to analysis as metaphysics is to physics." Black-box thinking surely is convenient. No matter that a meta-analysis with two unrecognized and undifferentiated subsets, e.g., nonacetylators and acetylators, one of which may exhibit strikingly positive and one equally strikingly negative changes in response to administration of a putative therapeutic agent, would lead to the spurious conclusion that a given drug exerts no effect. In fact, the drug may have exerted positive, therapeutic effects in one group and toxic effects in another because of unbeknownst mechanisms of action obfuscated by the meta-analytic approach.

It is this "statistical" expression of the results of multiple studies with a single point estimate that is the most criticized aspect of meta-analysis. When the outcomes of 12 large randomized assignment patient-controlled trials were compared with the results of previously published meta-analyses of studies addressing the same questions, results obtained with the meta-analyses were found to be inaccurate in 35% of instances. Often, substantial discrepancies were evident as well with respect to the magnitude of effects (12).

Like "evidence-based medicine," the terms "surrogate endpoints" and "outcomes research" were coined to influence not only thinking about patients, but also the design of research (13). We are told that blood pressure lowering is a surrogate (substitute) for reduction of mortality, but it is not. Elevated blood pressure may be a determinant of vascular disease, in turn a contributor to mortality in some patients. However, the lowering of blood pressure per se, e.g., as a consequence of occult hemorrhage (or iatrogenic blood letting) is certainly not a "substitute" for reduction of mortality. A decrease in blood pressure consequent to an intervention is, of course, a physiologic outcome variable that may be a primary endpoint. Assessment of such an endpoint can clarify mechanisms underlying disease and outcomes.

Outcomes research refers to the assessment of consequences of medicine as it is usually practiced, generally not under the stringent conditions of randomized patient assignment, controlled clinical trials. It employs the use of prespecified databases assembled for diverse socioeconomic and administrative purposes, not necessarily in a fashion designed to facilitate valid inquiry regarding specific questions addressed by clinical trials. Outcomes research is championed as a means for improving the quality and effectiveness of delivery of medical care and hence the health of patients (14). In addition to extending knowledge regarding the efficacy of a potentially therapeutic intervention, perhaps best defined in clinical trials, outcomes research may shed light on the anticipated and nonanticipated consequences of the intervention in clinical practice. Although the information provided is intended to be complementary, it is often used hierarchically. The most striking example is its use in formulating and implementing clinical practice guidelines.

Created under the mantra that variation in utilization and practice is bad and that cost containment is always good, practice guidelines are soon followed by algorithms (step-by-step decision-making and problem-solving prescriptions). Many such algorithms are complex and lengthy. Many fail to simplify patient care. Even worse, excessive reliance upon them and routine knee jerk-like applications fail to teach and can dull the skills in critical thinking required for a truly astute clinician. Paradoxically, such reliance can violate the Hippocratic admonition to "first do no harm." We believe that optimal care is facilitated by knowledge of the science underlying the practice of medicine. Individual patients exhibit profound biological variation and diverse constellations of interacting clinical phenomena. Compassion and elegant clinical judgment can be compromised by decisions made exclusively on the basis of algorithms often assembled in concert with insurers and managed care organizations.

Despite the insights provided by mechanistic research, "outcomes" require assessment for at least one compelling reason. That is the potential for drugs, not anticipated on the basis of their known mechanisms of action, to impair safety. We may know that an agent exerts a specific action (such as lowering blood pressure by dilating resistance vessels), but

we don't know, *a priori*, what else it may do. For example, if an agent capable of inducing hypotension also blunts cellular responses to infectious agents, an "outcomes" study might show that treated patients experience a greater number of, or more virulent, infections as well as reductions in blood pressure. The implications regarding the therapeutic value of the agent would obviously differ from those predicated exclusively upon mechanistic research in which the only end point analyzed was peripheral vascular resistance. Selection of a specific endpoint and inference regarding its importance often reflects observer bias. The same consideration applies to surrogate endpoints.

Surrogate endpoints are often laboratory values or physical signs used as substitutes for clinically important endpoints relevant to morbidity, functional capacity, or survival (15). They have been promulgated primarily because their use permits design of clinical studies that are efficient (smaller sample sizes) and relatively inexpensive. To be genuinely useful in relation to outcomes, the surrogate endpoints must meet criteria that link them biologically as well as statistically to specific clinical outcomes of interest (16). Reliance on surrogate endpoints with respect to prediction of outcome can be harmful or even lethal, however. Thus, although an intervention may exert a favorable outcome as judged from its effect on the surrogate endpoint, it may not improve or may even compromise health. Clinical science may be subverted as a consequence of cutting corners in research and reducing costs rather than performing definitive clinical trials with clinically relevant as well as surrogate endpoints. For example, in the Cardiac Arrhythmia Suppression Trial (CAST) study, suppression of premature ventricular complexes was induced by specific class Ic antiarrhythmic drugs tested (the surrogate end point) as had been hoped as judged from mechanistic considerations. Yet, mortality was unexpectedly increased (17).

Implications for Medical Education

The forces driving the social transformation of American medicine are immense. Unfortunately, they are also indiscriminate. Just as they have sought to curtail extravagance, they have led to a dumbing down of the profession. Numerous self-proclaimed health care commentators trivialize the nature of research. They fan the flames of unrealistic expectations for cures. They seduce the public into embracing alternative medicine without realizing that "alternative" applies to the lack of a scientific approach to assessment rather than a genuine choice rendered on a level playing field of adjudication.

A myriad of social, economic, and professional forces threaten medical education. They include the explosion in technology, heightened public expectations, and the loss of physician-scientists (18) as role models consequent to the deprofessionalization of physician to employee whose value is judged all too often by RVUs per hour rather than the quality of judgment or expertise. One other critical factor is the ascendance of "primary care internal medicine," perhaps the most difficult medical discipline to master because of its breadth, diversity, mercurial evolution, and requirement for profound attention to life-long education. The term itself conceals these difficulties, avoiding reference to general internal medicine, the genuine province of the primary care internal medicine physician. The implication is simplicity and initial triage with referral to "specialists," whom Hippocrates might have counted in the priesthood he so thoroughly rejected (19). In fact, the forces driving the social transformation of American medicine are driving the primary care physician to forsake consideration of scientifically established principles, rely on algorithms mandated by health care economists, triage patients, and mouth a litany of new objectives including market share, throughput, and productivity.

In 1965, Gordon Moore articulated Moore's law (Intel Corporation, personal communication). It described the exponential growth of computer chip capacity (and in fact, each element of computer technology) with a doubling rate at 18- to 24-month intervals. The law, greeted initially with skepticism, has proven to be remarkably prescient.

If anything, Moore's law underestimates the rate of growth of technology. Today's house officers must assimilate information that is growing comparably. The temptation is strong to resort to sound bytes, blurbs on the Internet, MD Consult, WebMD, and "authority," regardless of its stripes. The temptation to embrace algorithms as a substitute for ratiocination is enormous. So is the predilection to forsake problem solving based on elucidation and understanding of pathophysiological mechanisms. The danger is that training is replacing education. The danger is that development will be fostered of a trade rather than a profession. Unfortunately, the value of trained as opposed to educated physicians to the patients they serve declines precipitously with time. Unfortunately, none of this matters to the Captains of Managed Care.

How Can We Nurture Critical Thinking in Medical Education Programs?

Medical school and house staff education is beleaguered. The ratio of applicants to places available in first-year medical school classes is in decline. Medical school graduates are burdened by prodigious debt. More and more of our best and brightest young minds elect to enter industry rather than medical school. Our medical educational programs have often forsaken problem-solving skills, integration of basic and clinical science, and a focus on critical thinking, while embracing process and pedagogy devoted to enhancing performance on standardized tests. We need to counter those who pervert the original intentions of pioneers of evidence-based medicine and who portray basic science as a pariah, and algorithms and practice guidelines as sacred gifts from health policy gurus. "Integrative" curricula are blossoming. Yet, many well-intentioned attempts fail to adequately incorporate the seminal developments in genetics, molecular biology, and technology that offer so much prom-

ise. The public is bombarded by hyperbole implying that such advances will enhance health—now. At the same time, the managed care movement insulates the public from physician/scientists with "mid-level practitioners" deemed capable of fulfilling "primary care" needs. Medical schools must empower their graduates with skills in critical thinking and expertise in integrating advances in basic science into clinical practice if physicians are to retain any leadership role in the health care system's hierarchy. This goal can be reached only with education as opposed to training. It can be achieved only with emphasis on scientific principles and application of the scientific method as opposed to exclusive reliance on rote memory and the algorithm of the moment.

With respect to house staff education—residents are not apprentices. They must know how to search literature and assess outcomes. But, they must know also how to think critically and assimilate advances in science pertinent to their patients. Mouthing of the results of the "latest" (not necessarily the best designed) clinical trial cannot substitute for a broader epistemology with knowledge of "classics" in medical science and the logical fabric of mechanistic research and its contribution to clinical judgment. Residency programs must emphasize scholarly activity not because of Residency Review Committee requirements, but because such activity is the cauldron needed for development of critical thinking. Departments of medicine must potentiate interactions between residents and faculty, collaboration in vigorously mentored basic and clinical research, and personal participation of scientifically skilled faculty in the education of residents not because of the need to make each resident a researcher, but because participation in research is a powerful antidote to blind acceptance of the authority of the day or the algorithm of the moment. Research seminars should be part and parcel of the residency program. Medical education today is often guilty of suppression of expression of curiosity (20). A climate of research can stimulate curiosity and invigorate future clinicians as they prepare for life-long learning.

The term "primary care internal medicine" epitomizes the dumbing down of American medicine. It should be replaced by its antecedent, "general internal medicine." Residents aspiring to be generalists are perhaps those most in need of expertise in critical thinking. Recognition of biological variation and its implications, knowing how to tell whether a potential patient is "normal," commitment to evaluation and assimilation of results of original investigations as opposed to secondary sources, and remaining knowledgeable in fields as diverse as psycho-social medicine, women's health, office-based procedures, and the cognitive skills of all the medical subspecialties is a tall order.

The teaching of general internal medicine requires an extraordinary investment in faculty time and informatics. Technology alone cannot substitute for clinician/scientist role models. Selection of residents must focus on candidates with the capacity to think critically and solve problems. It cannot focus only on those who can simply learn a trade, as valuable as the trade might be. Fostering an iterative approach to problem solving and diagnosis, gathering and interpreting information, and continually revising hypotheses (21) occurs best in small groups, all too often considered to be "inefficient." Faculty leaders of such groups must be genuinely dedicated to teaching. They must be given an adequate amount of protected time to teach and to enhance their teaching skills. If we want medicine to remain a profession, if we want to reject a reductionist, cookbook approach to training physicians, and if we want to inculcate students with expertise in critical thinking that typifies the finest clinicians, we must nurture and incorporate physician/scientists in our educational programs. We must empower doctors to do what the etymology of the word itself designates—namely, to teach.

1. Starr P. The Social Transformation of American Medicine. New York: Basic Books, Inc., 1982.
2. Ginzberg E. From Physician Shortage to Patient Shortage. Boulder, CO: Westview Press, 1986.
3. Johnson LB. Comments made in an address at the National Institutes of Health, Bethesda, MD, 1968.
4. Carroll L. Alice's Adventures in Wonderland, 1865.
5. Ebert RH. Flexner's model and the future of medical education. Acad Med **67**:737–742, 1992.
6. Dickinson-Chamberlin M, Dickinson W, Richards MD. Through a granddaughter's eyes. Coron Artery Dis **12**:79–82, 2001.
7. Sackett DK, Rosenberg, WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: What it is and what it isn't. Br Med J **312**:1–2, 1996.
8. Evidence-Based Medicine Working Group. A new approach to teaching the practice of medicine. J Am Med Assoc **268**(27):2420–2425, 1992.
9. Rahimtoola SH. Some unexpected lessons from large multicenter randomized clinical trials. Circulation **72**:449–455, 1985.
10. Minot GR, Murphy WP. Treatment of pernicious anemia by a special diet. J Am Med Assoc **87**:470–476, 1926.
11. Banting FG, Best CH. Pancreatic extracts. J Lab Clin Med **7**:464–472, 1922.
12. LeLorier J, Gregoire G, Benhaddad A, LaPierre J, Derderian F. Discrepancies between meta-analyses and subsequent large randomized, controlled trials. N Engl J Med **337**:536–542, 1997.
13. Sobel BE, Furberg CD. Surrogates, semantics, and sensible public policy. Circulation **95**:1661–1663, 1997.
14. Canto JG, Kiefe C, Williams O, Barron HV, Robers WJ. Comparison of outcomes research with clinical trials using preexisting data. Am J Cardiol **84**:923–927, 1999.
15. Bucher HC, Guyatt GH, Cook DJ, Holbrook A, McAlister FA. Users' Guides to the Medical Literature. XIX. Applying clinical trials results. A. How to use an article measuring the effect of an intervention on surrogate end points. J Am Med Assoc **282**:771–778, 1999.
16. Gotzsche PC, Liberati A, Torri V, Rossetti L. Beware of surrogate outcome measures. Int J Technol Assess Health Care **12**:238–246, 1996.
17. Epstein AE, Bigger JT Jr, Wyse DG, Romhilt DW, Reynolds-Haertle RA, Hallstrom AP. Events in the Cardiac Arrhythmia Suppression Trial (CAST): Mortality in the entire population enrolled. J Am Coll Cardiol **18**:14–19, 1991.
18. Robertson RP. Medical education without physician scientists: Answers without questions. Proc Soc Exp Biol Med **223**:228–229, 2000.
19. Hippocrates. Ancient Medicine: Epidemics. Jones WHS (translated by). London: William Heinemann, 1923.
20. Fitzgerald FT. On being a doctor. Ann Int Med **130**:70–72, 1999.
21. Kassirer JP. Teaching problem-solving: How are we doing? N Engl J Med **332**:1507–1509, 1995.