

Virtual Mentor

American Medical Association Journal of Ethics
January 2011, Volume 13, Number 1: 55-60.

HISTORY OF MEDICINE

The Origins of Evidence-Based Medicine—A Personal Perspective

David M. Eddy, MD, PhD

Medical decision making has gone through a fundamental change in the last 40 years. Simply put, the foundation for decision making has shifted away from subjective judgments and reliance on authorities toward a formal analysis of evidence. In the past, treatments were recommended if physicians believed patients would benefit. Now, before recommending a treatment, physicians ask: what's the evidence?

What caused this change? Evidence-based medicine (EBM) as we know it today is the confluence of several related but initially separate lines of research, and everyone who participated in the movement has his or her own story. Mine began when I decided to drop out of residency in cardiovascular surgery to get a doctorate in engineering mathematics. One day in the fall of 1974, I was asked to give a talk on how physicians make decisions. I chose diagnostic mammography as the example, because Betty Ford and Happy Rockefeller had just been diagnosed with breast cancer. I planned to write out the decision tree I presumed their physicians had used, fully expecting to find strong evidence, good numbers, and sound reasoning that I could describe to my audience. But to my amazement I found very few numbers, no formal rationale, and blatant errors in reasoning. How could that be?

Perhaps it was because mammography was relatively new. I decided to look at a treatment that had been in use for more than 75 years—that for ocular hypertension. Tens of millions of people were receiving it; surely there would be solid trials to support those decisions. But there weren't. There were only eight controlled trials, all very small and poorly designed. But perhaps most startling, six of the eight trials showed that patients got worse with treatment, not better. I then tried to conduct a formal decision analysis of another treatment. Experts I consulted quickly pushed me away, saying that there wasn't sufficient evidence or data to do such an analysis. That clinched it. If there wasn't sufficient information to develop a decision tree, what in the world were physicians basing their decisions on? I then realized that medical decision making was not built on a bedrock of evidence or formal analysis, but was standing on Jell-O.

I wrote up the work on mammography as an informal report and sent it to some colleagues in 1975. I also submitted a paper to *JAMA* arguing that many widely used treatments and tests were in fact not backed by good evidence or reasoning, using ocular hypertension as the example. I received several letters and calls from ophthalmologists expressing their fury. The editors of *JAMA* said they would publish

the expose of ocular hypertension treatment, but recommended over the telephone that I not include the more general points about poor evidence. Because my main purpose was to call attention to the need for better evidence, not to embarrass ophthalmologists, I pulled the paper. It was widely circulated and debated in the underground, but never published. Indeed, few editors were willing to publish exposes of poor evidence and physician uncertainty in the early years, although some did [1-4]. The evidence about treatment of ocular hypertension was eventually published, but in the less-threatening context of screening [5].

Influencing National Guidelines

From that starting point, I worked to introduce formal analyses of evidence on several fronts. The unpublished paper on mammography somehow worked its way to the National Cancer Institute (NCI), which was working with the Blue Cross Blue Shield Association (BCBSA) to develop coverage policies for cancer screening. I was asked by the NCI and BCBSA to apply the methods I had described in the mammography report to analyze other screening tests. For that I built a mathematical model to analyze the appropriate use and frequencies of various tests (1976-1978). I know of no earlier documented use of formal evidence and analytical methods to design coverage policies.

The American Cancer Society learned of that work and asked me to help them rewrite its guidelines for cancer screening. It took 2 years to do the work and get it approved by the society's board. Again, I believe this is the first application of formal methods, evidence, mathematical modeling, and cost-effectiveness analysis in designing a national guideline. As the primary author of the report, I inserted a preamble:

In making these recommendations, the Society has four main concerns: First, there must be good evidence that each test or procedure recommended is medically effective in reducing morbidity or mortality; second, the medical benefits must outweigh the risks; third, the cost of each test or procedure must be reasonable compared to its expected benefits; and finally, the recommended actions must be practical and feasible [6].

For me, these are the seeds of evidence-based medicine, although I didn't use those words in the report. The conclusions—such as 3-year Pap smears, no mammography screening for women under 50, no screening of smokers with chest x-rays—were very controversial, with wide coverage in the national media [7]. Because of the controversy, the society and I went on the road to defend the guidelines; over the next few years I gave well over 100 speeches promoting and defending the use of evidence and formal methods in medical decision making.

Rethinking Guidelines Altogether

The second arena in which I worked was formalization of the concept of guidelines. The discovery that decisions were not based on formal thinking about evidence and numbers raised the obvious question: What were they based on? Reading medical

textbooks and journals with that question in mind quickly led to the answer. Medicine was riddled with thousands of very simple “if-then” rules that are still used today. For example, “Treat to BP less than 140/90 mmHg or BP less than 130/80 mmHg in patients with diabetes or chronic kidney disease” [8]. The value of this type of simplification is clear: a physician doesn’t have to think through evidence and numbers and make difficult tradeoffs; all he or she has to do is learn the rules of thumb. The difference between today and 1980 is that today’s rules of thumb tend to be backed up by formal analyses [9] whereas three decades ago they simply appeared, with the unstated justification that they were “standard and accepted practice.”

I wrote an article for the *New England Journal of Medicine* to draw attention to what I then called “clinical policies” (guidelines, maxims, dicta, indications and contraindications, recommendations) and their implications for the quality of care [10]. That paper was picked up by the Council on Medical Specialty Societies (CMSS), which held a national conference to push the idea of improving methods for clinical policies and guidelines. The CMSS asked me to conduct workshops to train specialty societies in the methods, which began in the late 1980s and resulted in a manual of methods [11].

The ideas and methods spread, receiving a big boost in 1993 when the Agency for Healthcare Research and Quality (then called the Agency for Health Care Policy and Research [AHCPR]) began its guideline program, which was later converted to evidence-based practice centers (1997). The ideas were publicized in the U.K. in 1991, when Richard Smith, then editor of *BMJ*, invited me to give a speech in Manchester and wrote a controversial editorial about it [12]. They became firmly planted with the founding of the Cochrane Collaboration in 1993.

Guidelines for Insurance Coverage

The third approach to promoting evidence-based medicine was reworking coverage policies. In 1984, BCBSA asked my help in creating a formal evidence-based process for determining whether new technologies should be covered. I worked with them to develop five criteria that every new technology must meet. The second and third criteria were that “scientific evidence must permit conclusions concerning the effect of the technology on health outcomes” and “the technology must improve the net health outcome.” In addition to requiring evidence, these criteria introduced a crucial distinction between evidence for health outcomes (e.g., heart attacks) as opposed to physiological variables (e.g., blood pressure, cholesterol). These criteria are still in place and became the starting points for criteria developed subsequently by other organizations such as Medicare.

In the 20 years I served as the chief scientist of the Medical Coverage Advisory Committee to BCBSA (1984-2005), the most serious test of the criteria was use of high-dose chemotherapy and bone marrow transplant for breast cancer. Because of lack of evidence about health outcomes [13], insurance companies held the line against coverage for several years until many buckled under intense legal, political,

and media pressure. But the holdout enabled initiation of controlled trials that eventually showed the treatment to be ineffective.

Performance Measurement

The fourth domain was performance measurement, which was introduced by the National Committee for Quality Assurance in 1993. I was appointed to their Committee on Performance Measurement, and chaired the subcommittee on methods where we introduced criteria that performance measures must meet. The cornerstone was good evidence that improving performance on some care process or treatment target would actually improve health outcomes.

While all this was occurring, others were working on related ideas. Some of the most prominent endeavors were Archie Cochrane's book on the importance of controlled trials [14], Jack Wennberg's work on wide variations in practice patterns [15], Weinstein and Stason's paper on cost effectiveness [16], the founding of the Society for Medical Decision Making (1979), and studies from RAND showing high rates of inappropriate care [17]. The realization that fundamental changes were needed in medical decision making was occurring on many fronts.

“Evidence-Based” Enters the Lexicon

Although I had been using the term “evidence-based” in speeches and workshops at least since 1985, I first published it in a 1990 article about evidence-based guidelines [18]. This was part of a series of 28 articles I was writing for *JAMA* about guidelines, evidence, and costs [19]. The term then got a big push when David Sackett and his colleagues used it in the title of a paper they published in *JAMA* 2 years later [20]. As the subtitle of their paper implies (“A New Approach to Teaching the Practice of Medicine”), their focus was on medical education and individual physician decision making, not the design of guidelines, coverage policies, or performance measures. Furthermore, their definition of evidence-based medicine (“The practice of evidence-based medicine means integrating individual clinical expertise with the best available external clinical evidence from systematic research” [21]) gave considerably more power to subjective judgment (“clinical expertise”) than I was willing to accept in the criteria I had been developing for guidelines and other types of policies.

I felt it was important to make a distinction between the evidence-based policies I had been promoting and evidence-based individual decision making being developed by Sackett and his colleagues, and offered a unified definition [22]. Today, the great majority of the time, the term “evidence-based medicine” is used in the context of guidelines and other policies, not medical education or individual physician decision making. But despite these differences, there is no doubt that the paper by Sackett and colleagues helped propel the term into common usage.

Evidence-based medicine now has a life of its own, with scores of books, courses, programs, and even departments in medical schools dedicated to it. Current students and younger physicians now take for granted the requirement for evidence and

explicit formal analysis, barely knowing that this was not the case just a few decades ago.

References

1. Eddy DM. Probabilistic reasoning in clinical medicine: problems and opportunities. In: Kahneman D, Slovic P, Tversky A, eds. *Judgment under Uncertainty: Heuristics and Biases*. New York: Cambridge University Press; 1982: 249-267.
2. Eddy DM. Flying without instruments: analyzing medical policies by consensus and expert opinion. *John A. Hartford Foundation Bulletin*. Winter 1983.
3. Eddy DM. Variations in physician practice. the role of uncertainty. *Health Aff*. 1984;3(2):74-89.
4. Eddy DM. The quality of medical evidence: implications for quality of care. *Health Aff (Millwood)*. 1988;7(1):19-32.
5. Eddy DM, Sanders L, Eddy JF. The value of screening for glaucoma with tonometry. *Surv Ophthalmol*. 1983;28(3):194-205.
6. Eddy DM. ACS report on the cancer-related health checkup. *CA Cancer J Clin*. 1980;30(4):193-240.
7. Brody JE. Cancer society reports it finds some detection tests unneeded. *New York Times*. March 21, 1980:1.
8. Reference card from the Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure (JNC 7). <http://www.nhlbi.nih.gov/guidelines/hypertension/phycard.pdf>. Accessed November 20, 2010.
9. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure (JNC 7). <http://www.nhlbi.nih.gov/guidelines/hypertension/jnc7full.htm>. Accessed on November 20, 2010.
10. Eddy DM. Clinical policies and the quality of clinical practice. *N Engl J Med*. 1982;307(6):343-347.
11. Eddy DM. *A Manual for Assessing Health Practices and Designing Practice Policies*. Philadelphia, PA: American College of Physicians; 1992.
12. Smith R. Where is the wisdom? The quality of medical evidence. *BMJ*. 1991;303(6806):798-799.
13. Eddy DM. High-dose chemotherapy with autologous bone marrow transplantation for the treatment of metastatic breast cancer. *J Clin Oncol*. 1992;10(4):657-670.
14. Cochrane AL. *Effectiveness and Efficiency: Random Reflections on Health Services*. London: Nuffield Provincial Hospitals Trust; 1972.
15. Wennberg JE, Gittelsohn A. Small area variations in health care delivery. *Science*. 1973;182(4117):1102-1108.
16. Weinstein MC, Stason WB. Foundations of cost-effectiveness analysis for health and medical practices. *N Engl J Med*. 1977;296(13):716-721.

17. Chassin MR, Kosecoff J, Solomon DH, Brook RH. How coronary angiography is used: clinical determinants of appropriateness. *JAMA*. 1987;258(18):2543-2547.
18. Eddy DM. Practice policies: where do they come from? *JAMA*. 1990;263(9):1265, 1269, 1272, 1275.
19. Eddy DM. *Clinical Decision Making. From Theory to Practice*. Sudbury, MA: Jones and Bartlett Publishers; 1996.
20. Evidence-Based Medicine Working Group. Evidence-based medicine: a new approach to teaching the practice of medicine. *JAMA*. 1992;268(17):2420-2425.
21. Sackett DL, Rosenberg VM, Gray JA, Haynes RB, Richardson WS. Evidence-based medicine: what it is and what it isn't. *BMJ*. 1996;312(7023):71.
22. Eddy DM. Evidence-based medicine: a unified approach. *Health Aff (Millwood)*. 2005;24(1):9-17.

David M. Eddy, MD, PhD, is the founder and chief medical officer emeritus of Archimedes, Inc., a health care modeling company in San Francisco. His research interests include mathematical modeling and quantitative policy analysis.

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

Copyright 2011 American Medical Association. All rights reserved.