

EVIDENCE-BASED MEDICINE

The Promise and the Pitfalls

Keith Berndtson, MD

Park Ridge MultiMed
keith@parkridgemd.com
mobile (847) 404-4868

EVIDENCE-BASED MEDICINE

The Promise and the Pitfalls

Keith Berndtson, MD

All of the big players in the American health care sector are charmed by the conceptual approach to improving the quality of health care known as *evidence-based medicine* (EBM). Insurance executives in particular feel tingly all over at the mere use of the phrase. Why are investors, health reformers, insurance companies, drug companies, hospital systems, physician groups, think tanks, and the like so enamored with EBM?

For starters, no one can argue against it and sound credible. The minute you try you'll be hit with questions like, "Why do you hate America?" or "Why do you love disease?" Defend yourself against that.

As it turns out, the evidence in favor of EBM is weak at best. This essay reviews the promise and the pitfalls behind medicine's hot new catchphrase.

The Meaning of EBM

Evidence-based medical practice is defined as the integration of best research evidence with clinical expertise and patient values to guide medical decision-making. The theoretical goal of this movement is to reduce medical errors and practice variations. In the name of best practices, however, the EBM movement runs the risk of throwing out the good practice variations along with the bad. Even worse, it courts subtle, self-serving manipulations by cash flow-advantaged players in the sector.

The hallmark of the EBM movement is its unabashed bias toward the *randomized controlled trial*, or RCT. The RCT is a study design that minimizes bias by randomizing research subjects into experimental and control groups. When subjects have an equal chance of being assigned to receive an intervention (experimental subjects) or to go without intervention (the controls), study results stand a better chance of determining cause-effect relationships.

If a medical approach to a problem has little RCT evidence to go by, then it has little chance of measuring up. It doesn't seem to matter whether that approach to medicine relies on uncontrolled evidence of patient care outcomes even when those outcomes are based on a solid grasp of foundational physiological imbalances that have power to explain and treat a given patient's problems.

The Dark Side of RCTs

RCTs can and often do suffer from lethal forms of bias, but they always appear to produce the closest approximations to the truth. Despite this apparent advantage, RCTs possess many disadvantages. They are expensive, difficult to do well, and most troubling, their findings apply to strictly defined groups of research subjects, not to individuals, making it difficult for the doctor to translate RCT findings to the unique patient in the room.

RCTs are touted as gifts that make decisions easier for practicing physicians. RCTs are supposed to take the guesswork out of practicing medicine so physicians can confidently move patients through “the system” at ever increasing speeds.

To be fair, EBM has its share of science-based success stories. Yet on the whole, evidence-based medicine amounts to speeding up the patient conveyor belt with the help of data that often do not extrapolate well from the experimental group to the patient in need of explanations, let alone a thoughtfully tailored care plan.

The EBM movement ignores the age-old medical precept that people get sick in different ways—that they can arrive at the same diagnosis by different routes—and that they should therefore be treated in a way that takes into account their unique history, biochemistry, and genetics.

When setting up experimental and control groups, the RCT design team attempts to equalize their differences. The biologist does just the opposite: clarify and factor in what makes each case different from the next.

There's nothing wrong with wanting to base medical practice on good scientific evidence, but good medical practice always requires skill to tackle health problems that RCTs have yet to illuminate. Given the complexity of the living human being, that amounts to most chronic health problems.

The risk to society if EBM is swallowed hook, line, and sinker is this: if the EBM movement simply puts new clothing on the old concept of “name the disease, name the drug for the disease,” doctors may be forced by higher powers to enact a mechanized, force-fed approach to chronic illness that neglects the importance of biochemical individuality and therapeutic lifestyle change—concepts essential to a reality-based, biology-minded physician.

Insurers, for example, are charging ahead with plans to discount the complexity of the physician's task, relying on the RCT as the be all and end all on matters of clinical complexity while throwing things like experience and inductive logic out the window.

EBM is a branch that has grown from the valuable root concept of *informational literacy*, which refers to a doctor's ability to recognize when additional information is needed, and to understand how to locate it, evaluate it, and effectively put it into play when that is the case. Practice-based outcomes research is a different branch grown from the same root concept that is powerful in its own way yet neglected for reasons that defy health care logic.

Chronic illness is a complex puzzle for which guesses and snap judgments are no match. Yet "best practices" in medicine have a habit of obeying the conclusions of masterfully spun drug company-sponsored RCTs. Skillfully marketed RCTs routinely neglect the web-like complexity of biological reality—a fatal flaw that renders them prone to miss what they're not looking for.

True and Useful, True but Useless, or Just Plain False?

Based on work by Stanford's John Ioannidis, MD, PhD, we know that roughly twenty-five percent of published RCT findings miss the truth. Ioannidis mathematically proved why eighty percent of all published medical research findings are false.

Scientific medical studies come varnished in several coats of bias, but because journals compete to publish novel and surprising findings, they are prone to discount bias when they receive a study whose results look true enough to grab attention. Years go by before we learn that other groups couldn't replicate the original findings—meaning that the original study was false all along.

In his 2005 *PLoS* paper, *Why Most Published Research Findings are False*, Ioannidis shows that the most obvious driver of false research claims is bias that comes in many flavors, including biased study designs, measurement error, researcher bias, presentation bias, and publisher bias that famously includes high-impact journals, which are known for publishing novel findings that turn out to be false.

Ioannidis then deduced corollary principles about the probability that a research finding is indeed true. When he applies these corollary principles to medical research studies, he concludes that eighty percent of them have a low probability of being true.

There's more. In *A Different Universe: Reinventing Physics from the Bottom Down*, Nobel Prize-winning physicist Robert M. Laughlin described a similar dilemma in experimental physics. He argues that experimental studies in physics have accumulated logically and statistically tight scientific conclusions that sit in dormant piles because their findings are not right or wrong about anything that matters in the complex world of reality.

The infinitely granular human bodymind is composed of organized networks of systems that integrate and simultaneously operate through networks from the bottom up and the top down at the same time. RCTs amount to very rough approximations of such truths. In most cases RCTs do not apply to the physician's task of restoring health for people whose unique differences matter.

Every time trumpets announce the latest new medical discovery, think public relations finesse. Wait a few years to find that the "truth" changed and that the trumpets were actually warning you to beware of false prophecies. Most novel and significant research findings fail to get replicated over time. Time goes by and we discover that the big news that once stirred excitement fizzled and found its way into a dusty pile of stuff that seemed important but turned out not to matter.

Publication Bias

Novel findings that stir excitement are rewarded with publication because they help the publisher draw attention through headlines, sound bites, and page views. This is publication bias. Show me a journal that doesn't want to be the first to come out with a new advance in scientific thought. And don't forget, studies more likely to get cited by other authors will help the journal's rankings and, in turn, the publisher's advertising fee structure.

From a mathematical standpoint (think "bell curve"), less than a third of all findings are likely to be viewed as somewhat novel. Less than one in twenty are likely to be truly novel. It is therefore no surprise that most novel findings cannot be replicated. Is it any surprise that doctors are beginning to resist being led by the nose toward malodorous methods of medical practice, reminded by insurers along the way that if they want to get paid for their medical services, those services had better be "evidence-based."

If that's not enough, scientific studies that truly serve the public interest are growing scarce. Almost all scientific results now get spun one way or the other to sell you on a product or idea that serves interests other than yours.

In 1992, the Cochrane Centre opened in Oxford. By 1993, a vision for an international Cochrane Collaboration had emerged. The Cochrane Collaboration is an international network of over 28,000 people from over 100 countries.

Thus far, the Cochrane Collaboration has produced over 5,000 reviews of medical evidence. They hold the world's largest database on health-related randomized controlled trials, and they are widely hailed as the benchmark source for high quality information about health care in the world, yet they have no system for assessing publication bias.

EBM 2.0

McMaster University's Gordon Guyatt is credited with embedding the term *evidence-based medicine* in the minds of physicians and policy makers. McMaster's David Sackett is credited with stating the reigning brief definition of evidence-based medicine EBM: "Evidence-based medicine is the integration of best research evidence with clinical expertise and patient values."

The 1990s witnessed an explosion of research, commentary, and policy papers on EBM. Since the turn of the century EBM has become the frenzied preoccupation of the medical profession. But a funny thing happened on the way to the crowning of EBM as king: patients asked, "What about us? Do you really care about our opinions, or is the 'patient values' part just window dressing?"

Consumer-citizens are now asking the most penetrating questions of all: What do we do when there is scientific uncertainty? What is the role of values when evidence gaps exist? Whose values count when? Should it be the values of the expert panel (rarely specified), the values of the treating physician (the interests of the patient), or the values of the patient (shared decision making)? What should happen when expert panels contradict each other?

And so began EBM 2.0, a new stage in the evolution of the EBM concept, one in which patients became more vocal about the untenable positions they found themselves in whenever medical evidence failed to resolve uncertainties about their diagnosis or treatment.

When doctors were divided about which guidelines applied to a patient with an ill-defined condition, the experts who came up with reasons *not* to treat almost always won the day, much to the delight of health insurers and their allies.

The experts who came up with reasons *to* treat were often vilified for basing their thought process on flimsy supports like clinical experience, inductive logic, and "mere" practice data. To many doctors it was as if skills developed over years of study and patient care didn't amount to a hill of beans in the EBM kingdom, where RCTs rule.

Physicians who carry the responsibility of caring for patients know how difficult it is to translate published research findings to the patient in the room. That a patient shares a diagnosis with other patients doesn't make them clinically identical.

Physicians who treat patients toil in the real world of biochemical individuality, where each patient with a given diagnosis differs from other patients with the same diagnosis, and for reasons that can greatly influence the treatment outcome for each patient.

Physicians did not want to get on the wrong side of the EBM movement, but at the same time, many EBM practice guidelines felt like edicts air-dropped at their feet by planes too high to see. Rarely were the values behind these guidelines transparent. Many seemed designed to promote the interests of drug companies and health insurers.

Thought guidelines would conflict on issues of how to treat angina, diabetes, high cholesterol, Lyme disease, sore throat, and many other conditions, the best wielders of EBM rhetoric held the most invincible positions. In truth, EBM reality was messy. Medical evidence has long been easy to skew in ways that serve the interests of corporations and experts better than the interests of patients and their treating physicians.

Exhibit 1: The Lyme Wars

Dueling guidelines have marred the field of Lyme disease diagnosis and treatment for over a decade. The prestigious Infectious Disease Society of America (IDSA) has allowed its guidelines panel on Lyme disease to take a strong position against more sensitive diagnostic criteria for Lyme disease, and in the face of growing evidence to the contrary, it clings to the position that there is no such thing as Lyme spirochetes able to survive antibiotic challenge. As a result of these positions, physicians feel paralyzed, unable to make a diagnosis of acute Lyme disease when it is present, and content to believe that persistent Lyme disease is a figment of the patient's imagination.

The most recent guidelines published by the International Lyme and Associated Diseases Society (ILADS) deconstruct the most recent IDSA guidelines, finding that only 16% of their guidelines are supported by properly designed randomized controlled trials. They offer nearly 800 references to peer-reviewed publications to support their case for a more evidence-based approach to Lyme disease, including the use of the GRADE method of analyzing medical evidence and formulating recommendations put forth by Gordon Guyatt and colleagues.

The GRADE system is a new advance in EBM by virtue of its call to make transparent the values underlying any given set of recommendations. The new ILADS guidelines do this. As of this writing, IDSA guidelines have not made the values of their panel transparent to the public.

It is not unusual for two independent expert panels provided with identical research evidence on a medical condition to arrive at conflicting sets of guidelines. The Helsinki Declaration of the World Medical Association has outlined the principles that should apply to medical research and the clinical care of patients in the face of evidence gaps or expert panel disagreements. Namely, in the treatment of the sick person, the physician must be free to use a new diagnostic and therapeutic measure if, in his or her judgment, it offers hope of saving life, reestablishing health, or alleviating suffering.

Exhibit 2: The Detective from Pocomoke

Evidence-based medicine is a noble calling so long as the data we're using is true and useful to doctors as they practice their craft to help patients get well or cope. Replication of novel findings gives us more confidence that the findings are true, but that doesn't necessarily make them useful. In response to their research translation dilemma, physicians who assess and treat chronic health problems find that they need to rely more on the natural mechanics of good judgment and decision-making.

The process of moving from a narrow-angle to a wide-angle view involves *inductive reasoning*—drawing limited tentative conclusions about larger patterns from a filtered assortment of relevant facts. When the problem solver reasons from facts, such as reported symptoms and lab test markers, to how interactions among physiologic systems could produce those symptoms and markers, the clinical picture that emerges is better able than most published research to explain the systemic health imbalances that are affecting a particular person's health as a whole.

Ritchie Shoemaker, a solo family physician from coastal Maryland, is a leading example of how such thinking, combined with a disciplined approach to collecting and analyzing patient outcomes data, can produce a major advance in the diagnosis and treatment of a highly complex form of chronic illness. Shoemaker used inductive reasoning and systematic data collection to identify the condition that has come to be known as chronic inflammatory response syndrome, or CIRS.

Most instances of CIRS are caused by poor clearance of mold toxins from the body. Shoemaker identified the genes that create this susceptibility and determined that such genes are present in roughly 1 in 4 people. Over fourteen years he described in slowly evolving detail the neuroimmune, vascular, and endocrine imbalances that occur in such patients.

Based on his patients' responses to cholestyramine—a medication that helps bind and remove various charged particles from the body—he induced that it could be used to treat patients with biotoxin illness. His interplay of inductive reasoning and practice data analysis produced a thorough description of the causes, diagnosis, and treatment of a previously undiscovered chronic illness that affects up to twenty-five percent of the population—not bad for a rural family doctor.

The medical mainstream maintains an uneasy relationship with his data for reasons that speak loudly about the cultural, political, and economic currents of science, whose strong undertow deters the adoption of new ideas on grounds of "insufficient evidence," a position that usually translates into, "Come back when you have a large RCT with at least two independent replications." That EBM trump card means, "See you twenty years." Meanwhile, patients suffer from a paradigm bias that serves to protect those whose status is threatened by the new paradigm.

Despite Shoemaker's discovery and thorough delineation of an important medical condition, *non-mainstream* physicians also have an uneasy relationship with CIRS. Divergent medical thinkers are used to treating medically unexplained or non-responsive conditions with methods based on educated guesses.

Shoemaker's published papers and ongoing practice-based outcomes research challenge both extremes in the EBM debate: the "nothing's valid until it's proven to be valid" mainstreamers, and the "anything's valid until it is proven to be invalid" alternative medicine crowd.

It's a broken application of *deductive reasoning* to apply the results of a randomized, controlled trial to a chronically ill individual whose illness is unique in ways not addressed by the RCT. It's as if the act of translating an RCT finding to the patient will, by itself, suffice to determine what treatments should or should not be advised. This assumes that results generated by a group of people should strictly apply to every individual in the same kind of group. Any clinician worth her salt knows that's horse hockey.

Thought Leaders, Expert Panels, and Forked Tongues

In the January 12, 2011 issue of the *Journal of the American Medical Association*, Robert H. Brook, MD, ScD, refers to evidence-based medicine as the "slogan of the day." Dr. Brook is Professor of Health Services and Medicine at UCLA's School of Public Health and a member of the Institute of Medicine. He directs the Robert Wood Johnson Clinical Scholars Program, and is also a Corporate Fellow of the RAND institute.

Dr. Brook has long had a special interest in the concept of health care quality. How should it be defined and measured? Are current measures adequate? How do consumers go about trying to find and access it? How are they to separate marketing rhetoric from real differences in quality that would reliably improve health outcomes and reduce costs? Is the evidence-based medicine movement helping consumers make better decisions or is it helping insurers to control cash flow and manipulate consumer demand?

"In the 1970s," wrote Brook, "for those with health insurance, no matter what kind of service a physician and patient agreed to, the service was covered without either evidence that it was effective or prior authorization." How times have changed. "Forty years later," he notes, "the Affordable Care Act requires that the federal government define an essential benefit package for those individuals who will obtain insurance through the new health insurance exchanges." Then, his wonderfully packed question: What is an essential benefit package, and who defines it, using what evidence?

Every day, health insurance executives, lobbyists, policy wonks, and politicos haggle over mutually agreeable answers to such questions and they do so behind closed doors. Missing from the equation are the patients and physicians who'll be forced to live with the decisions of the most powerful special interests in the health care sector. To make sure those special interests are properly served, the term "evidence" will be spun into policy argumenst that have the appearance of invincible logic, backed by well-tailored statistics that coincidentally maintain the money and power advantages of the sector's most elite stakeholders.

In his book, *stats.con: How We've Been Fooled by Statistics-Based Research in Medicine*, physician and writer James Penston, MD, argues that the claims of statistics-based research in medicine are, for the most part, imaginary, that the apparent success is based on the presentation skills of its advocates, and that the studies are devoid of anything of real value. On the presumed value of RCTs, he writes:

The components of the RCT—including randomisation, allocation concealment, double-blind administration of treatment, the handling of withdrawals and drop-outs, and the statistical tests—don't guarantee that the conditions for internal validity have been satisfied [and] the inference from a small difference in outcome to the presence of a causal relationship is highly questionable...The reliability of any generalisation from the results of an individual study to the wider population of patients—that is, the external validity—is always open to question. We can never know whether the results of a RCT apply to either a particular patient or to a specified group.

Doctors who practice the healing craft are always somewhat skeptical about translating research data, thought leader opinions, and expert panel recommendations to the patient in the room, because what was true for a narrowly defined study group rarely matches the uniquely complex life of an actual person. Physician healers strive to give the multiple dimensions of the patient's situation due consideration. Specialists aren't paid to practice multidimensional medicine. Primary care physicians are primed to practice multidimensional medicine, but their overlords won't afford them the time to do so. This is why physician healers are an endangered species.

Evidence-Based Survival of the Species

For any doctor aware of the clinical research translation problem, the challenge is to separate the wheat from the chaff when it comes to medical research with questions like, Is it valid? Is it valid internally and externally? Is it reliable? What's the likelihood of bias in any form? Are the findings meaningful? If so, to whom are they most useful? Patients? Physicians? Corporations? Statistical gymnasts? Thought leaders for hire?

The life sciences are so complex that most research is too non-definitive to support the policies that insurers use to fight claims and harass doctors. The way to benefit from medical science over the long term is to temper it with clinical experience about what works based neither on fallible evidence concerning randomized groups, nor on the fallible judgments of physicians who shoot from the hip, but on unconstrained human wisdom, experience, and judgment supported by practice data.

The pop culture of EBM could be laying a trap for our society. By crowning the RCT as king of medical science, we vastly oversimplify the complexity of the physician's cognitive task. If we discount the experience, inductive powers, and multidimensional wisdom of seasoned clinicians in the trenches, we run the risk of dehumanizing medicine altogether, giving big data crunchers free reign to give predetermined answers the appearance of good science.

We all know that science is not a purely intellectual enterprise. It is contaminated by a host of conflicting interests. Only fools trust science to pump out nothing but pristine analyses of the objective world around or within us. The wise find it safe to assume that many, if not most academic, business, and political institutions will use their power and influence to manipulate interpretations of scientific data to serve their own purposes.

There are many things we need to fix in our dysfunctional health care sector. A key need is to extend concern beyond the health care system itself to the health and wellness of living systems as a whole. Which corporations will weave that goal into their mission statements? Where are the leaders willing to champion the health needs of individuals?

If we look down upon the earth from space shuttle altitude, we realize that behind that mask of incredible beauty simmers a cauldron of man-made trouble that is threatening the health of many species, including our own. Why is the medical profession looking so deformed? Whatever happened to science in the public interest? Where are the leaders to champion the kinds of change needed to support the health of living systems as a whole? Is humanity even capable of developing a more sustainable relationship to the world? What does the best evidence have to say on the subject?