

11–16

Evidence-based medicine (EBM): origins and modern application to spine care

Authors Andrea C Skelly¹, Jens Chapman²
Institutions ¹ Spectrum Research Inc, Tacoma, WA, USA
² Harborview Medical Center, Seattle, WA, USA



Since the words *evidence-based* are so pervasively used in our present day practices of medicine in general, and spine specifically, we thought it might be of interest to explore the roots of this movement in greater detail. It is probably always a good idea to look for the origins of medical developments in ancient writings on scientific methods from around the world, particularly Greece and ancient Chinese medicine.

In more modern times, it has been 20 years or more since some of the first references to and basic descriptions of evidence-based medicine (EBM) started appearing in the medical literature [1–3]. Since then, the concept of EBM has gone through various interpretations and applications, myths and misconceptions, uses and abuses. Tenets of EBM have seeped in to most aspects of patient care in a variety of forms and have impacted areas from clinical practice guidelines to reimbursement and health policy. The precepts of EBM and their application continue to evolve.

Prior to the 1950s, health care decisions were based primarily on anecdotal information, pathophysiology, and the expert opinions of leaders in the profession. In a 1992 article in the *Journal of the American Medical Association (JAMA)*, the author pointed to a “new paradigm for medical practice that is emerging,” and defined this paradigm as the shift away from “intuition, unsystematic clinical experience, and pathophysiologic rationale as sufficient grounds for clinical decision making” [2]. This opinion-based article introduced the concept of examining evidence from clinical research and applying the findings to clinical practice. Important to the introduction of EBM in clinical practice was that EBM “requires new skills of the physician, including efficient literature searching and the application of formal rules of evidence evaluating the clinical literature.” Some of the earliest arguments in the 1990s for an “evidence-based” approach to medical practice were in the context of systematic development of clinical guidelines [3].

Modern “founders” of EBM, include Archie Cochrane, David Sackett and Gordon Guyatt, and David Eddy among others. Each has contributed to unique areas of what we currently perceive as EBM.

What EBM is and is not

EBM is not intended to replace clinical experience and judgment, nor is it “cookbook” medicine. Rather, it is a “set of principles and methods intended to ensure that to the greatest extent possible, medical decisions, guidelines, and other types of policies are based on and consistent with good evidence of effectiveness and benefit” [3]. It considers the validity of, and gaps in, research. It facilitates *correct, informed* interpretation of the literature and how to best apply it in a given clinical situation and in discussing options for care with patients. EBM has the potential to enhance the overall quality of care by providing clinicians with the “best” current evidence to support decision making without replacing the need for crucial consideration of individual specific patient presentation, available resources as well as expertise and combining it with the evidence to formulate a clinical judgment on how to treat a given patient.

Evolution of study design—some perspectives

Most early medical literature consisted of case reports and case series in addition to description of techniques and information from non-human studies. Over the past several decades, much work has been done on developing study methods in epidemiology and health services research to provide a more rigorous framework for study design to reduce bias. Sophisticated statistical techniques for evaluating and combining studies have advanced the ability to estimate study effects while controlling for confounding and accounting for a variety of study factors.

Historically speaking, the randomized control trial (RTC) dates back to 600 BC when Daniel of Judah compared the health effects of the vegetarian diet with those of a royal Babylonian diet over a 10-day period. The trial had obvious deficiencies by contemporary medical standards (allocation bias, ascertainment bias, and confounding by divine intervention), but the report has remained influential for more than two millennia [4]. Credit for the modern RTC is usually given to Sir Austin Bradford Hill for his clinical trials on streptomycin for pulmonary tuberculosis and is rightly regarded as the landmark study that ushered in a new era of medicine. Since Hill’s pioneering achievement, the methodology of the RCT has been increasingly accepted and the number of RCTs reported has grown exponentially [5].

All study designs have limitations, even RCTs. No one study (or study design) can definitively address all aspects (efficacy, effectiveness, safety) of a topic. RCTs, if done well, have the greatest potential for providing the highest-quality evidence regarding efficacy (by reducing bias and confounding). However, RTCs may not have the statistical power to detect differences in rare adverse events and may not have sufficient length of follow-up to describe long-term benefits and complications. (One only need look at the recent well-publicized recall of certain pharmaceuticals to see that RCTs do not answer all questions.) This is where well-designed and reported nonrandomized comparative studies (cohort and case control) and even case series can help fill gaps in understanding. Nonrandomized studies also facilitate the development of hypotheses and next steps for adding to the overall evidence base on a topic. While all study designs have some role in contributing to the evidence on topic, the tenets of EBM suggest that the focus be on the highest quality of evidence, not necessarily the whole laundry list of studies that have been done on a given topic.

The development of systematic reviews has been crucial to EBM as they focus on research questions that synthesize all high-quality research evidence relevant on the clinical questions. In the early 1990s, Archie Cochrane, a British physician, began advocating the use of systematic reviews in medicine. Cochrane is best known for his influential book, *Effectiveness and Efficiency: Random Reflections on Health Services* published in 1972. He suggested that because resources would always be limited they should be used to provide equitably those forms of health care that had been shown in properly designed evaluations to be effective. In particular, he stressed the importance of using evidence from RCTs because these were likely to provide much more reliable information than other sources of evidence [6]. Cochrane's simple propositions were soon widely recognized as seminally important—by lay people as well as by health professionals and policymakers. The Cochrane Collaboration is named in his honor.

Well-designed and methodologically sound systematic reviews and meta-analyses have become a mainstay in EBM. They have become the basis for clinical guideline formation, health technology assessment and comparative effectiveness reviews, all of which are used increasingly for health policy and reimbursement decisions.

Critical appraisal: the foundation of EBM

A cornerstone of EBM is the critical appraisal of clinical research. As mentioned above, EBM facilitates *correct, informed* interpretation of the literature. The critical appraisal provides important context around the findings and potential biases of a study (or studies) so as to help one put the results of a study in perspective.

During the 1980s, Canadian physician, David Sackett, began to develop a system for the critical analysis of the medical literature. Sackett organized the principles of the systematic review into six steps: (1) decide what information is needed; (2) formulate one's information needs in the form of a question that a research study could answer; (3) search the published literature to find the evidence; (4) decide which studies are valid and applicable to the patient at hand; (5) apply the findings to the patient; and (6) evaluate the outcomes [7]. This approach is mirrored in a series of articles describing "how to" use various types of research articles to answer clinical questions about diagnosis, prognosis, and treatment to make clinical judgments for a given patient based on the best evidence.

Many systems for the critical appraisal of clinical studies have been described in the literature [8, 9]. Overall, the intent is to evaluate the potential sources of bias in studies, allowing the astute reader to assess the extent to which such biases may influence the results. In general, all schemes assess components of studies that may bias studies such as the following:

- Study design (eg, prospective cohort, RCT)
- Patient selection and evaluation methods (including outcomes evaluation)
- Patient follow-up (length and completeness)
- Sample size and ability to detect differences beyond the role of chance
- Consideration of and controlling for potentially confounding factors

All these are important when considering the extent to which the results of a study are valid and believable. EBSJ's format for critical appraisal takes into account the primary sources of potential bias and is described at the back of each issue. For each of the clinical research studies published, a critical appraisal is performed.

Within EBSJ, readers will get additional detail about the importance of these and other factors in the "Science in Spine" section which is a regular feature.

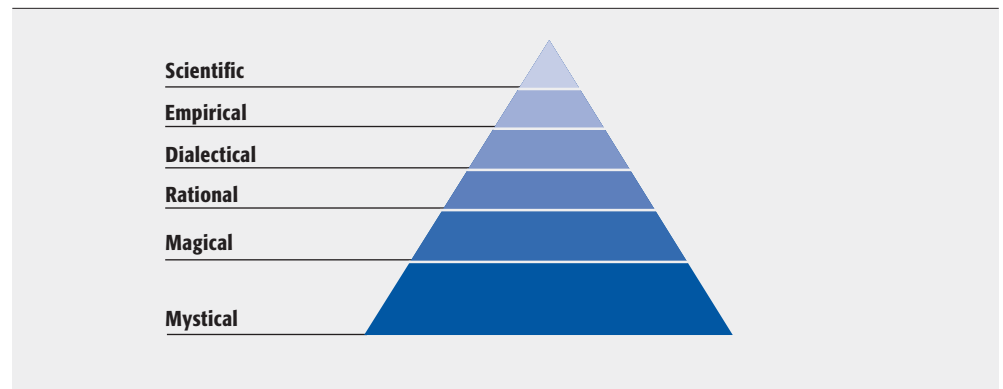
As described below, the basics of critical appraisal form the foundation for determining the "best" evidence.

Hierarchy of evidence

Evidence is the knowledge on which we base our beliefs (and disbeliefs) and in many instances our actions. It stands to reason that we would want to base our actions on the “best evidence.” So how do we know what type of “evidence” gets us closest to the “truth” and what quality of evidence do we base our actions?

René Descartes (1596–1650), philosopher, scientist, and mathematician, provided an important basis for the scientific method. Descartes recommended that, “When it is not in our power to follow what is true, we ought to follow what is most probable.” What is most probable is that the “best evidence” on a given spine-care topic may not be the highest quality and its proximity to the “truth” may not be clear. Yet, it is the “best” we have and we need to combine it with our clinical judgment and act accordingly. Descartes even described a hierarchical pyramid of evidence and process for moving toward the top (**Fig 1**). So, at least as far back as the early 1600, the concept of aspiring to consider the highest quality of evidence existed.

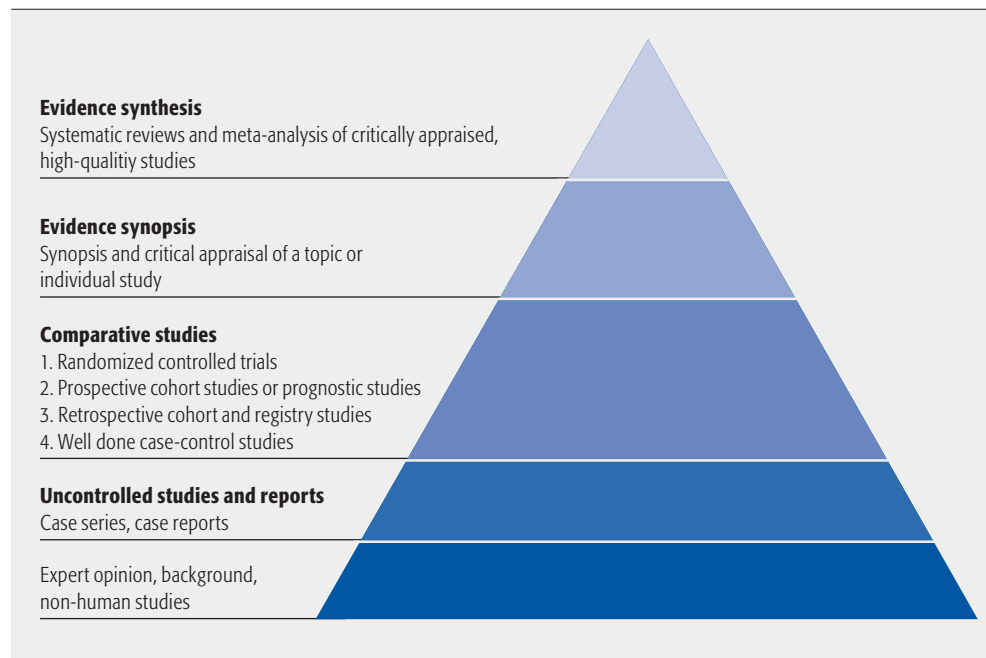
Fig 1 Descartes' hierarchical pyramid of evidence.



Readers are most likely familiar with some version of the more modern “evidence pyramid” (**Fig 2**). This too has evolved and there are a number of different versions. Study design plays a key role in all. The intent is to focus on studies with the least potential for bias as being the highest quality and, where possible, synthesis of high-quality studies on a given topic. The highest-quality studies are not necessarily RCTs, as suggested earlier. Where there is “higher-quality” evidence for efficacy, such as RCTs or methodologically rigorous prospective studies comparing treatment, the usefulness of case series for efficacy is questionable. This was part of Cochrane’s contention. Lower-quality comparative studies (eg, retrospective cohort studies which do not control for confounding) may have conflicting results versus the RCTs that may be attributable more to bias than to at-true effect. So an evidence-based synthesis would logically include only the highest-quality studies. For surgical specialties, the applicability of RCTs in helping answer important clinical questions has increasingly been called into question as even in major, well-designed RCTs lack of adherence to assigned treatment groups may be a problem. For example, in the Spine Patient Outcomes Research Trial (SPORT), comparing nonoperative and operative treatments for lumbar degenerative spondylolisthesis, 54% of those assigned to nonsurgical treatment crossed over to surgical options by 4 years; while 66% of those assigned to surgery received it by 4 years, resulting in no statistical difference between treatments when intention-to-treat analysis was performed. In the corresponding observational cohort, 97% of those who chose surgery received it and 33% of those who chose nonoperative treatment eventually underwent surgery [10].

The likely cause for this phenomenon lies in the increasing emancipation of patients as part of the decision-making process in an era of increasing availability of information. This quest for increased ‘shared-decision’ making, however, does not call into question the validity of the evidence pyramid per se, rather it underscores the critical role of well-conceived and thought-out studies for the future.

Fig 2 Hierarchy of evidence: one modern perspective.



Strength of evidence—evolving application of EBM

In the past 5 years in particular, methods for describing the overall “strength of evidence” for a topic have continued to evolve. The intention is to assess the overall body of literature on a given topic by the quality of the studies (based on critical appraisal components above), the quantity of studies (which takes into account the comparisons and outcomes of primary importance with a focus on higher-quality studies) and consistency (including constancy, magnitude, and direction of effect size). This is used to determine how likely new studies are to change the confidence in the effects estimated in the included studies.

All the systematic reviews included in EBSJ are assessed with respect to the overall strength of evidence based on the GRADE system [11]. The overall Strength of Evidence (SoE) ranges from high for a body of evidence if new studies are unlikely to change the effect estimates to very low if estimates from the currently available body of evidence is very likely to change as new data become available. This system and its variations are being used more frequently in clinical guideline development and health policy. Extensive discussion of this application of EBM is beyond the scope of this article, and the reader should consult the references.

Why is the evolution of EBM important to clinicians and clinical researchers?

- The evolution of the EBM has important implications for the practice of medicine. It is a valuable component of clinical decision making and facilitates discussions with patients.
- An important implication relates to the use of EBM components as the foundation of health policy and reimbursement.
- For researchers, understanding aspects of critical appraisal and the hierarchy of evidence can help facilitate designing, executing, and reporting higher-quality studies.

Conclusion

Evidence-based medicine is here to stay. Its various forms and applications will increasingly impact medical practice, patient care, and health policy.

References

1. **Bhandari M, Zlowodzki M, Cole PA** (2004) From eminence-based practice to evidence-based practice: a paradigm shift. *Minn Med*; 87(4):51–54.
2. **Rosenberg WM** (1992) Evidence-based medicine: a new approach to teaching the practice of medicine. *JAMA*; 268:2420–2425.
3. **Eddy DM** (2005) Evidence-based medicine: a unified approach. *Health Aff (Millwood)*; 24:9–17.
4. **Jadad A** (1998) *Randomised Controlled Trials: A User's Guide*. London: BMJ Books.
5. **Stolberg HO, Norman G, Trop I** (2004) Randomized controlled trials. *AJR Am J Roentgenol*; 183(6):1539–1544.
6. **The Cochran Collaboration** (2010) Archie Cochrane: the name behind The Cochrane Collaboration. Available at: www.cochrane.org/about-us/history/archie.cochrane.
7. **Sackett DL, Rosenberg WM, Gray JA, et al** (1996) Evidence based medicine: what it is and what it isn't. *BMJ*; 312(7023):71–72.
8. Oxford Centre for Evidence-based Medicine Levels of Evidence (2009) Available at: www.cebm.net/?o=1025.
9. **West S, King V, Carey TS, et al** (2002) Systems to rate the strength of scientific evidence. *Evid Rep Technol Assess*; 1–11.
10. **Weinstein JN, Lurie JD, Tosteson TD, et al** (2009) Surgical compared with nonoperative treatment for lumbar degenerative spondylolisthesis: four-year results in the Spine Patient Outcomes Research Trial (SPORT) randomized and observational cohorts. *J Bone Joint Surg Am*; 91(6):1295–1304.
11. **Atkins D, Best D, Briss PA, et al** (2004) Grading quality of evidence and strength of recommendations. *BMJ*; 328(7454):1490.