Evidence-based Dentistry

David W. Chambers, EdM, MBA, PhD, FACD

Abstract

Both panegyric and criticism of evidence-based dentistry tend to be clumsy because the concept is poorly defined. This analysis identifies several contributions to the profession that have been made under the EBD banner. Although the concept of clinicians integrating clinical epidemiology, the wisdom of their practices, and patients’ values is powerful, its implementation has been distorted by a too heavy emphasis of computerized searches for research findings that meet the standards of academics. Although EBD advocates enjoy sharing anecdotal accounts of mistakes others have made, faulting others is not proof that one’s own position is correct. There is no systematic, high-quality evidence that EBD is effective. The metaphor of a three-legged stool (evidence, experience, values, and integration) is used as an organizing principle. “Best evidence” has become a preoccupation among EBD enthusiasts. That overlong but thinly developed leg of the stool is critiqued from the perspectives of the criteria for evidence, the difference between internal and external validity, the relationship between evidence and decision making, the ambiguous meaning of “best,” and the role of reasonable doubt. The strongest leg of the stool is clinical experience. Although bias exists in all observations (including searches for evidence), there are simple procedures that can be employed in practice to increase useful and objective evidence there, and there are dangers in delegating policy regarding allowable treatments to external groups. Patient and practitioner values are the shortest leg of the stool. As they are so little recognized, their integration in EBD is problematic and ethical tensions exist where paternalism privileges science over patient’s self-determined best interests. Four potential approaches to integration are suggested, recognizing that there is virtually no literature on how the “seat” of the three-legged stool works or should work. It is likely that most dentists choose to wait for collective professional standards to reveal acceptable practice or follow a strategy of punctuated equilibrium, only switching out established practice habits when very conspicuous advantages are identified. Integration in medicine appears to follow the statistically sophisticated practice of updating estimates of clinical parameters (probabilities) for diagnoses, treatments, prognoses, and side-effects. This approach is likely beyond the skill or interest of clinical dentists and it fails to incorporate values in the integration. The use of decision trees to integrate both research and experiential parameters and values is illustrated and it is shown that such a technique identifies why there are very few cases in dentistry where evidence needs to be consulted and indicates what such cases are.

The renaissance thinker Buckminster Fuller suggested that in order to change the way people function, it is better to give them a tool than an argument. He must have been thinking particularly of dentists. Evidence-based dentistry certainly qualifies as one of the hot tools in the profession today. But it is well-known that large numbers of dentists have garages filled with materials and devices that were once the “must have” tools of their day.

Evidence-based dental practice (hereafter EBD) is a three-legged stool. Straus, Richardson, Glasziou, and Haynes (the successors to EBM founder, David Sackett) define the concept as “the integration of the best research evidence with our clinical expertise and our patient’s unique values and circumstances” (Straus et al, 2005). This is essentially the definition adopted by the American Dental Association. One of the legs of EBD is practitioners’ professional judgment derived from experience. Another, usually described as patients’ values, is the presenting condition of the patient, including the patients’ personal preferences regarding their treatment. The third leg is high-quality clinical epidemiology that permits justifiable estimates of clinical parameters such as risk factors and success rates associated with various interventions. These three legs support a seat, the place where dentists take a position, that consists of the prudent, rational, and judicious integration of information derived from weighing and balancing information from all three legs. EBD is decision mak-
ing done by practitioners; it is not research done by academics.

If a dentist accessed a systematic review or meta-analysis from the literature showing that treatment W produced incontestably superior results in numerous, large randomized controlled trials, and preceded based solely on this information to perform procedure W, he or she would not be engaging in EBD (clinical experience, patient preferences, and integration have been omitted). If another practitioner proceeded with treatment X because it had always worked “in his hands” in the past and the patient strongly favored this general approach, the dentist would probably not be recognized as practicing EBD. If dentists reviewed the literature and discovered that performance characteristics of material Y are consistent with what is known generally about physics, materials science, biochemistry, and human anatomy, they would still not qualify as evidence-based practitioners (although we might praise them for being “science-based”). If the same dentist who labored to integrate clinical experience and patient factors conducted a literature search and discovered that there were no good studies proving the efficacy of Z or that there were a few studies and they were slightly inconsistent, it would be entirely a matter of definition whether the dentist is practicing EBD. If the dentist had at hand a rich repertoire of clinical experience, was oriented toward honoring the patient, and found definitive, high-quality clinical research, he or she may still fall short of evidence-based practice. It would be necessary to calculate number needed to treat values (NNT), patient value-adjusted likelihood of help and harm (P*LHH), unique patient expected event rates (PEER), and then use Bayesian logic or some other form of decision science for integration. If a researcher or group of subject matter experts in a consensus conference setting declare the evidence overwhelming in favor of treatment R on average, they are not doing EBD. Only practitioners do it, and they only do it in the context of treating individual patients. Certainly, unreflective adherence to clinical guidelines generated by others is not EBD. Nor is EBD to be equated with the approach to improving performance known as best practices or benchmarking. Importing what has been proven to work does not count because the information comes from other, similar practices and not from the research literature.

Only the integration by a dentist in a unique practice setting of the relevant clinical epidemiology literature, professional judgment from experience, and the patients’ circumstances and values counts as EBD.

EBD enthusiasts generally offer something like the three-legged stool definition and then focus their energies on summarizing what high-quality clinical epidemiology exists, making it available to practitioners, and propagandizing that what they have done is the complete package for improving dental practice. The Hackshaw, Paul, and Davenport (2006) book Evidence-based Dentistry is a standard statistics textbook with an introductory chapter on the philosophy of EBD and some standards for literature searches. Enthusiasts are fond of rehearsing statistics about the rapidly explosion of the dental literature. The volume of science exceeded practitioners’ practical capacity many years before the Internet became a tool EBD experts could demonstrate to dentists.

The three-legged stool of EBD is not currently a comfortable sit. Far and away the best developed leg, and the one capable of bearing the greatest weight, is the dentist’s professional judgment. By contrast, incorporation of patient values into treatment decisions seems to be rather stubby, and we would gain much
by learning to engage patients more fully in their care. The availability of “best evidence,” a comprehensive body of high-quality clinical trials showing a consistent superiority of one or another approach to treatment, is more a promise than an actuality in dentistry. Perhaps the stability of the system could be improved if we add a general understanding of science and the practitioner’s own circumstances and values as the fourth and fifth legs.

The situation that makes EBD so discomforting is the complete lack of a seat, an integration of the support provided by three types of evidence. There is no consistent theory for how to integrate diverse sources of information in practice. (There has been extensive work on criteria for ruling out of court research findings that do not meet standards for experimental rigor: but that is another matter.) My own paper in the Journal of the American Dental Association (Chambers et al, 2010) is the only study I have been able to locate that explores how dental practitioners actually integrate information in deciding on care. My fellow researchers and I found that dentists place unwarranted emphasis on research data, that such inaccurate use of evidence decreases as practitioners gain greater clinical experience, and that practitioners have a natural suspicion of procedures that limit their professional judgment. There is virtually no evidence that the adoption of evidence-based medicine improves the quality of medical care (Straus et al, 2005). At best, there are anecdotal stories about once held but now discarded theories (without appreciation that today’s evidence may be the source of snickers in future generations). It is not uncommon to find EBD researchers advancing what “works in their hands.” There is a small body of research intended to identify the factors that retard the introduction of evidence-based practice (Cabana, 1999; Houser & Oman, 2011; Tracey, 2005; Webster, 2005). Generally the evidence suggests that practitioners make trade-offs between external data sensitivity to the clinical context and their own judgments.

Evidence-based approaches began in medicine in the 1980s, and have been imitated to some extent in dentistry. It has not been popular in the discipline of education, where several professional groups have formal policies denying its usefulness (Educational Researcher, 2008). In her 2005 presidential address to the Academy of Management, Denise Rousseau (2006) proposed that the field of academic management might profitably follow medicine’s lead. After a few rebuffs (see Ashkanasy, 2007), the subject has been largely dormant. One can only speculate as to the reasons evidence-based language would be congenial to some in medicine and a few in dentistry, but not elsewhere. Just possibly, it is the money. Almost all of medicine and much of dentistry is driven by procedures and materials or drugs in which there is a tremendous commercial stake. That is not the case in disciplines that have been slower to respond to the potential of empirical justification for treatments that can be purchased. (See www.provationordersets.com/common/flash/pvosdemo/demo.html for a commercial about “evidence vendors”.) Alternatively, medicine is a practice where the contribution of technology is larger relative to the contribution of operator variance (professional experience). The balance between operator and technology is different in dentistry, and strongly in favor of the “operator” in educational and business.

The Evidence Leg

It is certain that the recent increased attention to EBD has been a boon to the profession. Multisite practice collaborations are producing answers...
to clinical questions. The emphasis on methodological rigor in research design has raised the quality of clinical science. New money and recruiting of talented researchers to address clinical questions is probably happening. Publication of systematic reviews showing lack of sound studies or inconclusive results has called attention to areas where research is needed. A new group of academics skilled in accumulating and summarizing clinical epidemiology has been formed. Internet tools are more readily available to practitioners to guide focused, quick searches for information. There is renewed suspicion regarding the claims of commercial interests, both corporate and private—although everyone who has a solution to sell now seems to be tacking “evidence-based” onto their claims.

These are blessings to dentistry, but they are not the whole story. Recently we have stretched the evidence leg of the stool, but it remains to be seen whether the sit is any more secure. We certainly would not want to mistake the brightness of this tool for the work it does or does not actually accomplish.

What Makes Evidence Good?
Notice that EBD is not about the scientific base for practice or even about research. It is not about understanding why this or that approach works or fails to produce any worthwhile effect. EBD is about evidence: the factual justification for choice (Sivia, 2006). What is to count for one contemplated action and against others? If research is just objectively true but fails to guide behavior, it is not evidence. But there is no way to determine whether a set of numbers is evidence without knowing its intended use. Unlike a datum that simply is, evidence lives or dies depending on what we intend to do with it.

In a word: there is no set of research method or experimental design criteria that allow us to say whether this or that dataset constitutes good evidence, or even “best evidence.” This is Lee Cronbach’s classic observation (Cronbach & Meehl, 1955), later codified by the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education (1985), that the validity of evaluation data is in the decision made using the evidence, not in the evidence itself. [See also Educational Researcher; 2007, entire issue.] Does the leg fit the stool: not does the stool fit the leg (Messick, 1994)?

Certainly, every observation is not good evidence. It must pass “admissions standards.” Confessions and physical evidence obtained at crime scenes are rules out of court in legal systems unless obtained via recognized procedures. Hearsay is not evidence in court or in dentistry. When diverse actions are contemplated, it is only prudent to look to the quality of the evidence, or even the existence of any evidence, that supports alternative choices. The pyramid of evidence developed for grading evidence is helpful because it draws our attention to the fact that, other things being equal, certain types of evidence (meta-analyses and RCTs) are stronger than others (one-off correlational studies and case reports). Methodological rigor is a necessary condition for good evidence, but it is not a sufficient one. Although it is fun to jab at the gaffs of the past, proving others wrong or not useful does not prove us right or of value.

There are innocent individuals in our prisons who were convicted based on good evidence. Evidence does not constitute proof. It supports predictions that may or may not prove to be sound guides to action.

Infernal and External Validity
Straus, Richardson, Glasziou, and Haynes (2005) propose a two-step process for considering evidence. First, is the evidence internally valid (in the sense of meeting criteria for methodological rigor); second, is it applicable to our clinical situation? This evaluation procedure is understood to apply only in the order given. This rule trades on the paradigm-shifting monograph by Donald Campbell and Julian Stanley, Experimental and Quasi-experimental Designs for Research on Teaching (1965). Still taught widely today in lit review courses, Campbell and Stanley identify some types of flaws or uncertainty that block conclusions from research because of the design of the study, pure and simple. These are called issues of internal validity. There is a second type of validity, external validity, where flaws or uncertainty block application of research findings to practical application. In the standard test of EBM proposed by Straus and colleagues, if the standard of internal validity for the research in and of itself is satisfied, that a “good enough” estimation of external validity can be applied to determine whether to use the research on a particular patient. The authors frame the decision about external validity in these terms: “Are the study patients so different from ours that we should not use the results at all in making predictions for our patients?” (Straus et al, 2005, p. 112).

It makes equal sense to reverse these screens on the evidence. Does the research speak exactly to the patient I have in the chair? If so, is there any reason why the found evidence should not be used to guide treatment? The pyramid of evidence may be sometimes upside down when considering the issue of applicability of the evidence to various clinical practice settings. The order of privileging methodological rigor over clinical applicability in EBD makes sense only where it can be demonstrated that a little help is worse than none. Low-quality research is not as effective as high-quality research in reducing the
variance in treatment outcomes, but all such information benefits patients unless a clear case can be made that the information is biased.

**Does Evidence Make or Support Decisions?**

Not all high quality dental research counts as evidence. Understanding the properties of materials, common biological reactions to classes of chemical agents, and natural maturation patterns in young patients, for example, are good science that practitioners need to master, but they are wide of the EBD net.

What does it mean to say that evidence justifies a decision? There are two ways to answer this question. On one hand, evidence might be taken to mean that one treatment is superior to another (in a large number of cases similar to the one faced by the practitioner). In this sense, the burden of justification rests on the practitioner who decides to treat contrary to the evidence, as it currently does with respect to the standard of care. On the other hand, evidence may be understood as estimates of clinical parameters that inform the practitioner’s decision for each particular patient. The latter is the sense of evidence most often understood in EBM as it is presented from a theoretical point of view.

Specifically, evidence-based physicians are interested in getting information that will allow them to more accurately predict any of the following four factors in treatment: (a) diagnosis and screening—what is the likelihood that a patient presenting with a set of characteristics has any of several disease conditions? (b) prognosis—what is the most likely course of the condition, given both reasonable interventions and no intervention? (c) therapy—what are the most likely outcomes of various interventions? and (d) harm—what are the potential downsides of intervention? In each case, a parameter (probability, average life expectancy, quality of life, and so forth) is determined along with an estimate of the variation in the parameter. Physicians understand that accurate (close to the real case) and precise (close most of the time) estimates of parameters are helpful in caring for patients. Such information can be combined with clinical judgment and patient values to guide treatment decisions.

What about the other interpretation of evidence, the one that says this treatment is better than that one? That is what my PhD advisor called a “horse-race research design,” and it is uniquely suited to the randomized controlled trial (RCT) that is thought of as the “gold standard” in dental research. It is certainly correct that the RCT is good at picking winners. But it is not effective for deciding how other horses would run, or whether the track and weather make a difference, or even whether the owners of the horse want to accept the odds on offer. The “one-size-fits-all” search for superior treatments in head-to-head competition does not inform the practitioner or patient so much as pretend to make the choice for them. It is certainly not in the spirit of EBM. But commercial interests have certainly been quick to recognize the unreflective embrace of EBD and now regularly attach this adjective to marketing claims as if to relieve the practitioner of the need to integrate the information in making a clinical judgment in practice.

**The “Best Evidence” Problem**

Imagine that an important visitor is coming to town and you wish to entertain him or her properly. You intend to take your guest to the best restaurant in town. But your spouse correctly observes that you have some wonderful places you enjoy, but no French Laundry, Inn at Little Washington, or Tour d’Argent. Reluctantly, you tell your visiting friend that he or she is on their own because you do not have a “best restaurant.” That is silly, you say. That is not what “best” means; it means the best that is available.

Of course this is right, but it is not in tune with EBD enthusiasts. “Best” to them means research that meets standards for research rigor. What is envisioned is systematic reviews of large numbers of RCTs and other methodologically rigorous research.

It is assumed in the EBD sense of the qualifier “best” that all high quality research is determinative of practice and that low quality research is not. As Straus, Richardson, Glasziou, and Haynes (2005) put it: “If the study wasn’t randomized, we’d suggest that you stop reading it and go on to the next article in your search.” It is unlikely that we will ever have RCTs involving the effects of smoking or pregnancy on periodontal health because true randomization is not possible over some variables.

There are three problems here. First, the best evidence in the EBD sense may not justifiably influence practice decisions. It may be “proven” that too little is known with certainty to alter what professional judgment and patient values urge. It is very common to read in journals regarding contain systematic reviews: “1,562 articles were identified in the initial PubMed search; 341 of these met inclusion criteria for the review; of these 244 were excluded after review because they did achieve standards for “best evidence”; the 17 remaining studies were inconclusive.”

Second, the out-of-hand dismissal of research that does not meet the abstract criteria for “best” research assumes that excluded studies contain bias. The multiplication of biased studies will always misguide clinical practice; but the multiplication of low quality but unbiased studies will reduce clinical variance in...
judgment. That means multiple, unbiased studies can be used to provide evidence in EBD.

Third, “best evidence” may not exist. EBD supporters are silent on this point. The examples they present almost always demonstrate that the clinician’s question can be answered in a few minutes or even seconds with a computerized search. If “best evidence” does not exist, the clinician must still proceed with treatment. But the problem is a little deeper than “nothing ventured, nothing gained.” No stopping rule exists for EBD searches. How is it known whether it is worthwhile initiating or continuing a search?

Reasonable Doubt

John Iannidis (2005) has argued persuasively that over half the research published in medicine makes false claims. Regrettably, he is not able to clearly identify which these are.

But do knowledgeable people really have serious doubts about the role of evidence in professional practice? Who could be against EBD? Of course, the question is malposed. Who could be against fluoride, for example? Most scientists recognize that it is not just the fluoride that matters but how much, for whom, by what delivery system, etc.

Nancy Cartwright, in her 2009 presidential address to the Western Division of the American Philosophical Association, said: “RCTs are just the bad penny” (Cartwright, 2009, p. 18), meaning that they circulate more frequently than other currency because their actual value is less than their face value. Cartwright’s point is that randomization is effective in controlling those dimensions of a clinical decision over which one has randomized. But practice is so multidimensional that conspicuous randomization in one direction can actually be dangerous in lulling others to believe that all relevant dimensions have been randomized. We have a preoccupation for randomization across subjects in clinical trials, but not always for randomization regarding operator, setting, clinical protocol, and other factors. Counting subjects as the critical dimension also substantially inflates the sample size, thus exaggerating statistical significance. In the end, the application of evidence in practice always involves an n of 1.

Occasionally, EBD is used in a normative sense. Those with whom one disagrees are said to “lack evidence”; not publicly embracing EBD is sometimes understood as being unreasonable. Philosophers usually make a clear distinction between reasons and evidence. The reason is what causes an action (Bratman, 1999; Kolodny, 2005; Sober, 2008). “I was so tired that I did not notice the change in the patient’s color in time to administer pure oxygen.” By contrast, evidence is normative; it offers a prima facie justification for a course of action. Certain changes in patient color are an indication that pure oxygen, and other recovery techniques, should be administered immediately. But evidence controls action only in a prima facie fashion. Not all changes in patient coloration preclude continuation of treatment. Practicing in ways that do not match researchers’ understanding of best evidence is not automatically unreasonable. Considerations of cost, patient selection, comorbidities, and practice delivery systems may differ from the conditions that prevailed on average in controlled research settings.

Evidence contributes to patient care. Better evidence contributes more than motley evidence. But evidence does not dictate patient care, and characteristics of the evidence in the abstract are insufficient to determine the extent to which they are deterministic of treatment in particular situations.

Physicians understand that accurate (close to the real case) and precise (close most of the time) estimates of parameters are helpful in caring for patients. Such information can be combined with clinical judgment and patient values to guide treatment decisions.
The Clinical Experience Leg
Sometimes it is thrown up as an epithet that one must not be much of a practitioner if all one can say is “it works in my hands.” On the other hand, patients always want to go to dentists whose treatments work in the hands of those dentists. There is a deeper implication in the criticism, one that suggests that dentists may not be the best judges of what actually works. It is unlikely that EBD will get very far in promoting an integration of evidence and clinical experience when it begins from a position of disparaging rather than understanding the power of clinical judgment.

Observation Bias
There is a voluminous social psychology literature about unconscious bias, self-confirming hypotheses, herd mentality, placebos, and selective searching for information. [See Chambers, 2003; 2009 for summaries with dental examples.] There is no reason to believe that dentists would be immune from these “all-too-human” character traits. Dentists may misinterpret what actually works in their hands.

But this is also a dangerous argument for advocates of EBD. Literature searches are known to be biased, especially in terms of going with only the first positive hit (Kuiper et al, 2005). There is almost certainly a “Hawthorne effect” in new movements such as EBD. This phenomenon was named for a Western Electric plant that assembled telephones. When volunteers in the 1930s were recruited to participate in a study to assess the effects of brighter lighting on productivity, the results of the RCT were impressive. The trouble was that when the study was repeated to test for lower levels of lighting that reduced glare, the results were equally effective. In neither case did the results last. Novelty and a public expectation of change are often temporarily effective in the absence of any true causal intervention (Mayo, 1933).

Policy-Based Dentistry
There is, however, an extreme form of EBD that aims to drastically curtail or even eliminate clinical experience from practice. It might be called PBD, for policy-based dentistry. Critical paths or clinical guidelines are sometimes proposed as a means for reducing practitioner judgment. Once a patient has been classified, even tentatively, one or a few alternative treatments are initiated automatically. Some organizations have published “evidence-based consensus statements.” Often, these are advanced by specialty groups and have the effect of limiting involvement in certain procedures by general practitioners. Third-party carriers are also interested in PBD.

There are some advantages in using PBD. Treatment variation is reduced, allowing for more rapid detection of approaches that do not work. Policy is a platform from which opinions and evidence can be compared. Policy also allows more ready management of care for third parties by unitizing disease conditions. The role of diagnosis relative to treatment is also elevated in this system. It is worth noting that systematic reviews in dentistry are more often inconclusive than in medicine and that reviews in medicine are dominated by drug trials. It is more difficult to get con-
clusive evidence when operator variance plays a part in the treatment outcome.

PBD seems well suited to medicine. Physicians practice in hospitals, and staff working in committees there review evidence and craft critical paths that the entire medical staff agrees to. This is a nice blending of EBM and PBM, but it may not be so easy to achieve this synthesis in dentistry.

But PBD should be carefully distinguished from EBD. In the policy version, groups of practitioners pool the evidence and their practice experience to create guidelines. The patients’ and the practitioners’ values and the circumstances in the practice are not considered or are only allowed to the extent that a decision is made that the guidelines are applicable in the particular case or they are not. Exceptions require external approval.

Outcomes-Based Practice
There is a close cousin to EBD that does not rely on either primary or screened evidence from the research literature. All the evidence comes from within the individual practice itself. It certainly ought to be possible to apply objective and systematic methods to individual practices in order to minimize bias, just as such methods are applied in classical research. There are some very simple techniques for these purposes. The focus here is on practice outcomes (rather than practice inputs drawn from other people’s evidence). I have written a 12-part series of articles showing how existing charts, staff members, and simple tallies and computations can help identity what works reliably in a particular dentist’s office (Chambers, 2001b). The techniques lack the sophistication and complexity of publishable research. Most involve simple tallies and computations that can be completed on a hand-held calculator. The focus is not on evaluating the usefulness of what researchers are doing. It is on isolating the activities that make the largest differences in practice outcomes and then reducing the variations and unwanted surprises on those activities.

My outcomes-based practice (OBP) alternative to EBD has been criticized by some purists because the results are not generalizable across practices. The outcomes certainly apply only to the office where the data are collected and the findings cannot be published. There actually is not need to make claims that apply in other settings: all that is being done in OBP is applying some research rigor to observing evidence in order to improve outcomes in one practice only. That is what most practitioners are interested in.

The Patient Value Leg
When I first heard the term EBD, I was overjoyed. “At last we will give proper attention to ethics in dentistry.” There is little doubt but what patient values are so much the shortest leg of the stool that the seat is unstable.

There are two reasons for this lacuna. First, patient values are not included in the designs for research that constitutes the evidence in EBD. EBD is widely believed to be a discipline in the natural sciences, so intangibles such as values are difficult to study. (If it is correctly understood as centering on the integration of multiple courses of information in clinical decision making, EBD is actually a social science discipline.) The problem is not acute with regard to the relationship between evidence and practice judgment since it is often assumed that research evidence will simply substitute for clinical experience. The second reason why so little is heard of with regard to patient (and practitioner) values in EBD is that the application of values to diagnostic and treatment decisions requires a different logic from the logic of clinical research. Inferential tests of hypothesis, p-values, confidence intervals, measures of effect, and such are essential for research. But they are not the stuff of integrating this evidence in clinical decision making. This is a fatal impediment to the advance of EBD: attempts are being made to use the tools of research for the work of dental practice.

In medicine, the procedures for integrating evidence with patient values are better worked out than for integrating evidence with clinical judgment. What has not been developed is good systems for soliciting reliable expressions of patient values. This is a rich field in economics, where there is a good theory of utility functions (Binmore, 2009; Keeney & Raiffa, 1993). Marketing research also knows a thing or two about how to elicit the expressions of values that will drive behavior (Kotler & Clarke, 1987). The ADA Code of Professional Conduct specifically states that “the dentist’s primary obligations include involving patients in treatment decisions in a meaningful way, with due consideration being given to the patient’s needs, desires and abilities, and safeguarding the patient’s privacy.” Informed consent is the ethical position and legal requirement that patients (not practitioners) make the final selection of when to begin and end diagnostic testing and whether to engage in therapeutic interventions.

The implication is that use of evidence, absent consideration of patient values, is indefensible practice. Dlugokinski and Browning (2001) present research that bears on this point. They reported on the informed consent practices of dentists who were major users of composite for posterior restorations and those who were not. The driving force in informed consent was
the practitioners' personal beliefs about the procedures: those who preferred to use composites for posterior restorations found more patients in need of these restorations and systematically altered the presentation of benefits and costs to favor treatments the dentist preferred. This is a case where, at least potentially, an Internet search and an outcomes-based practice database of the efficacy of posterior composites in the practice could provide useful information. Such activities would not, however, furnish all the information needed to make the correct decision.

Paternalism is the ethical position that professionals are justified in selectively presenting information or even substituting their own decisions for those of the patient when patients are thought to lack the capacity to exercise their own judgment (Beauchamp & Childress, 2009). Unless one is prepared to take an extreme position that scientific evidence always overrides patient's values, some structure is needed for integrating such values into clinical decision making. It is evident that something more that what EBD has so far developed will be needed to manage this successfully.

Integration
Regardless of the irregularity in length and strength of the three legs of the EBD stool, if there is no place to sit comfortably that spreads the weight uniformly, this tool will be set aside as nonfunctional. There is virtually no research showing how dentists actually integrate diverse sources of information in their approaches to treatment.

At this point in our immature understanding of how practitioners integrate evidence, experience, and values in clinical decision making, perhaps we can do no better than inventory and sketch several alternatives. There are at least four such positions: (a) policy matching, (b) punctuated equilibrium, (c) probability updating or advance on baseline, and (d) decision trees. Very likely, future research will show that practitioners use these and other approaches in various combinations depending on circumstances.

Policy Matching
The simplest approach to grounding practice is to conform to a standard. The decision rule here is “follow the standard” or come close enough so that one could not be criticized for ignoring it. The standard of care is the obvious example, and those who depend on this approach appear to accept the position that research evidence will move the standard of care, in its time and to the extent the professional as a whole approves. The weight of professional opinion will emerge and until it does, it is safe to follow convention.

Everett Rogers's pioneering work on the *Diffusion of Innovations* (1995) describes the characteristics of early and late adopters of innovation. Dentists appear to have many of the socioeconomic features of early adopters, such as high formal education and socioeconomic status (but not working in large organizations), but also many of the personality features of late adopters, such as discomfort with uncertainty and preference for concrete over abstract issues (Chambers, 2001a). Following the suggestion that John Colobolos (1989) made regarding physicians, dentists who accept the policy matching strategy will prefer a minimal standard or one that allows the maximum of professional freedom in choice. The driving question is not what some researchers might think is slightly better practice but what the profession at large feels is allowable practice. On this view, many of the nuances in EBD will be entirely too fine to deserve the practitioner's time.

Punctuated Equilibrium
The concept of punctuated equilibrium is due to Niles Eldredge and Steve Gould (1985) and it refers to the view that biological evolution is not a steady and gradual process. Instead, there are long periods of stability, broken by episodes of rapid innovation. The analogy being suggested here is that dentists maintain stable practice patterns, and make changes only occasionally. They may retain a second-generation material and only move to a new product in the fourth or fifth generation. They may perform the occasional endodontic procedure long after they might have logical reason to discontinue this activity altogether.

The point about equilibria is that the entire practice exerts forces for stability and against change, even when a slight objective advantage might be demonstrated for doing something other than what habit recommends. Again, we come face to face with the strong possibility that the objective desirability of a product, material, or technique demonstrated in the literature cannot be evaluated only in the context of the literature.

The practice logic of punctuated equilibrium is counterfactual. That is a fancy way of saying that the practitioner has to make a bet on the value of EBD. Actually, there are two bets. The first one is whether it is worthwhile to search the literature. The second is whether it is worthwhile to change habits based on new evidence.

There is a simple rule for evaluating the wisdom of searching for information. Is the likely benefit of a search and change, minus the cost of the search and the cost of making the change, greater than the likely benefit of continuing
with the status quo? One of the great benefits from EBD has been to reduce the cost of searches. The increase in useful information from clinical epidemiology has not been as noticeable. The very real possibility remains that practitioners will not be able to find any information that specifically addresses many of the issues they face. As individuals tend to exaggerate the cost of those activities they do not understand, it is reasonable to expect that EBD searches will be rare unless there is a clear and unavoidable cost of continuing with the present approach. A similar cost formula applies to making changes once relevant information is found. The cost of implementing the new approach may not justify switching (including the cost of the learning curve and uncertainty) if there is a statistically significant difference in the literate, on average, across many patients treated in other settings.

**Advance on Baseline**

The first two approaches to integration involve all-or-none changes in practice habits, usually based on recommendations made by others who have no direct knowledge of the individual practitioner’s circumstances. We can call this the exogenous understanding of integrating evidence, experience, and values to improve practice, meaning “coming from outside the practice.” The next two approaches are endogenous in the sense that the evidence, experience, and values are evaluated afresh within the practice with each potential application. Of course, this decision making need not be from scratch with each situation. Sometimes, a challenging diagnosis or decision about whether to continue with diagnostic tests may present itself as a novel element in a familiar clinical context with well understood values. Perhaps, the patient presents with values that challenge evidence and experience, as might be case with a Christian Scientist.

The essence of endogenous integration of evidence, experience, and values is that a new and particular decision is called for that extends beyond habit.

In medicine, EBM approaches to clinical decision making are endogenous, typically focused on reducing the variance around an estimated parameter. Perhaps the physician knows little or nothing of the patient’s five-year survival rate given a certain form of cancer. A good literature search could provide a more accurate probability estimate. In dentistry, one may be confused by competing claims from two bonding agents. A review of several good RCTs may show that one has sheer strength of 32.5 MPa and the other 33.1 MPa. The idea is that more accurate estimates of relevant parameters facilitate better decisions.

Sometimes the analysis and interpretation of such parameters is complex, often well beyond the training or time available for practitioners. Sometimes the exact type of estimate sought cannot be located (usually because it does not exist). Sometimes the results are expressed in units or for conditions that are not exact matches with the presenting patient condition. Sometimes the parameter estimates from the literature are at odds with the estimates from clinical experience.

The reflex response to complex or conflicting information is usually to pick one or another value and ignore the rest. That, however, is not an integration of available information; it is a privileging of one source over another. There is a well-established procedure for integrating probability estimates for parameters that is statistically sound and perfectly accurate. It is called Bayesian updating (Hoff, 2009; Silvia, 2006), and a simple formula is used to combine new evidence.
with what one already “knows.” Thus, new evidence either permits an “advance on baseline” or it does not. The expected advance on baseline should always be greater than the cost of a search for new evidence.

As is the case with “best evidence,” there is a subtle, but significant confusion in EBD over what it means to apply evidence from research in the clinical setting. It may seem obvious that the parameter estimates (probabilities of one treatment being superior to another, a diagnostic test having a certain accuracy, or the odds of a risk factor or side effect mattering, for example) should slide over just fine from summarized RCTs. Statisticians would state the relationship between evidence and the research observations they are based on with the expression \( Pr(E|O) \), where \( E \) is a research claim and \( O \) are the data researchers use to justify their claim and the “|” means “given” or “based on.” From the practitioner’s perspective, the expression looks more like \( Pr(O|BE & PJ & PV) \). Here \( O \) stands for clinical outcomes, \( BE \) is the best evidence the clinician can find, \( PJ \) is professional judgment, and \( PV \) expresses the patient’s and the practitioner’s values. This express just states that clinical outcomes are a function of the integration of evidence, judgment, and values—exactly what EBD is supposed to be. It is immediately obvious that \( Pr(E|O) \neq Pr(O|BE & PJ & PV) \). It is even true that \( Pr(E|O) \neq Pr(O|E) \), if we were to bracket out professional judgment and patient and practitioner values. The observant reader will say, this last inequality is simply due to fact that \( O \) in one equation stands for observation in a research study and \( O \) in the other stands for clinical outcomes in somebody else’s practice. That is correct: research observation and clinical outcomes are not the same thing. But even if they were, in the case where EBD enthusiasts are academics with part-time practices, \( Pr(E|O) \neq Pr(O|E) \) by mathematical necessity. There are ways around this problem, but they involve gathering information or making assumptions about factors beyond the research evidence and using Bayesian techniques.

What cannot be done with Bayesian approaches to integrating evidence and experience is to manage values.

**Decision Trees**

Although research shows that healthcare professionals are not especially good at Bayesian updating as would be required by EBD, they do seem to possess an intuitive and workable approach to practical decision making. One of the oldest and best established findings in decision science is that the expected value (what one expects to get) is the product of the value of the outcome should it occur and the probability that this outcome will occur (Bodily, 1985; Luce and Raiffa, 1957). This is abbreviated \( EV = Pr * V \); think of EVs as bets. And to make matters even more wonderful, complex paths through multidecisions networks are simply chains of EVs. Better still, the probability of any of the possible events occurring is always equal to 1.0. An example of a decision tree is shown in the side bar. In EBM this is known as clinical decision analysis or CDA. An example in dentistry is developed in the context of informed concern and agreement to participate in clinical research in my paper “Confusions in the Equipoise Concept and the Alternative of Fully Informed Overlapping Rational Decisions” (Chambers, in press).

As one might have guessed by now, the value part of decision trees is deter-
mined by the patient values and the likelihood is determined by the evidence, as adjusted by the practitioner’s experience. The overall structure of tree, knowing which alternatives to consider it the special province of the practitioner. Patients and dentists agree on which paths to pursue based on which will give both the patient and dentist the likely most desirable outcome. The whole tree can be evaluated using a pocket calculator, or more often by natural intuition.

Aside from the compelling fact that standard decision theory is the only approach so far considered that permits the systematic integration of evidence, judgment, and values, there are some further benefits. Decision analysis is action-oriented, rather than concerning itself with abstract claims. The researcher pays the price of advancing a claim that does not follow the canons of science; the practitioner pays the price of performing a procedure that damages a patient’s health. Only potential courses of action (as identified by the practitioner) have meaning. Thus, practitioners are always deciding between alternative actions, not theories.

Because the plausible course of action is the essential question in decision making, evidence can have meaning well beyond the accuracy and precision reported in the literature. This is called sensitivity analysis in technical terms, but the idea is very simply that information is only of interest if it might change the course of action chosen. In the case of two materials that had shear strengths above 30 MPa cited earlier, further RCTs to determine which really is superior would be a waste of time since both are clinically acceptable. The decision between the materials will be based on other considerations, such as cost, ease of use, familiarity, and “whether it works in the clinician’s hands.” That is exactly as it should be. When one course of action is dominated by advantages in another, differences, regardless of the relevant evidence, on the dominated path should not be considered at all. Several Nobel prizes in economics have been awarded for work that assumes this statement among the givens (Arrow, 1963; Nash, 1950). This is just the familiar argument that reasonable people will look for their missing wallet where they believe they have lost it rather than where engineers have placed streetlamps.

The argument for decision analysis and the for practitioners to perform sensitivity analyses to determine whether information about dominated alternatives could reasonably change the action taken is not a council of despair about EBD. Rather, what is being proposed is a rule of thumb for the wise use of evidence in practice: If it is plausible that evidence exists that would change the decision about how the patient in the chair should be treated, consider getting this evidence and integrating it with what else is known. There are two parallel rules: If the practitioner’s experience might be biased due to haphazard notice of treatment outcomes, a system of outcomes-based practice should be initiated, especially so since one of the steps in this process is to isolate those activities that most strongly influence clinical outcomes. Finally, if it is likely that the values of various patients or the values of the practitioner may make one or another alternative preferable, those values should be identified and integrated into treatment and ethics texts should be consulted for discussions of the limits of paternalism.

References
Confidence Intervals: It is common and useful practice to express the variation of the parameters estimated by clinical epidemiology in terms of confidence intervals. The problem is that such intervals that figure so prominently in the forest plots of systematic reviews and meta-analyses are generally taken to mean that the estimated parameter will fall within a CI\textsubscript{95} 95% of the time in individual cases. This, of course, is wrong. The correct interpretation is that if researchers conducted 100 research studies, the data reported in their papers would be replicated in 95 of the studies, on average. The CI\textsubscript{95} reported in the literature (unless it is specifically stated that Bayesian statistics were used) does not inform the practitioner about his or her patient. The problem is very simple: researchers present confidence intervals in systematic reviews that are appropriate to the research context and imply or allow practitioners to believe that this is a measure of the variation the practitioner should expect in his or her office. It is not.

Randomization: Most RCTs are semirandomized. Inclusion and exclusion criteria on patients may alter generalizability of findings in material ways. In other words, steps taken to increase internal validity normally reduce external validity. It is also necessary to randomize over all factors that may have a measurable effect on outcomes in the application context. I have published research showing that a substantial, if not the largest, source of variation in technical dentistry outcomes is the operator (Chambers, 2005; Chambers et al, 2009; Re et al, 2009). Most published research randomizes across patients, but not conditions, operators, or even the underlying scientific features of the materials and techniques used. The statistical models necessary to isolate the sources of variance that matter in clinical practice are complex. Performing large meta-analyses on studies that ignore these factors, however, does not correct this deficiency.

Confusions in the equipoise concept and the alternative of fully informed overlapping rational decisions. Medicine, Healthcare and Philosophy.


