

Debate

A philosophical analysis of the evidence-based medicine debate

Scott R Sehon*¹ and Donald E Stanley²Address: ¹Department of Philosophy, 8400 College Station, Bowdoin College, Brunswick, ME 04011, USA and ²Associates in Pathology, 500 West Neck Road, Nobelboro, ME 04555, USAEmail: Scott R Sehon* - ssehon@bowdoin.edu; Donald E Stanley - dstanley@tidewater.net

* Corresponding author

Published: 21 July 2003

Received: 01 April 2003

BMC Health Services Research 2003, 3:14

Accepted: 21 July 2003

This article is available from: <http://www.biomedcentral.com/1472-6963/3/14>

© 2003 Sehon and Stanley; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

Abstract

Background: The term "evidence-based medicine" (or EBM) was introduced about ten years ago, and there has been considerable debate about the value of EBM. However, this debate has sometimes been obscured by a lack of conceptual clarity concerning the nature and status of EBM.

Discussion: First, we note that EBM proponents have obscured the current debate by defining EBM in an overly broad, indeed almost vacuous, manner; we offer a clearer account of EBM and its relation to the alternative approaches to medicine. Second, while EBM proponents commonly cite the philosophical work of Thomas Kuhn and claim that EBM is a Kuhnian 'paradigm shift,' we argue that such claims are seriously mistaken and unduly polarize the EBM debate. Third, we suggest that it is much more fruitful to understand the relationship between EBM and its alternatives in light of a different philosophical metaphor: W.V. Quine's metaphor of the web of belief. Seen in this way, we argue that EBM is an approach to medical practice that is indeed importantly different from the alternatives.

Summary: We can have a more productive debate about the value of EBM by being clearer about the nature of EBM and its relationship to alternative approaches to medicine.

Background

Over the past decade, the term "evidence-based medicine" (or EBM) has gained considerable currency. EBM has been described as a "paradigm shift" [1] that will "change medical practice in the years ahead." [2] Some suggest that the principles of EBM should be part of the standard training of all physicians and that those physicians who violate its precepts should ultimately face license suspension. [3] Critics of EBM claim that "there is no evidence (and unlikely ever to be) that EBM provides better medical care," and that EBM is simply "following its own political agenda." [4] Other critics use even harsher rhetoric, claiming that EBM's assumptions are "absurd" and "irrational". [5] Some commentators claim a middle ground by

saying that EBM and other approaches should be "harmonized". [6]

The critics argue that EBM proponents overemphasize the value of clinical trials. These critics suggest that there are other aspects to medicine. Some note the fundamental role of basic science in understanding the physiologic mechanisms of the body, the biology of disease, and the cellular targets of drugs. [5,7] There are also critics who emphasize the value of clinical experience and the judgment of individual physicians; [7-10] these critics sometimes emphasize the *art* of medicine, [10] and contrast this with the *science* of medicine, or they speak of *techne* vs. theory or compassion vs. reason. And other critics might argue that there are circumstances in which obser-

vational studies (or outcome studies) are a better choice than clinical trials [11].

The questions raised by this debate are fundamental: How should physicians practice medicine? Should they do so in accordance with the principles of EBM? If so, how should this be accomplished? How should health care dollars best be spent? While there has been considerable debate about these questions, it seems to us that the debate also involves some confusion. The confusion is a good indication that some conceptual clarification is in order, and that is our principal aim in this paper. Our philosophical analysis will be limited to conceptual and epistemological issues; there are ethical issues that arise in the practice of evidence-based medicine, but we will not be addressing those questions.

Philosophers distinguish different types of questions. The questions just raised are, in this context, the *basic* questions, or the *first order* questions. By contrast, *second order* questions are questions about the first order questions, e.g., questions about the concepts employed in the first order questions. We can also refer to the second order questions as *conceptual* questions. In the present context, the second order questions would include: what is EBM? What are the alternatives to EBM? What is the relationship between EBM and the alternative approaches to medicine? Does EBM represent a paradigm shift? Without clear and compelling answers to these second order questions, the debate about the first order questions will be an exercise in futility. Obviously, if both sides in the debate have different ideas about what EBM is, then the opposing sides will be talking past each other when they argue about whether or not EBM should be practiced. We have to agree on what we are talking about before we can make progress on the question of whether it is a good thing. This seems obvious, but in the first section of our discussion we will argue that some overly broad characterizations of EBM have in fact done much to obscure the debate. In section two, we will argue that it is unhelpful and misleading to portray EBM as a paradigm shift and that such portrayals have polarized the debate in an unfortunate way.

We will claim that it is much more useful to see EBM and its alternatives in light of a different philosophical picture: W.V. Quine's web of belief metaphor.[12] Seen in this light, we can see that EBM is an approach to medical practice that is indeed importantly different from the alternatives. The Quinean metaphor helps us to understand these differences, and this in turn puts us in a position to evaluate much more effectively the first order normative questions about how medicine ought to be practiced. We will begin in the third section by explaining certain aspects of

his view that are most relevant to the EBM debate before applying the view in the fourth and final section.

Discussion

What is EBM?

As has been noticed by others [7,13], the debate about the value of EBM has been muddied by an unfortunate tendency to define the term "evidence-based medicine" in an overly broad manner. For example, here is one definition of EBM that has some claim to being authoritative: "The conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients" [14]. In a similar vein, Brian Haynes writes "Advocates of evidence-based medicine want clinicians and consumers to pay attention to the best findings from health care research that are both valid and ready for clinical application".[15] Taken at face value, these definitions seem merely to say that EBM is the wise use of the best evidence available. Given that characterization alone, it would be astonishing that there is any dispute about EBM. It would be equally astonishing that anyone could think EBM, defined in this manner, revolutionary or even useful. After all, who could possibly be opposed to using the best evidence wisely?

The very term "evidence-based medicine" seems similarly vacuous – as if any alternative to EBM means doing medicine based on something other than evidence. Even homeopaths base their treatments on evidence of a sort, and we suspect that homeopaths would even claim that they are conscientiously and explicitly using the best evidence available. Of course they are wrong about the latter claim, but nothing in these standard definitions of EBM tells us why homeopathy fails to be evidence-based medicine. Furthermore, few would suggest that the very idea of using evidence in medicine constitutes a revolutionary paradigm shift. Perhaps advocates of the term "evidence-based medicine" have a philosophical theory of evidence, a theory according to which everything that has passed for evidence prior to the EBM movement was not evidence at all. However, if they have such a theory, we have not seen it expressed.

In the Background section, we mentioned a number of criticisms that have been raised against EBM. With such critics, we have the beginnings of an intelligible debate, for we have several reasonably distinct approaches to intervention questions, and we can ask substantive questions about the strengths and weaknesses of each, and we can ask which should be used when or how they should be combined. The present point is simply this: there is a tendency for some proponents of EBM to duck these questions and avoid this debate by *defining* "evidence-based medicine" such that it includes the best possible combination of basic science, clinical experience, and clinical tri-

als. In so doing, the proponents of EBM come awfully close to simply *defining* EBM as the best way to practice medicine, whatever that may be. In other words, they respond to the second order conceptual question ("what is EBM?") by saying that EBM is whatever approach to medicine best answers the first order normative question ("how ought we to practice medicine?"). The proponents of EBM thereby give the illusion of having answered both sets of questions when in fact they have answered neither. They thus walk away from an important and substantive debate.

We emphasize that our concern about the definition of EBM stems not merely from an academic desire for philosophical clarity. Our concern is eminently practical: we believe that the debate between EBM and alternative approaches is important and can change medical practice. But if EBM is explicitly defined in a way that is essentially vacuous, the debate will be unproductive. If this debate is to yield fruit, we must therefore clarify what we mean when we speak of EBM.

The heart of EBM is the use of randomized controlled trials (RCTs) to determine the safety and efficacy of interventions, e.g. treatment, diagnosis, and prevention. Of course, EBM advocates have taken care to allow a role for other factors beyond evidence from RCTs. In particular, one must take into account the clinical state and circumstances of the patient, as well as the patient's own preferences and actions.[14,16–18] Even for EBM practitioners, the research evidence is but one component of medical decision making. But of course alternative approaches to medical practice will also take into account the patient's clinical state and values; this is not what separates EBM from the other approaches. What separates EBM from other approaches is the priority it gives to certain forms of evidence, and according to EBM the most highly prized form of evidence comes from RCTs (including systematic reviews) and meta-analyses of RCTs. So, for the purposes of this paper, we will take the term "evidence-based medicine" to refer essentially to the practice of taking RCTs as the strongly preferred form of medical evidence.

Is EBM a paradigm shift?

Proponents have hailed EBM as a "paradigm shift" or "revolution", in the sense of those words put forward by historian of science Thomas Kuhn.[19] At the same time, Brian Haynes acknowledges that "most scientists and EBM advocates are ignorant of the philosophy of science and give little or no thought to constructing a philosophical basis for their activities." [15] Indeed, Haynes goes on to write that "One hopes that the attention of philosophers will be drawn to these questions" [15]. In that spirit, we want to examine more closely the claim that EBM is a Kuhnian paradigm shift. The Kuhnian metaphor is famil-

iar to philosophers, but its implications are presumably less familiar to physicians. We have nothing against a metaphorical characterization of EBM and its alternatives; well-chosen metaphors can do much to illuminate conceptual (i.e., second-order) questions, and thereby can help us answer the normative (i.e., first-order) questions. However, we will suggest that the Kuhnian metaphor is a poor one and serves to obscure more than it illuminates.

According to Kuhn, "normal science" takes place when the scientific community shares what Kuhn calls a "paradigm". A paradigm is "entire constellation of beliefs, values, techniques, and so on shared by the members of a given community"[19]. The paradigm represents the shared context, and within that context, the scientists will work on remaining puzzles or problems. But within normal science, scientists do not seek "major substantive novelties"; [19] rather, "Mopping-up operations are what engage most scientists throughout their careers".[19] Kuhn suggests that the history of science is mostly constituted by periods of normal science. However, sometimes an accepted paradigm undergoes a crisis, followed by a 'scientific revolution' which overthrows the old paradigm and replaces it with a new paradigm. With the new paradigm, the scientific community has a new set of assumptions and a new set of puzzles to be solved. Kuhn refers to these revolutions as "changes of world view"; he suggests that "after a revolution scientists are responding to a different world".[19] Somewhat less metaphorically, Kuhn says that the post-revolutionary paradigm is, at least to some extent, *incommensurable* with the previous paradigm.

Is EBM a paradigm shift? To answer this question, we have to specify the alternative with which we are comparing EBM. One alternative to EBM is the basic science approach: studying the physiological mechanisms of the body and the biochemical properties of drugs. We think that EBM is clearly not a paradigm shift in relation to this. As we will explain more fully in section four, EBM is so clearly intertwined with and complementary to basic science that it would make little sense to see EBM as a paradigm *shift* away from basic science.

On the other hand, when we compare EBM with clinical experience and observational studies, we do have a shift in the sort of evidence that is most highly valued for diagnosis, therapy, and prognosis questions. Clinical experience and observational studies have much in common with each other and with RCTs. In all three cases, physicians attempt to answer the question of whether intervention X is effective for condition Y, and they do so by *trying* X in a certain number of cases and see whether it works. If a physician relies on her clinical experience, she reflects on past cases that were similar in certain respects and in which intervention X was tried. Observational studies are an

attempt to make this more systematic: a large number of patients with Y are given intervention X, and the results are observed and summarized. This improves upon unsystematic clinical experience because it avoids some problems of bias or selective memory. RCTs take clinical experience several steps further by imposing fairly elaborate controls.

So, the move to RCTs certainly constitutes a shift in some of the techniques of medical researchers. The move also reflects a change in some beliefs as well, for it reflects a change of view about the reliability of uncontrolled studies or clinical experience. So in moving from clinical experience to RCTs, we do have a change in many of the elements which make up a paradigm, according to Kuhn. Is the change enough to count as a paradigm shift? That is difficult to say, and we don't think the difficulty can be resolved by a more detailed description of the changes in medical research brought about by RCTs. The difficulty lies in the fundamental unclarity of the Kuhnian notion of a paradigm.

The question is this: at what point do changes in theory and practice of science amount to a paradigm shift? Where is the dividing line between changing one's beliefs and practices *within* a paradigm and moving to a *new* paradigm? If we do not have any idea what sort of dividing line there might be, then it is not clear that talk of paradigms comes to anything at all. The move to RCTs changed considerably the specific techniques behind studies of treatment effectiveness. On the other hand, there was no immense change in doctrines. Instead, there was a recognition of the fact that evidence gathered through clinical experience and observational studies was subject to a number of flaws: expectation effects (on the part of both the patients and the physicians), the fact that most medical conditions improve regardless of treatment, etc. So was this sufficient to count as a paradigm shift? Absent a clearer dividing line or a clear account of the *point* of calling some advances paradigm shifts, we suggest that any answer to this question would be arbitrary and the exercise pointless.

On behalf of Kuhn, one might suggest that a scientific community employs a revolutionary new paradigm precisely when the new theory and practice are *incommensurable* with the old. However, this suggestion creates two problems: one for the proponent of EBM specifically; and one for Kuhn's account more generally. The first problem is this: it is simply not credible to think that the theory and practices introduced by EBM are incommensurable with the previous practices and theory. Two theories would be incommensurable if they are untranslatable – i.e., if there are significant statements within each theory that cannot be translated at all into the language of the other theory.

Thus neither theory could be completely understood using only the terms and concepts of the language of the other theory. Accordingly, incommensurability suggests that scientists working under distinct paradigms would not be able to comprehend one another's claims. However, by this standard, it seems clear that there is no incommensurability between medical research as practiced before EBM and medical research after the emphasis on RCTs. Physicians working under each of the supposedly different paradigms would have no trouble understanding one another's claims. Any number of extant texts [14,20,21] describe the theory and practice of EBM, and the description is couched in terms that would be fully comprehensible to a medical researcher in the early 1940s – prior to the first well publicized randomized controlled trial. Physicians can also understand and benefit from observations made centuries ago e.g. in physical diagnosis. So, if we adopt incommensurability as the indicator of a revolutionary paradigm shift, then we have no reason to say that EBM constitutes such a shift.

The second problem with Kuhn's talk of incommensurability is that it is by no means obvious that there have been *any* new scientific theories that can be seen as genuinely incommensurable with the previous theories. After all, when Kuhn himself describes various alleged paradigm shifts, he always manages to tell us about both the old and the new scientific theories. Were his theory correct, this means he would have succeeded in doing that which is impossible. While the issues are too deep and complicated to be adequately addressed here, we will just say that Kuhn and his followers have not convinced us that the notion of incommensurable paradigms is even coherent.[22]

Why it matters

It seems to us that when proponents grandly describe EBM as a paradigm shift, they not only make a philosophical mistake, but they inadvertently have an unfortunate effect on the debate about the first order questions, for talk of paradigm shifts unduly polarizes the debate about the value of EBM.[23] Recall that Kuhn describes a paradigm as the "entire constellation of beliefs, values, techniques, and so on shared by the members of a given community".[19] When proponents of EBM suggest that it is a new paradigm, this certainly fosters the impression that an entire set of beliefs, values, and techniques is to be left behind, and that the whole world of medical research and clinical practice is completely different than it was in the days before EBM was recognized. This in turn strongly suggests that physicians have a stark choice: accept the new regime and completely reject the old, or defensively hold on to the old and reject EBM entirely. If something as grand as a complete shift of paradigm is in the offing, then the old physicians, with their old understandings,

training, and paradigm, are relics to be removed as soon as possible. The prospect of this kind of consequence would certainly lead to a defensive retrenchment and even anxiety. If you exaggerate the claims for EBM, you intensify the natural defensiveness of those who feel themselves not to be completely in line with the new regime. This is not a productive atmosphere in which to hold a debate about the merits of EBM.

In the next section, we suggest a different philosophical framework with which to examine the EBM debate. But we think that some progress would be made simply by avoiding overly broad definitions of EBM and by abandoning all talk of paradigm shifts.

Quine and the web of belief

We can make more progress by turning away from Kuhn and turning to the work of philosopher of science, W.V. Quine. Quine describes the totality of our beliefs – from ordinary common sense claims to the laws of physics – as a *web*, and he thus speaks of the *web of belief*. [12] A web, e.g., a spider web, has an exterior edge or frame, and an interior, consisting of radii and connecting points. At the edges, the web is secured to an already existing structure, e.g., a branch or wall. According to the analogy, the web of belief comprises statements; certain of the statements are closer to the periphery or frame, which consists of our sensory experience.

Those sentences that lie closest to the periphery of the web Quine calls *observation sentences*, while the interior of the web contains more theoretical claims. The observation sentences are those that are more likely to be denied or affirmed on the basis of immediate sensory experience. For example, a quick glance out of the window will determine whether or not we accept the sentence "It is raining." With other sentences, we are much less immediately influenced by sensory experiences. For example, consider the sentence, "Ronald Reagan was President in 1987." One might have a sensory experience that is apparently at odds with this sentence; one might see a list that includes Walter Mondale as the 41st President serving from 1985 through 1988. But we would not allow such apparently recalcitrant sensory experience to change our minds about the original sentence. We would conclude instead that the list was simply wrong. This is not to say that the sentence "Ronald Reagan was president in 1987" is immune from all possibility of doubt; sufficient evidence might indeed convince us that the sentence is false. But two points are in order. First, it would take much more than a simple one-time observation of anything, and, second, we would have to change many other beliefs as well. This is characteristic of theoretical sentences within the sciences too: to convince us of the falsity of a firmly accepted claim, we will need repeated experiments with congruent results,

and rejection of theoretical sentences typically requires the revision of many other related beliefs. Some theoretical claims will be even closer to the center of the web, e.g., laws of nature like "E=mc²" and laws of logic or mathematics. According to Quine, even claims as central as these are revisable. But because such claims are connected to so many other beliefs, their rejection would typically require revision to many, many sentences within the web.

So according to the Quinean metaphor, we use a vast network or web of beliefs with intricate logical and evidential relations; at the periphery of the web are the observation sentences, which represent claims that are the most dependent on sensory observation. In the medical context, observation sentences would include simple findings during a physical examination, e.g., a patient's temperature or blood pressure. Moving inward from the periphery of the web would be hypotheses about these findings, e.g., that the patient has an infection, or that if the systolic pressure is 160 mmHg, we should repeat the measurement before concluding that it represents that patient's normal blood pressure. Even further towards the interior of the web would be theories about disease occurrence, e.g., that HIV is a retrovirus that targets CD4(+) T cells; that the differential diagnosis of hypertension includes idiopathic, renal stenosis and even pheochromocytoma; even further in the interior would be very general claims, e.g., the physiological model for elevated blood pressure or the germ theory of disease.

It is important to note that Quine is *not* suggesting a foundationalist picture, according to which we build our theory on supposedly indubitable reports of sensory experience. Quine writes:

Any statement can be held true come what may, if we make drastic enough adjustments elsewhere in the system. Even a statement very close to the periphery can be held true in the face of recalcitrant experience by pleading hallucination or by amending certain statements of the kind called logical laws.[12]

One of Quine's central claims is "that our statements about the external world face the tribunal of sense experience not individually but only as a corporate body".[12] Thus our theories are put to the test as a whole; we cannot simply test individual claims in isolation from the rest of our theory. This Quinean view, which seems to us perfectly correct, is sometimes referred to as *holism*, or as the Quine-Duhem thesis (for Pierre Duhem had argued for a similar view.[24]).

Nonetheless, observation sentences do have a certain privileged status. Observation sentences are those that are most directly tied to our immediate sensory experience;

given the right sensory experience, all competent observers fluent in our language will unhesitatingly assent to the appropriate observation sentence. For example, while looking at a thermometer, all competent observers could agree that it reads 37.0 C.; and given the right visual, aural, and tactile stimuli while using a sphygmomanometer, all competent medical professionals could agree that the patient's blood pressure is 130/80 mmHg at the time of the measurement. Having made such observations, we will also be very reluctant to change our minds about the truth of those sentences. Of course, we might well come to conclude that the patient's blood pressure has *changed*, or that the reading does not represent the patient's normal blood pressure (if, for example, we had reason to think that the patient was under stress at the time). However, having made our observation, we will remain quite confident that the original reading was correct at the time. Naturally, we might be forced by other observations to conclude that we had misread the sphygmomanometer, the cuff was wrongly inflated, or even that the instrument was not functioning properly. But in general, we will be very reluctant to reject a previously accepted observation sentence. Our developing scientific theory must somehow accommodate the observation sentences we accept. In this sense, sensory experience and observation sentences are the ultimate checkpoint for scientific theory; they are that which secures the web to fixed ground.

Application to EBM and the alternatives

EBM is principally concerned with what we have called *intervention questions* of the form, 'is X safe and effective for condition Y?'. Answers to these questions will fit somewhere in the web of belief. They will clearly not be located at the very center of the web like principles of mathematics or laws of physics; in the face of recalcitrant empirical data (e.g., data from RCTs) we will revise our views about treatment efficacy much more readily than we will reconsider what we take to be basic laws of nature. On the other hand, answers to intervention questions will also not be at the very periphery of the web, for questions about the safety or efficacy of therapy are also clearly not answerable by simple observation statements. With something as complex as the human body, we will not be able to simply *observe* the exact effect of a particular therapy. (Of course, some people fail to understand this. Some people think that they can just *see* that, e.g., their breast implant caused some awful condition that developed subsequently. [25]) So questions about treatment efficacy and safety will not be on the outer periphery of the web, though they will also not lie too deeply in the interior.

What is the relationship between EBM and its alternatives, as seen from the perspective of the Quinean metaphor of the web of belief? The answer depends on which of the alternatives we are considering. We'll first look at the third

alternative to EBM, namely, deduction from basic scientific knowledge.

EBM and Basic Science

From our knowledge of human physiology, disease, and pharmacology we might be able to infer whether a particular drug would be effective in treating a given condition. With the basic science approach, we work up from our knowledge of physiology and biochemistry to a prediction of what will happen. Clinical trials essentially ignore, or suspend temporarily, all of that and determine whether a treatment works by *trying* the treatment in a large number of cases under controlled conditions.

Basic science and RCTs thus represent two different paths to answering intervention questions. Seen explicitly in terms of the web metaphor, RCTs seem to take the shortest path possible from the periphery of the web towards the mid-web space where clinical decisions are made. Within the course of a clinical trial, we will make observation statements that seem very directly relevant to our original treatment question, for we will be observing clinical endpoints for other patients with condition Y when given treatment X. The basic science approach, on the other hand, looks rather different. On the basis of biochemistry and physiology, we might predict certain clinical endpoints for patients with condition Y when given treatment X. These predictions will be based on a very different set of observation statements than those involved in RCTs. The observation statements that ultimately support the basic science approach will be those that support our knowledge of biochemistry and physiology.

Although EBM and the basic sciences embody different approaches, this does not mean that they are competitors. In fact, the two approaches need each other; neither can stand alone. This is perhaps easiest to see with basic science. If basic science could give us perfect confidence in the safety and efficacy of an intervention, then clinical trials would indeed be theoretically superfluous. However, we rarely, if ever, can be certain of both the safety and efficacy of a treatment without clinical testing, for our knowledge of the human body and how it interacts with the environment is far from complete.

While it is perhaps less obvious, the method of clinical trials is also not able to stand alone. RCTs can sometimes give us confident answers to intervention questions even when our basic scientific knowledge is insufficient for this purpose (e.g., on many questions concerning diet and exercise); however, this does not mean that RCTs are completely independent of basic science. If nothing else, a controlled trial needs a hypothesis to test. Typically, treatments will be suggested as candidates because of other more basic scientific research. A classic case of this is the

accidental discovery of penicillin by Alexander Fleming, though Fleming himself initially paid little attention to the possible therapeutic uses of the mold.[26] Contemporary basic scientific work on thalidomide provides another example.[27]

But there is also a more fundamental way in which RCTs cannot stand apart from basic science. Even when a clinical trial returns positive results in the treatment arm that satisfy tests of statistical significance, we will have more confidence in these results when they have some antecedent biological plausibility.[28,29] Put more generally, we would suggest that the degree of confidence appropriate for a clinically tested claim is a function of *both* the strength of the clinical result and the claim's antecedent biological plausibility. This relationship is perhaps obscured by cases in which we understand very little about the pharmacology of a drug (e.g., treatment of ulcerative colitis with sulfasalazine). In such cases, it is perfectly appropriate that strong clinical results yield a reasonably high degree of confidence. But we shouldn't forget that we would be even *more* confident if the results were exactly what biochemistry and physiology would lead us to expect.

More importantly yet, if the basic sciences gave us very strong reason to believe that a drug would *not* be effective, then it is appropriate to be very cautious when interpreting apparently positive clinical results. For example, there have been some RCTs of homeopathic remedies according to which subjects receiving the homeopathic remedy did better than those receiving a placebo, and where the P-values are less than .05 or even less than .001.[30] However, given that homeopathic remedies are typically composed of nothing more than water, accepted basic science should lead us to seriously doubt that these 'drugs' can be effective. In any event, we cannot simply accept that pure water (in the form of a homeopathic medicine) is medically effective unless we are also willing to revise much of our basic scientific conception of chemistry and the body. The interior of the web acts as a compass we use to interpret the direction we should take from signals from a new RCT. *Statistical information from an RCT is virtually uninterpretable and meaningless if stripped away from the backdrop of our basic understanding of physiology and biochemistry.* The dependence of RCTs upon the backdrop of basic science is unsurprising from the Quinean perspective. The point is essentially an application of the Quinean doctrine of holism, the claim that our theories are put to the test as whole bodies rather than being tested sentence by sentence. (Tonelli and Callahan [31] make a similar point about the need for theory in interpreting empirical results. However, they suggest that evidence concerning alternative medicine should be weighed and interpreted in terms of *its* theory of disease, e.g., theories involving undetecta-

ble spirit essences or the like. This does not follow at all. If *we* are judging the effectiveness of alternative medicinal therapies, then we evaluate the evidence in light of the theory about disease that we believe to be true. It would be a silly and pointless exercise to evaluate evidence in light of a theory we believe to be false.)

The dependence of RCTs on basic science appears not always to have been fully appreciated by some EBM proponents. In some discussions of hierarchies of medical evidence, basic physiology and biochemistry is put at the very bottom or is ignored altogether.[14,32] Thus, these proponents of EBM seem to say that we should *only* look at evidence from RCTs (when we have it), and that medical practice is conducted in isolation of our understanding of physiology and biochemistry. We claim that such views are wrong. On the other hand, this is not to say that if a clinical result flies in the face of our physiological understanding that we should just reject it out of hand, as suggested by Couto [5]. In general, given the complexity of the human body and our nascent knowledge of it, if we use basic science alone to judge the safety or efficacy of an intervention, then our judgment will be fraught with a high degree of uncertainty. For example, some laboratory evidence suggests that estrogen might have a positive effect on brain cells, and thus suggesting that hormone therapy might help to ward off dementia. However, a recent controlled trial strongly suggests that whatever positive effect such hormones have, these effects are outweighed by negative effects.[33] On the other hand, in an extreme case like homeopathy, where the biological plausibility is so low, our background biological knowledge does and should affect our interpretation of the evidence from RCTs.

The outcome of our discussion above is that basic science and EBM are deeply intertwined within the web and complementary. In Quinean terms, each of these represents a different path to the intervention question through the web of belief. We clearly need both approaches. Which of the approaches should receive the greater share of our limited health care resources? There is little useful to be said about this question at this level of generality. The best allocation of resources is affected by a broad range of contextual factors.

EBM and the other alternatives

As noted earlier in the discussion of Kuhn, clinical experience, observational studies, and RCTs have much in common. All are attempting to ascertain the safety and efficacy of interventions, and all do so by trying the intervention and noting the results. From the Quinean perspective, we could also say that all three are anchored to observation statements at the periphery of the web in similar ways. In each case, we will observe the treatment received by a

patient, and then we will observe the outcome or end-point for each patient. Of course, in an RCT, we place restrictions on the way data are gathered: we require that there be a control group of patients who do not receive the target experimental therapy, and that patients are randomly assigned to either the control or experimental group; we typically 'blind' researchers (and other participants), by not allowing them to know, during the course of the study, which patients are in the control group; and we blind the patients themselves, so that they do not know which therapy they received. In this section, we will look at how these restrictions are to be understood and justified from the perspective of the Quinean metaphor.

In observational studies or clinical experience, the physician knows which subjects received the therapy or exposure, and then observes the patient's subsequent condition. Thus, statements about the patient's condition constitute the evidential base, and these statements are treated as observation sentences. But recall that an observation sentence is one that would elicit assent from any competent observer. The advocate of double-blind trials contends, in effect, that simple statements about the patients' condition are not observation sentences after all, for the sentence might *not* elicit assent from all competent observers – especially for 'soft' end points such as pain scales. Physicians may be unconsciously biased in favor of seeing or eliciting a treatment effect when none exists. A doctor who knows that the patient has received a certain therapy, and who thinks that the therapy is likely to be effective, might make observations that would not be made by a physician who knew nothing of the treatment. So the criticism of clinical experience and observational studies is this: the evidence in those cases is tainted in that physicians take certain sentences about the patient's condition to be observation sentences (and hence firm checkpoints) when in fact they do not qualify as observation sentences. Hence, according to the proponent of EBM, the conclusions reached through clinical experience are less reliable.

In a double-blind trial, we also blind patients from information about what sort of therapy they received. Here, the main concern is not that the patients will be subconsciously biased in reporting their condition – though this might be a concern in some cases. The issue is very different: we want to rule out placebo effects (or the "meaning response" [34]), meaning that we want exclude the possibility that the observed effect stemmed from the patient's *belief* that she was given a particular therapy. We might have reliable observation sentences about the patient's subsequent condition, but, unless we control for placebo effects, the improvement in a patient's condition might have nothing to do with the actual treatment received. In Quinean terms, in an uncontrolled study, the observation

statements (about the patient's condition) and the theoretical statement (about the efficacy of the treatment) may be linked too tenuously along the radii of the web; in the extreme case, the apparent linkage may be purely by chance. If we control the conditions under which we make observations, we can then compare the observation statements concerning the condition of patients who received the treatment with statements concerning patients who received a placebo. In other words, we use observation sentences similar to those that would be obtained in clinical experience, but we are more careful in the inferences we draw from these statements. The EBM proponent contends that by including observations about patients in the control group, we put ourselves on firmer ground in determining the efficacy of the treatment; thus the web is strengthened.

Perhaps the most basic feature of an RCT is the use of randomized control groups. When research is done without using control groups at all, researchers might infer treatment was effective, when in fact the patient's condition would have waned on its own or would have improved just as much with the standard therapy. In other words, the EBM proponent claims that in the *absence* of a control group, the inference from observation statements (about the patient's condition) to the conclusion (about the efficacy of the intervention) is greatly weakened. Moreover, if we do not *randomly* assign patients to either the experimental or control groups, then we again run the risk of researcher bias affecting the outcome, e.g., by selecting patients for the experimental group who have the best chance of improvement.

Returning to the first-order, normative question, we can now ask to what extent should physicians rely on RCTs as opposed to clinical experience or observational studies. Unlike the comparison with basic science, we cannot simply say that they are complementary and that we need both. As discussed above, observational studies essentially are clinical experience made systematic, and RCTs are observational studies subject to further controls. The extra time and expense involved in performing RCTs is justified only if their answers to intervention questions are more likely to be true. For the reasons laid out in the preceding paragraphs, the proponent of EBM contends that RCTs are indeed more reliable than the alternatives. We are inclined to agree. It is clear that for clinical experience and observational studies, there is a risk of different sorts of bias; in Quinean terms, the controls imposed by RCTs give us firmer observational checkpoints and stronger inferences from observations to conclusions, particular conclusions about interventions. Experience confirms that uncontrolled studies can give different results than RCTs [35,36]; the Quinean analysis gives a diagnosis of those differences.

Still, that doesn't mean we should reject clinical experience or observational studies altogether. Much of our web of medical theory was spun from these sources. Moreover, as a number of authors have pointed out, there are significant limitations on when RCTs are practical or even ethical. Where limitations of this sort intervene, we have to do the best we can, and that means relying on the experience we have through observational studies and clinical experience, even though we know that there may be problems of bias and selective memory, etc. If there are EBM proponents who think that medical practice can be reduced to an algorithmic application of evidence from RCTs, they are certainly mistaken.[37]

Summary

Because we wished to clarify the debate about EBM, we have spent the bulk of this paper exploring the second-order conceptual questions concerning the nature of EBM. We have argued that it is both a philosophical and practical mistake to see EBM as a paradigm shift. We have also attempted to explain the differences between EBM and its alternatives in terms of Quine's metaphor of the web of belief. EBM and the basic sciences offer different paths through the web to claims about intervention safety and efficacy: each begins with a very different set of observation statements as its evidential checkpoint; however, neither is completely self-standing and independent. There is a sharper contrast between EBM and the methods of clinical experience and observational studies, according to the EBM proponent: uncontrolled studies take certain reports of clinical endpoints as observation sentences – i.e., sentences that would be agreed to by all competent observers, sentences that are the checkpoint and basis for science – when in fact some of those reports do not count as observation sentences; and in other cases, uncontrolled studies may base their conclusions on observation sentences that are too tenuously linked to the claimed result, resulting in a *post hoc ergo propter hoc* fallacy. This does not mean that RCTs are always appropriate or are always feasible, and there is much more to be said about the first order questions of when RCTs are worth the time and expense. But we can only have that discussion when we are clear about the nature of the questions themselves.

Competing Interests

None declared.

Acknowledgments

We would like to thank Dr. Henrik Wulff and Dr. Ross Upshur for detailed comments on earlier drafts of this paper.

References

- Guyatt GH, Haynes B, Jaeschke R, Cook D, Greenhalgh T, Meade M, Green L, Naylor CD, Wilson M, McAlister F and Richardson WS: **Introduction: The Philosophy of Evidence-Based Medicine** *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice* Edited by: Guyatt G and Rennie D. Chicago, AMA Press; 2002:3-12.
- Eisenberg John M.: **Evidence-Based Medicine** *Expert Voices* 2001.
- Muney Alan M.: **Evidence-Based Medicine Needs To Be Promoted More Vigorously** *Managed Care* 2002.
- Goodman NW: **Who will challenge evidence-based medicine?** *J R Coll Physicians Lond* 1999, **33**:249-251.
- Couto JS: **Evidence-based medicine: a Kuhnian perspective of a transvestite non-theory** *Journal of Evaluation in Clinical Practice* 1998, **4**:267-275.
- Shaughnessy AF, Slawson DC and Becker L: **Clinical jazz: harmonizing clinical experience and evidence-based medicine** *J Fam Pract* 1998, **47**:425-428.
- Charlton BG and Miles A: **The rise and fall of EBM** *QJM* 1998, **91**:371-374.
- Tanenbaum SJ: **What physicians know** *The New England Journal of Medicine* 1993, **329**:1268-1271.
- Miles A., Grey J.: **New perspectives in the evidence-based healthcare debate** *Journal of Evaluation in Clinical Practice* 2000, **6**:77-84.
- Miettinen OS: **The modern scientific physician: 1. Can practice be science** *Canadian Medical Association Journal* 2001, **165**:441-442.
- Benson K and Hartz AJ: **A comparison of observational studies and randomized, controlled trials** *Am J Ophthalmol* 2000, **130**:688.
- Quine WV: **From a Logical Point of View** 2nd edition. Cambridge, MA, Harvard University Press; 1953.
- Tonelli MR: **The philosophical limits of evidence-based medicine** *Academic Medicine* 1998, **73**:1234-1240.
- Sackett DL, Straus SE, Richardson WS, Rosenberg W and Haynes RB: **Evidence-Based Medicine: How to Practice and Teach EBM** 2nd edition. Edinburgh, Churchill Livingstone; 2000.
- Haynes RB: **What kind of evidence is it that Evidence-Based Medicine advocates want health care providers and consumers to pay attention to?** *BMC Health Serv Res* 2002, **2**:3.
- Haynes RB, Devereaux PJ and Guyatt GH: **Physicians' and patients' choiced in evidence based practice** *British Medical Journal* 2002, **324**:1350.
- Haynes B, Devereaux PJ and Guyatt G: **Clinical expertise in the era of evidence-based medicine and patient choice** *APC Journal Club* 2002, **136**:A11-A14.
- Guyatt GH, Meade MO, Jaeschke R, Cook D and Haynes B: **Practitioners of evidence based care** *British Medical Journal* 2000, **320**:954-955.
- Kuhn Thomas S.: **The structure of scientific revolutions** 2d, enl.th edition. Chicago, University of Chicago press.; 1970:xii, 210.
- Guyatt GH and Rennie D: **Users' Guide to the Medical Literature: A Manual for Evidence-Based Clinical Practice** Chicago, AMA Press; 2002.
- Wulff HR and Gotzsche PC: **Rational Diagnosis and Treatment: Evidence-Based Clinical Decision Making** 3rd edition. Blackwell Publishers; 2000.
- Davidson D: **On the Very Idea of a Conceptual Scheme** *Proceedings and Addresses of the American Philosophical Association* 1974, **47**:5-20.
- Miles A., Bentley P., Polychronis A., Grey J.: **Evidence-based medicine: why all the fuss? This is why.** *Journal of Evaluation in Clinical Practice* 1997, **3**:83-86.
- Duhem Pierre: **La Theorie physique: son objet et sa structure** Paris; 1906.
- Angell Marcia: **Science on Trial: The Clash of Medical Evidence and the Law in the Breast Implant Case** W.W. Norton & Company, Inc.; 1996.
- Kiester E: **A Curiosity Turned into the First Silver Bullet Against Death** *Smithsonian* 1990, **21**:173-188.
- D'Amato RJ, Loughnan MS, Flynn E and Folkman J: **Thalidomide is an inhibitor of angiogenesis** *Proc Natl Acad Sci U S A* 1994, **91**:4082-4085.
- Goodman Steven N.: **Towards Evidence-Based Medical Statistics. 1: The P Value Fallacy** *Annals of Internal Medicine* 1999, **130**:995-1004.
- Goodman Steven N.: **Toward Evidence-Based Medical Statistics. 2: The Bayes Factor** *Annals of Internal Medicine* 1999, **130**:1005-1013.

30. Linde K, Clausius N, Ramirez G, Melchart D, Eitel F, Hedges LV and Jonas WB: **Are the clinical effects of homeopathy placebo effects? A meta-analysis of placebo-controlled trials** *Lancet* 1997, **350**:834-843.
31. Tonelli M.R., Callahan, T.C.: **Why alternative medicine cannot be evidence-based** *Academic Medicine* 2001, **76**:1213-1220.
32. Centre for Evidence-Based Medicine: **Levels of evidence and grades of recommendation** 2003, **2003**: [http://www.cebm.net/levels_of_evidence.asp]. *Centre for Evidence-Based Medicine*
33. Shumaker SA, Legault C, Thal L, Wallace RB, Ockene JK, Hendrix SL, Jones BN 3rd, Assaf AR, Jackson RD, Morley Kotchen J, Wassertheil-Smoller S, Wactawski-Wende J: **Estrogen Plus Progestin and the Incidence of Dementia and Mild Cognitive Impairment in Postmenopausal Women: The Women's Health Initiative Memory Study: A Randomized Controlled Trial.** *JAMA* 2003, **289**:2651-2662.
34. Moerman Daniel E. and Jonas Wayne B.: **Deconstructing the Placebo Effect and Finding the Meaning Response** *Annals of Internal Medicine* 2002, **136**:471-476.
35. Schulz K.F., Chalmers, I., Hayes, R.J., Altman, D.G.: **Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials.** *JAMA* 1995, **273**:408-412.
36. Guyatt G: **Therapy and harm: Why study results mislead--bias and random error** *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice* Edited by: Guyatt G Rennie D. Chicago, AMA Press; 2002:223-229.
37. Upshur RE: **Seven characteristics of medical evidence** *Journal of Evaluation in Clinical Practice* 2000, **6**:93-97.

Pre-publication history

The pre-publication history for this paper can be accessed here:

<http://www.biomedcentral.com/1472-6963/3/14/prepub>

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

