

## **PART I**

# Introduction

COPYRIGHTED MATERIAL



## CHAPTER 1

# The philosophy of evidence-based medicine

This is a thorough analysis of the justification for using evidence-based medicine (EBM) methodology. Why should we believe that EBM methods provide more reliable knowledge than other methods? While many have criticized various aspects of EBM, the system as a whole has, with a few notable exceptions [4,5], escaped careful scrutiny. One can, of course, raise critical questions about the foundations of EBM without denying its value [1]. And, in fact, my overall conclusions are mostly sympathetic with the EBM position and a central aim of this book is to clarify misunderstandings of what EBM actually involves. Much work in the philosophy of science is relevant to this analysis, including the logic of scientific discovery, the problem of underdetermination, the nature of causal inference and above all the logic of evidence (confirmation theory). Philosophers who are interested in how these central issues in the philosophy of science apply to contemporary medical science should find new and relevant material here. At the same time, medical professionals who would like to examine the underlying reasons why they should (or should not!) use EBM methods to determine whether the treatments they prescribe “work” will find this analysis useful.

### 1.1 What on earth was medicine based on before evidence-based medicine?

Loosely speaking, three overlapping methods for determining whether treatments are effective have competed for dominance in the history of medicine. One school has insisted that the effects of medical treatments must be

observed directly, usually by comparing groups of people who receive the treatment with groups who do not [6–8]. Another school has demanded that the underlying causes (“mechanisms”) of health and disease must be specified before concluding that a treatment caused a cure [6,9]. In parallel with these two schools, authoritative pronouncements of clinical “experts” have often played a powerful role, sometimes trumping external evidence. The EBM movement recently weighed in heavily on the side of the first method.

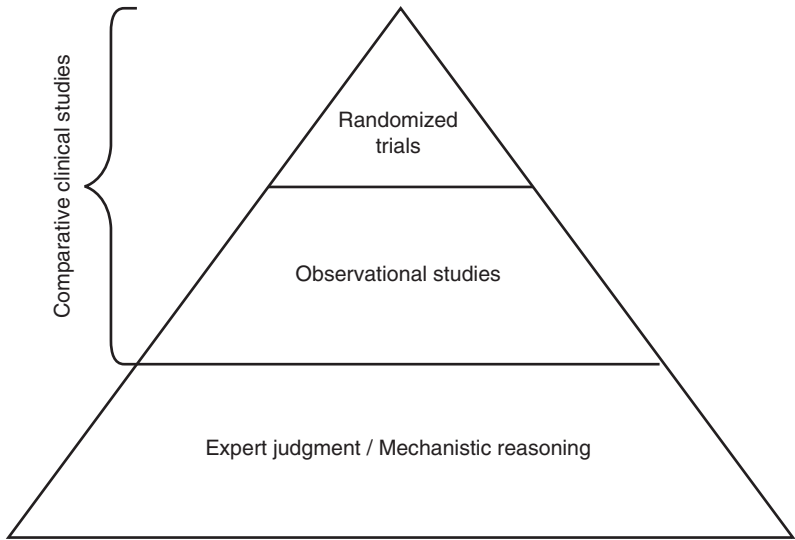
With a rhetorical *tour de force*, EBM was introduced as a “new paradigm” in the early 1990s [10–12]. Less than two decades later, there are at least seven journals, a dozen books, thousands of new citations to EBM each year, and a growing number of international research centres dedicated to the practice, teaching, and dissemination of EBM. Prominent medical journals, including the *British Medical Journal*, *Journal of the American Medical Association*, and *Annals of Internal Medicine*, endorse editorial policies encouraging researchers to follow the EBM rules of evidence [13], and the *New York Times* judged EBM to be the idea of the year in 2001 [14]. EBM has also colonized other disciplines. Social scientists [15], policy-makers, and even chaplains [16] are eager to demonstrate that their practices are “evidence”-based.

But what on earth was medicine based on before 1990? Given that “evidence” simply means “grounds for belief” [17], medicine has always been evidence-based by definition. Barring cases of deliberate deception, even physicians deemed to be quacks have had grounds to believe that their therapies worked. If EBM is something new, and its proponents insist it is, it must be a specific view of what counts as (good) evidence.

The EBM “philosophy” of evidence is best expressed in the EBM “hierarchies” [18–23]. The idea behind the many different hierarchies can be summed up quite simply with three central claims (Figure 1.1).

- 1 Randomized trials (RCTs), or systematic reviews of many randomized trials, generally offer stronger evidential support than observational studies.
- 2 Comparative clinical studies in general (including both RCTs and observational studies) offer stronger evidential support than “mechanistic” reasoning (“pathophysiologic rationale”) from more basic sciences.
- 3 Comparative clinical studies in general (including both RCTs and observational studies) offer stronger evidential support than expert clinical judgment.

Early EBM proponents showed that many widely used therapies that had been adopted based on “lower” forms of evidence proved to be useless or harmful when subjected to randomized trials. In a particularly dramatic (but not unique) example, antiarrhythmic drugs became widely used based on what was (believed to be) understood about the causes of sudden death after heart attack (“mechanistic reasoning”). However, a randomized trial



**Figure 1.1** Simplified EBM hierarchy of evidence (systematic reviews of all study types is assumed to be superior to single studies).

suggested that the drugs *increased* mortality, and had killed more people every year than died in action during the whole of the Vietnam War [24].

In spite of the compelling rationale, the EBM hierarchy leads to several paradoxes. The first is that many of the treatments in whose effectiveness we have the most confidence – that we consider most strongly supported by evidence – have never been supported by randomized trials of any description. These treatments include automatic external defibrillation to start a stopped heart, tracheostomy to open a blocked air passage, the Heimlich maneuver to dislodge an obstruction in the breathing passages, rabies vaccines, penicillin for the treatment of pneumonia, and epinephrine injections to treat severe anaphylactic shock. Meanwhile we often lack confidence in some treatments that are supported by evidence from higher up the hierarchy. The antidepressant Prozac, for instance, has proven superior to placebo in some double-blind RCTs, yet the effects of Prozac (over and above “placebo” effects) are hotly disputed [25–29]. Exploiting this irony, Gordon Smith and Jill Pell wrote a spoof article entitled “Parachute use to prevent death and major trauma related to gravitational challenge: a systematic review of randomised controlled trials” [30]. They concluded that:

Advocates of evidence-based medicine have criticised the adoption of interventions evaluated by using only observational [not RCT] data.

We think that everyone might benefit if the most radical protagonists of evidence based medicine organised and participated in a double blind, randomised, placebo controlled, crossover trial of the parachute.

Strictly speaking, this critique is unfair since the EBM movement has always acknowledged that treatments with dramatic effects do not require support from randomized trials [31–34]. Yet with one recent exception [19] current hierarchies have ignored this paradox: (systematic reviews of) randomized trials still feature categorically at the pinnacle of the EBM hierarchies.

The EBM rationale for the view that comparative clinical studies provide stronger evidential support than mechanistic reasoning and clinical expertise is also problematic. While EBM proponents have always acknowledged that mechanistic reasoning is important for generalizing (see Chapter 10), and that expertise should be integrated with external evidence (see Chapter 11), the view that comparative clinical studies provide stronger evidence for efficacy than mechanistic reasoning or clinical expertise is unsupported by a defensible rationale. A stubborn objector could always claim that the conclusions from mechanistic reasoning and expert judgment were more reliable than conclusions from randomized trials. This leads to the paradox that the EBM hierarchy itself appears to be supported by “weak” (according to EBM) evidence, namely the opinion of EBM experts!

These problems suggest that although EBM is compelling on many levels, a sustained analysis is wanting. While critics have attacked various aspects of the EBM methodology, the system as a whole has, with two notable but brief exceptions [4,5], escaped scrutiny. Indeed most critics have focused on the EBM view that randomized trials are less biased than non-randomized studies [30,35–48]. While I believe there is more to say about the relative value of randomized trials, this debate is, in many ways, independent of the EBM philosophy. Bayesian philosophers of science and statisticians have been debating the relative value of randomization [49,50] since long before the EBM movement was born. More importantly, these critiques of the EBM stance on randomized trials leave the central message of EBM – that comparative clinical studies *in general* provide better evidence than mechanistic reasoning and expert judgment – untouched.

To be sure, some philosophers [5,51–56] and medical professionals [57–61] have addressed the EBM stance on “mechanistic” reasoning from the basic sciences. Yet these critiques focus on the importance of mechanistic reasoning for generalizing the results of randomized trials, a view the EBM movement has accepted since the outset [12,32–34]. The EBM view that mechanistic reasoning is inferior to comparative clinical studies for establishing that a therapy has an average clinical effect in a study population (*efficacy*) has been altogether ignored.

Similarly, while some have flagged possible problems with the EBM stance on expert judgment [62–64], and EBM proponents have proposed a model for incorporating clinical expertise [65], there have been no sustained investigations of the EBM stance on expert judgment. Indeed little critical analysis of expert judgment has been undertaken at all since the late 1970s [66].

## 1.2 Scope of the book

EBM raises many compelling issues that are closely linked to the question of whether their theory of evidence is acceptable. These include the practical feasibility and uptake of EBM [64,67], the supposed hijacking of EBM methodology by special interests [64,67], EBM's relationship to alternative medicine [68–70], the ethical implications of EBM [1,46,71–74], whether it is possible to adapt EBM for social science and public policy [52,72,75], how EBM can be implemented [76,77] and various other social and historical aspects of EBM [78–81].

While this book will touch on these issues in various ways, I believe it is important to analyze the EBM methodology separately for two reasons. For one, as we shall see throughout the book, many of these other problems turn on issues of strict methodology [47,74,82]. Consider, for example, the relationship between EBM and research ethics. Some accuse the EBM movement of promoting randomized trials even when we “know” that an experimental therapy is effective [46,47,83] (see Chapters 4, 7, 8 and 11). But whether or not we already know that a treatment “works” depends, to a large measure, on whether we possess sufficient supporting evidence. This, in turn, relies on our account of what counts as sufficient evidence. Hence to attack EBM on ethical grounds is parasitic on an attack of the EBM philosophy of evidence. Likewise, the relationship between EBM and alternative medicine might depend on what counts as legitimate “placebo” controls (see Chapter 7).

Then, some of these other controversies are altogether independent of the EBM philosophy. A common critique, for example, is that EBM has been hijacked by special interests. Since randomized trials are expensive, potentially profitable (i.e. patentable) treatments will be more likely to be investigated in the first place [84]. These factors are important and influence the nature and quality of research produced (see Chapter 12). If the EBM movement is serious about producing the best evidence and improving patient outcomes, its proponents should engage more actively with the powerful forces involved in producing and disseminating evidence. At the same time, special interests will attempt to influence *any* (EBM or non-EBM) methodology. Imagine for the sake of argument that the EBM philosophy was violently rejected in favor of the view that palm reading

experts possessed the unassailable authority to decide whether an intervention had its putative effects. Special interests would then presumably focus on influencing palm reading experts, which could turn out to be far cheaper than conducting several large randomized trials. In brief, the problem that special interests corrupt medical research is a real problem independent of methodology. Once we address the corrupting sociological forces, we will still be left with the essential task of determining which methods most reliably detect an intervention's clinical effects.

One might argue, of course, that EBM is *particularly* prone to being hijacked by certain special interests. It is undoubtedly true, for example, that the EBM methodology is more easily used as a device to hold clinicians accountable than, say, a methodology insisting on the absolute authority of clinical experts. At the same time, if the EBM methodology is more reliable at detecting treatment effects – say it leads to saving many more lives – then the control over the medical profession allegedly attributable to EBM might be acceptable. Nobody complains that airline pilots are held accountable to a large number of rules and protocols because we believe that these rules save lives.

### 1.3 How the claims of EBM will be examined

Each of the three central claims of the EBM philosophy of evidence require distinct methods that I will outline separately in the relevant chapters. To summarize, I will evaluate the EBM claim that randomized trials offer superior evidential support to observational studies by appealing to the general rule that *good evidence rules out confounding factors*. Then, I will appeal to empirical evidence and analysis of relative strengths and weaknesses of mechanistic reasoning and expert judgment to evaluate the EBM claims that comparative clinical studies *generally* provide superior evidence to mechanistic reasoning and expert judgment. Contrary to what the EBM movement seem to concede, there is a strong justification for their position on the *evidential* roles of mechanistic reasoning and expert judgment.

However, there is one particular methodology that applies to the entire book and it is this: I will insist that all problems be stated clearly. With that in mind, I will spend the rest of Part I clarifying what EBM is, and what it means for a medical treatment to “work” in a clinically relevant sense. Failure to understand the nature of EBM and the nature of claims about treatment effects has led to much confusion in the critical literature.

### 1.4 Structure of what is to come

This book is divided into four parts. The remaining three chapters of the first part investigate what EBM is (Chapter 2) and how a claim that a treatment

“works” *should* be unpacked (Chapter 3). Part II is dedicated to analyzing the EBM claim that randomized trials provide stronger evidence than observational studies, and resolving the paradox that our most effective therapies are only supported by “lower-level” comparative clinical studies. After defending a method for evaluating the relative strength of comparative clinical studies (Chapter 4), I argue that the EBM position on randomized trials is, with a slight modification, sustainable (Chapter 5). The modification involves replacing categorical hierarchies with the requirement that comparative clinical studies should reveal an effect size that is greater than the combined effect of plausible confounders. In the next three chapters I evaluate the claims that double blinding (Chapter 6) and “placebo” controls (Chapters 7 and 8) enhance the quality of randomized trials. I then introduce Part III (Chapter 9), where I examine the EBM position on mechanistic reasoning (Chapter 10) and expert judgment (Chapter 11). I argue that mechanistic reasoning, while beleaguered with often unrecognized problems, should be admitted as evidence, perhaps alongside evidence from comparative clinical studies. Meanwhile, I defend the EBM view that expert judgment is not reliable as evidence, but that expertise plays several other important roles that deserve more emphasis in discussion in the EBM literature and practice. In the conclusion (Chapter 12) I summarize the findings then point out two new classes of methodological difficulties EBM faces in the near future.

A unifying theme of the book is that ethics and epistemology are intertwined. Randomized trials are unethical if we already have sufficient evidence from observational studies (Chapter 5) or mechanistic reasoning (Chapter 9), or if we *would have* sufficient evidence had we conducted a systematic review (Chapter 2). Likewise, the debate over “placebo” versus “active” controls (Chapter 7) has important ethical implications for the approval of trials, and using expert judgment *as evidence* (judgment is required for many other roles) could be unethical if it can be proven to be harmful (Chapter 10).

By the end of the book the reader will be able to evaluate the evidence for the EBM methodology and answer the question “What is the evidence for the EBM philosophy of evidence?”