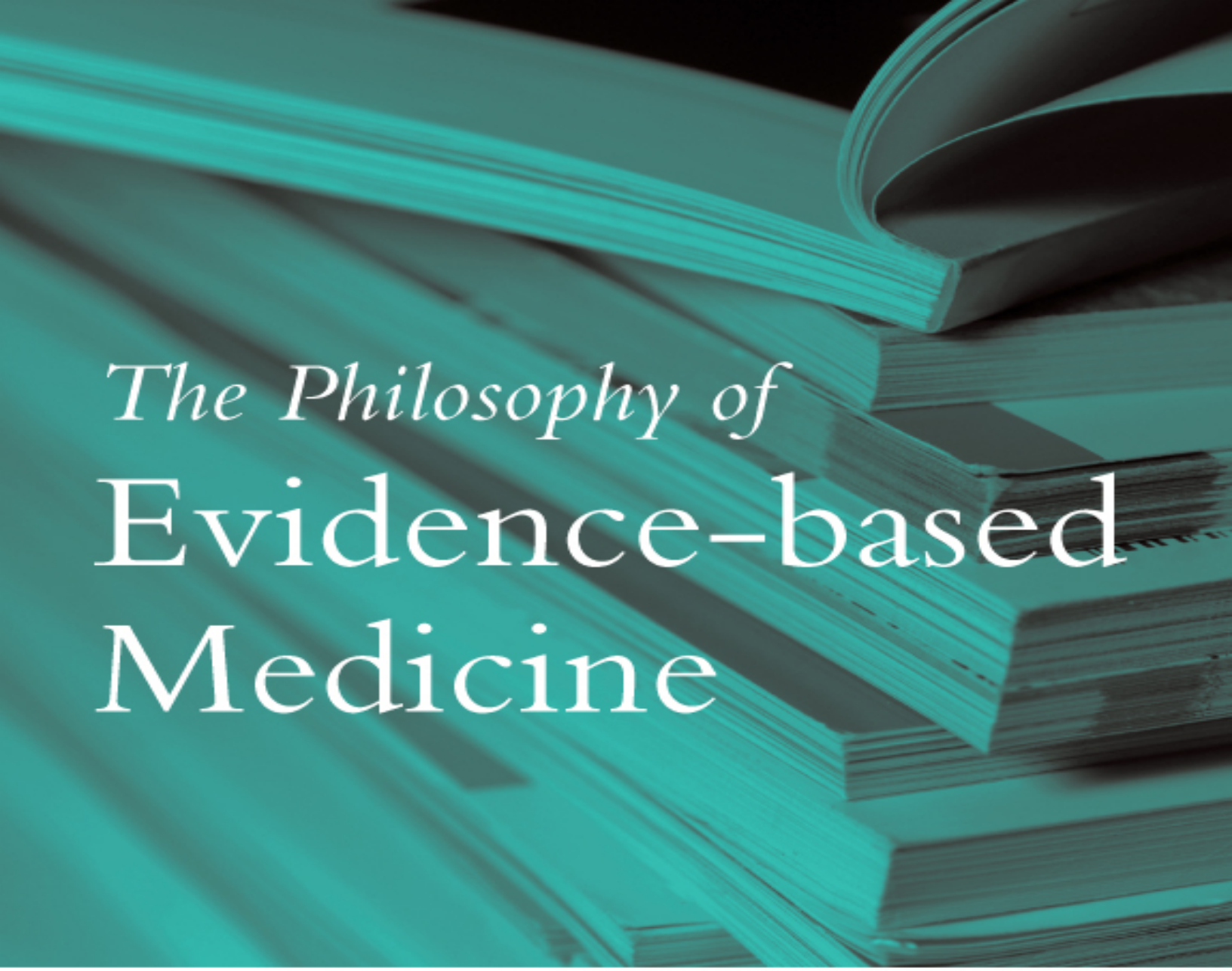


The Philosophy of Evidence-based Medicine

JEREMY HOWICK

With a foreword by Paul Glasziou

WILEY-BLACKWELL

BMJ Books

Contents

To question the foundations of a discipline or a practice is not necessarily to deny its value, but rather to stimulate a judicious and balanced appraisal of its merits.

—R. Ashcroft & R. ter Meulen [1]

It would be difficult to put the case for the clinical trial of new (or old) remedies more cogently or more clearly. The absence of such a trial in the past may well have led, to give one example, to the many years of inconclusive work on gold therapy in tuberculosis, while. . .the grave dangers of much earlier and drastic methods of therapeutics, such as blood-letting, purging, and starvation, would quickly have been exposed by comparative observations, impartially made.

—A.B. Hill & I.D. Hill [2]

The central problem of epistemology has always been and still is the problem of the growth of knowledge. And the growth of knowledge can be studied best by studying the growth of scientific knowledge.

—Karl Popper [3]

The Philosophy of Evidence-Based Medicine

Jeremy Howick
Centre for Evidence-Based Medicine
Department of Public Health and Primary Care
University of Oxford
Oxford, UK

and

Department of Science and Technology Studies
University College London
London, UK

Foreword by Paul Glasziou

WILEY-BLACKWELL
A John Wiley & Sons, Ltd., Publication

BMJ|Books

This edition first published 2011, © 2011 by Jeremy Howick

BMJ Books is an imprint of BMJ Publishing Group Limited, used under licence by Blackwell Publishing which was acquired by John Wiley & Sons in February 2007. Blackwell's publishing programme has been merged with Wiley's global Scientific, Technical and Medical business to form Wiley-Blackwell.

Registered office: John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

Editorial offices: 9600 Garsington Road, Oxford, OX4 2DQ, UK

The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

111 River Street, Hoboken, NJ 07030-5774, USA

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com/wiley-blackwell

product or vendor mentioned in this book. This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

The contents of this work are intended to further general scientific research, understanding, and discussion only and are not intended and should not be relied upon as recommending or promoting a specific method, diagnosis, or treatment by physicians for any particular patient. The publisher and the author make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation any implied warranties of fitness for a particular purpose. In view of ongoing research, equipment modifications, changes in governmental regulations, and the constant flow of information relating to the use of medicines, equipment, and devices, the reader is urged to review and evaluate the information provided in the package insert or instructions for each medicine, equipment, or device for, among other things, any changes in the instructions or indication of usage and for added warnings and precautions. Readers should consult with a specialist where appropriate. The fact that an organization or Website is referred to in this work as a citation and/or a potential source of further information does not mean that the author or the publisher endorses the information the organization or Website may provide or recommendations it may make. Further, readers should be aware that Internet Websites listed in this work may have changed or disappeared between when this work was written and when it is read. No warranty may be created or extended by any promotional statements for this work.

Library of Congress Cataloging-in-Publication Data

Howick, Jeremy.

The philosophy of evidence-based medicine / Jeremy Howick, with a foreword by Paul Glasziou.

p. ; cm.

Includes bibliographical references and index.

ISBN 978-1-4051-9667-3 (pbk. : alk. paper)

1. Evidence-based medicine. 2. Medicine—Philosophy. I. Title.

[DNLM: 1. Evidence-Based Medicine. 2. Philosophy, Medical. WB 102.5]

R723.7.H693 2011

610—dc22

2010047393

A catalogue record for this book is available from the British Library.

This book is published in the following electronic formats: ePDF 9781444342659; Wiley Online Library 9781444342673; ePub 9781444342666

1 2011

Acknowledgments

Many of the ideas in this book about evaluating the differences between randomized trials and observational studies are John Worrall's. I am also indebted to John for extensive feedback on several chapters in the book. I am equally indebted to Nancy Cartwright and Paul Glasziou. Nancy encouraged me to write the book, and without her feedback on my chapter on mechanistic reasoning it would have been a mess. Paul, together with Jeffrey Aronson, helped me conceptualize how to combine various types of evidence and were also encyclopaedic sources of fascinating medical examples. My collaboration with Paul and Jeffrey resulted in two publications.

Most of the book was written whilst I was a Medical Research Council (MRC) funded postdoctoral research fellow (G0800055) hosted by the Centre for Evidence-Based Medicine (CEBM) at Oxford's Department of Public Health and Primary Care (DPHPC). Many people at the CEBM and DPHPC gave me invaluable advice, including Paul Montgomery, Rafael Perera, Jason Oke, Richard Stevens, Merlin Wilcox, David Mant, Amanda Burls, Alison Ward, Carl Heneghan, Su-May Liew, and Olive Goddard. Thanks to Kay, Claire, Suzie, and Andrew for encouraging frequent tea and biscuit breaks. Sir Iain Chalmers of the James Lind Library was enormously generous with his time and enriched my work in many ways, especially by encouraging me to be succinct. On one occasion, he claimed I took 30 pages to express a message he could express in 3; I succeeded in reducing it to 15.

Concurrent with my postdoctoral fellowship at Oxford I have been teaching philosophy to medical students at University College, London (UCL). My students provided me with a platform to teach my ideas (Emily Sweetman pointed

out several typos in earlier drafts), and many of my colleagues at UCL gave me useful feedback. These include Steve Miller, Josipa Petrunic, Brian Balmer, Jon Agar, Norma Morris, and especially Donald Gillies.

Many other people have listened or read the ideas in these chapters at various conferences, meetings, or email exchanges. Space doesn't permit me to mention all of them, but Gordon Guyatt, Brian Haynes, Dave Sackett, Murray and Eleanor Enkin, Alejandro Jadad, Ted Kaptchuk, Elizabeth Silver, Eileen Munro, Jon Williamson, Federica Russo, and Beatrice Golomb (many of the ideas in the conclusion are borrowed from Dr Golomb) were all helpful.

I would not have been able to write the book without (some) personal distractions, usually in the form of rowing. Thanks to Jonny Searle for making sure I woke up at 5.30 on Saturday mornings in the dark of the winter, to Paul Kelly for encouraging me to join the Wolf Pack, Colin Smith for being a recent training partner, and especially to the Great Eight (Phil, Fred, Nick, Ted, John, Sohier, Dave, Hirsh and most of all Scott Armstrong) for continuing to remind me that the world is the oyster of the bold.

Various friends, including Foad Dizadji-Bahmani, Jesse Elzinga, and Josh West, put up with me for various periods of time while I moved flats. Other friends, including Sebastien, Mark, Qarim, and Renaud, were always there when I needed them. Dick Fishlock has been a continued source of entertainment and inspiration, and Dusan gave me a break. Lastly, my mother (who I'm becoming increasingly convinced is an Angel), father (who generously supported my expensive education), sisters, Brett, John, Jack and Raven provided unconditional love that helped me keep going.

Stephen Hlophé, Mingy, and Dr Bali have helped keep me almost sane. Thank you to everyone I forgot to mention.

Foreword

In 1991 an international group formed to encourage clinicians to consider results of recent research when treating patients. They commenced writing a series of User's Guides to reading research for JAMA, the Journal of the American Medical Association, and needed a new term to signal the intention of the series. After several suggestions, the group's leader, Gordon Guyatt, proposed the term Evidence-Based Medicine. The new term was to ignite a movement that spread rapidly around the world. The methods of evidence-based medicine (EBM) have evolved since then, but the focus of the inventors – mostly clinicians – was the practical concern of bedside decision making. Understandably they paid less attention to the psychology, sociology or philosophy that might underpin EBM. However, now that EBM is well established in the medical world, deeper exploration by different disciplines seems warranted. This book is an examination and extension of the philosophy of EBM: a modern conversation between Aristotle and Hippocrates.

While the term Evidence-Based Medicine has a short history dating back to the 1990s, the ideas behind it have been evolving for centuries. A large part of the vocabulary of EBM – bias, confounding, randomization, placebo, confidence interval, etc. – has been invented and developed by statisticians and epidemiologists. But philosophers have been grappling with many of the same issues that lie behind the ideas, including the nature, and proof for, causal relationships, justification for induction, and errors in human observation, models, and reasoning. Many of these terms appear and are explained inside. Other ideas less familiar in the routine EBM books also enrich this text; for example, Phillip's paradox, nocebo effects, and probabilistic causality.

The book is a rich treasure of examples. Some are akin to zen koans: thinking about them can be a struggle but considerably deepen understanding of EBM. Consider the randomized comparison of nicotine versus placebo but where both groups were also randomised to be told they received nicotine, received placebo, or not told anything (see Figure 8.4) – a 2×3 factorial design. What do the various possible comparisons tell you? Which is better: to have the nicotine patch but be told it is placebo, or have the placebo patch but be told you received nicotine? Considering these comparisons may change the way you think about the placebo effect and the place of placebos in trials.

EBMers have been focused on teaching, and getting the evidence in practice. However, less attention has been given to the philosophical roots of EBM. In particular, we have ignored or belittled the role of mechanism. The battle between mechanists and empiricists is long standing in both philosophy and medicine, but what have the two opposing ideas to offer each other, to researchers, and to the users of research? Chapter 10 is an excellent synthesis of both camps. This chapter is a crystallization of many long afternoons of stimulating discussion between the author, Jeffrey Aronson and myself. Besides the many insights developed in those conversations and set down here, I also learned the value of having the input and insight of other disciplines on the work of EBM. And had fun in the process. The challenge of working across disciplines though is great: basic assumptions are different, purposes are different, and even the vocabulary can be different. “Proof” means different things to philosophers, doctors, detectives and distillers. But with a generous dose of good-will, we found the interdisciplinary exploration fruitful for both philosophy and medicine. And worth continuing.

This work represents an important dialogue between EBM and philosophers of science. There has been too little. I searched MEDLINE for titles which include EBM and philosophy and found only six, but all from the last 6 years. Let me end with a quote from the earliest of these articles: Ashcroft and Ter Meulen introduce a special issue of the Journal of Medical Ethics that reported on a symposium on EBM by saying: “To question the foundations of a discipline or a practice is not necessarily to deny its value, but rather to stimulate a judicious and balanced appraisal of its merits; we offer the present selection of papers in that spirit.” So I hope you enjoy and learn from reading this, and seek out your local philosopher for a cup of tea or a pint of ale, and some stimulating discussion.

Professor Paul Glasziou PhD FRACGP MRCGP

Director, Department of Evidence-Based Medicine

University of Oxford, Oxford, UK

Preface

Most EBM “hierarchies” of evidence rank comparative clinical studies (including systematic reviews of randomized trials) above mechanistic reasoning (“pathophysiologic rationale”) and expert judgment. Within comparative clinical studies, randomized trials are considered to offer stronger evidence than observational studies. 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 evaluation by randomized trials. In spite of the compelling rationale, the EBM philosophy of evidence leads to several paradoxes. Perhaps the most striking is that many of the treatments in whose effectiveness we have the most confidence – that we consider to be 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, and the Heimlich maneuver to dislodge airway obstructions. While critics have attacked various aspects of the EBM methodology, the system as a whole has, with few exceptions, escaped scrutiny. After outlining the paradoxes (Chapter 1), I investigate what EBM is (Chapter 2), and how a claim that a treatment “works” *should* be unpacked (Chapter 3). Next, I defend a method for evaluating the relative strength of comparative clinical studies (Chapter 4), and I argue that the EBM position on randomized trials is, with a slight modification, sustainable (Chapter 5). The modification is to replace categorical hierarchies that place randomized trials on top with the requirement that comparative clinical studies should reveal an effect size that outweighs 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

comparative clinical studies. I then examine the EBM position on mechanistic reasoning and expert judgment (Chapters 9–11). 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 serious discussion in the EBM literature. My conclusion (Chapter 12) is that strict hierarchies should be replaced by the requirement that all evidence of sufficiently high quality should be admitted as evidential support, and that the various non-evidential roles of expertise deserve more discussion in the EBM literature.

PART I

Introduction

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.

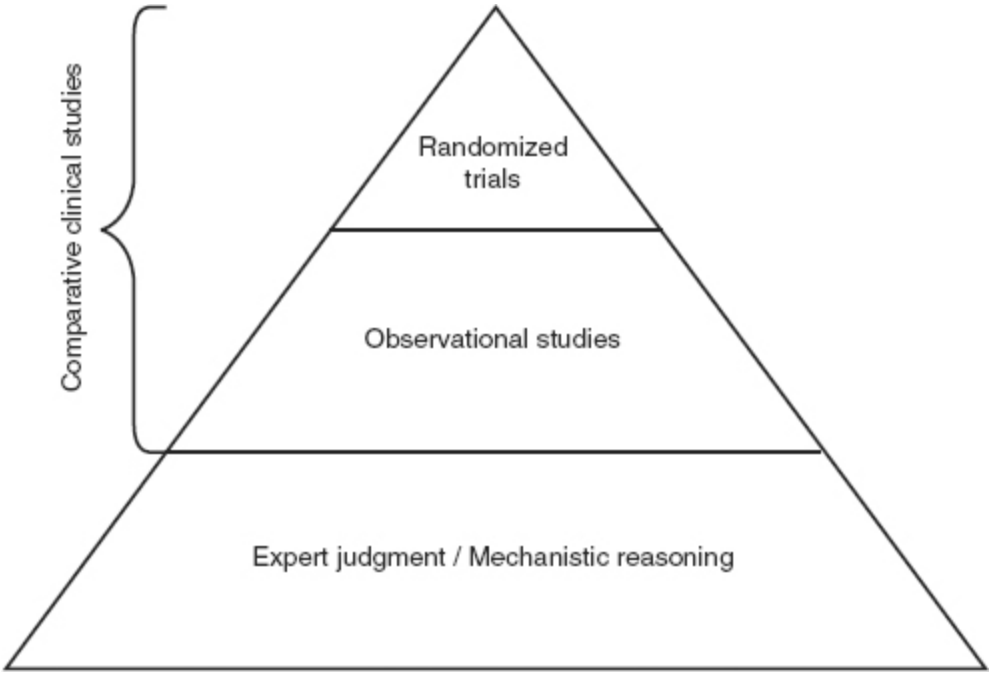
With a rhetorical *tour de force*, EBM was introduced as a "new paradigm" in the early 1990s [10–12]. Less than two decades later, there are at least seven journals, a dozen books, thousands of new citations to EBM each year, and a growing number of international research centres dedicated to the practice, 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).

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



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 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