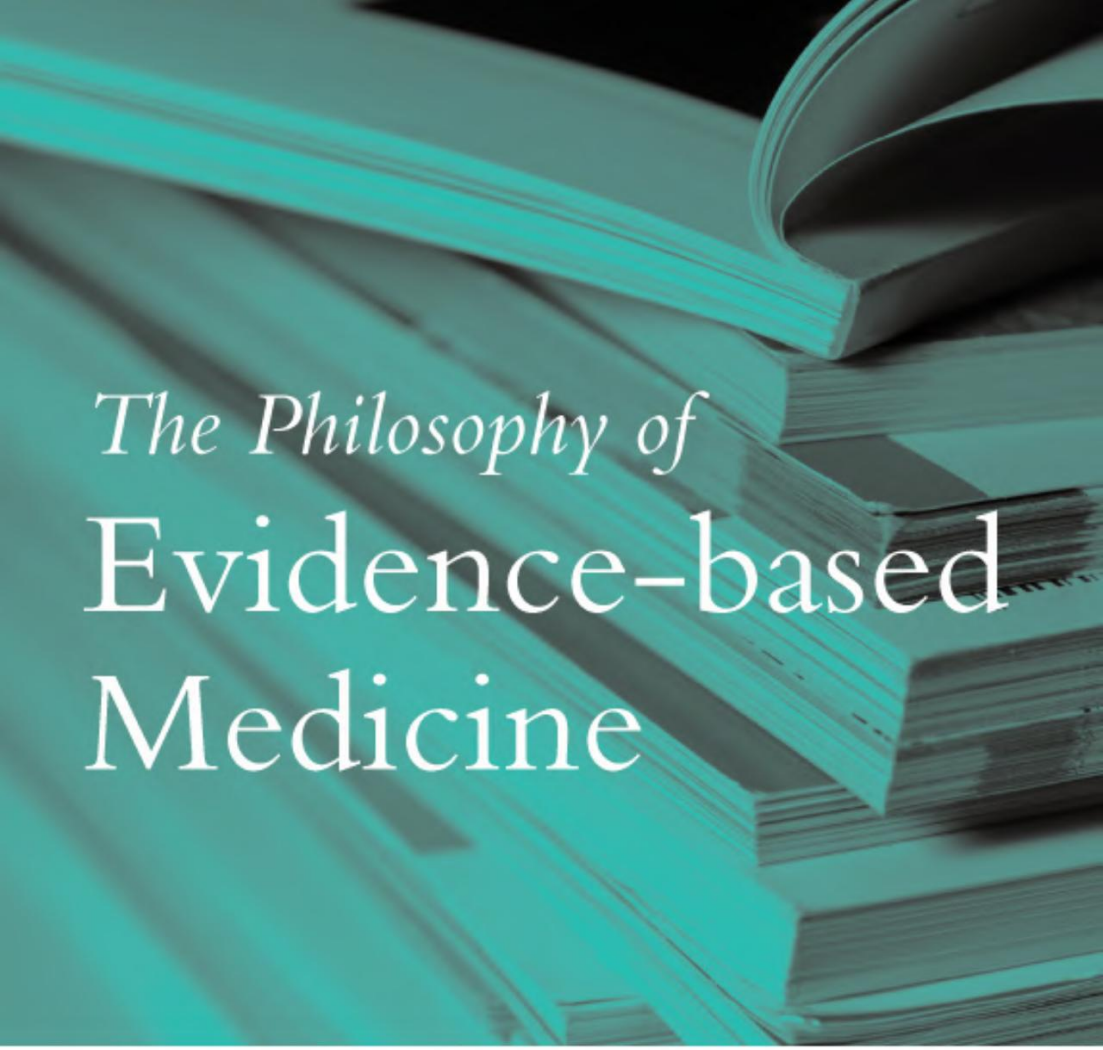

The Philosophy of Evidence-based Medicine

JEREMY HOWICK

With a foreword by Paul Glasziou

WILEY-BLACKWELL

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

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

Set in 9.5/12pt Meridien by MPS Limited, a Macmillan Company, Chennai, India

1 2011

Contents

Acknowledgments, ix

Foreword, xi

Preface, xiii

Part I: Introduction

1 The philosophy of evidence-based medicine, 3

1.1 What on earth was medicine based on before evidence-based medicine?, 3

1.2 Scope of the book, 7

1.3 How the claims of EBM will be examined, 8

1.4 Structure of what is to come, 8

2 What is EBM?, 10

2.1 EBM as a self-proclaimed Kuhnian paradigm, 10

2.2 The motivation for the birth of EBM: a sketch, 11

2.3 Original definition of EBM, 15

2.4 Reaction to criticism of the EBM system of evidence: more subtle, more or less the same, 20

3 What is good evidence for a clinical decision?, 24

3.1 Introduction, 24

3.2 Evidence *for* clinical effectiveness, 25

3.3 Strong evidence tells us what?, 29

Part II: Do randomization, double masking, and placebo controls rule out more confounding factors than their alternatives?

4 Ruling out plausible rival hypotheses and confounding factors: a method, 33

5 Resolving the paradox of effectiveness: when do observational studies offer the same degree of evidential support as randomized trials?, 39

5.1 The paradox of effectiveness, 39

5.2 Observational studies: definition and problems, 40

5.3 Randomized trials to the rescue, 43

5.4 Defending the EBM view that randomized trials provide better evidence than observational studies, 48

5.5 Overcoming the paradox of effectiveness, 56

5.6 Conclusion: a more subtle way to distinguish between high- and low-quality comparative clinical studies, 59

Appendix 1: types of restricted randomization, 60

Appendix 2: Worrall's arguments that randomization is required for classical hypothesis testing and establishing probabilistic causes, 61

6 Questioning double blinding as a universal methodological virtue of clinical trials: resolving the Philip's paradox, 63

6.1 The problems with double masking as a requirement for clinical trial validity, 63

6.2 The many faces of double masking: clarifying the terminology, 65

6.3 Confounders that arise from participant and caregiver knowledge, 66

6.4 The importance of *successful* double masking, 69

6.5 One (and a half) solutions to the Philip's paradox, 75

6.6 The full solution to the Philip's paradox: challenging the view that double masking rules out confounding factors when treatments are evidently dramatic, 76

6.7 Double masking is valuable unless the treatment effects are evidently dramatic, hence the Philip's paradox does not arise, 78

7 Placebo controls: problematic and misleading baseline measures of effectiveness, 80

7.1 The need to control the placebo, 80

7.2 Legitimate placebo controls, 82

7.3 How placebo controls often violate the first condition for legitimacy, 83

7.4 How placebo controls often violate the second condition for legitimacy, 86

7.5 Special problem for constructing placebos for complex treatments: case studies of exercise and acupuncture, 87

7.6 Summary and solution to the problem with illegitimate placebo controls, 94

8 Questioning the methodological superiority of "placebo" over "active" controlled trials, 96

8.1 Epistemological foundations of the ethical debate over the use of placebo-controlled trials, 96

8.2 Problems with the assay sensitivity arguments against ACTs, 98

8.3 Problems with the first assay sensitivity argument against ACTs, 98

8.4 The second assay sensitivity argument, 106

8.5 Challenging the view that PCTs provide a measure of absolute effect size, 107

8.6 Questioning the claim that PCTs require smaller sample sizes, 112

8.7 Conclusion: a reassessment of the relative methodological quality of PCTs, 113

Appendix: more detailed explanation of why the second assay sensitivity argument fails, 114

Part III: Examining the paradox that traditional roles for mechanistic reasoning and expert judgment have been up-ended by EBM

9 Transition to Part III, 119

9.1 Summary of Part II, 119

9.2 Introduction to Part III, 120

10 A qualified defence of the EBM stance on mechanistic reasoning, 122

10.1 A tension between proponents of mechanistic reasoning and EBM views, 122

10.2 Clarifying the terminology: comparative clinical studies, mechanisms, and mechanistic reasoning, 124

10.3 Why the strong view that mechanistic reasoning is necessary to establish causal claims is mistaken, 130

10.4 Two epistemological problems with mechanistic reasoning, 136

10.5 Why EBM proponents should allow a more prominent role for high-quality (valid and based on "complete" mechanisms) mechanistic reasoning in their evidence hierarchies, 143

10.6 Mechanisms and other roles in clinical medicine, 146

10.7 Recommending a (slightly) more important role for mechanistic reasoning in the EBM system, 153

Appendix: cases where mechanistic reasoning led to the adoption of therapies that were either useless or harmful according to well-conducted clinical research, 154

11 Knowledge *that* versus knowledge *how*: situating the EBM position on expert clinical judgment, 158

11.1 Controversies surrounding the EBM stance on expert clinical judgement, 158

11.2 General clinical judgment belongs at the bottom of (or off) the hierarchy of evidence, 161

11.3 Individual clinical judgment also belongs at the bottom of the hierarchy, 167

11.4 The equally important non-evidential roles of expertise, 176

11.5 Conclusion, 182

Part IV: Conclusions

12 Moving EBM forward, 187

12.1 Summary of findings: the EBM philosophy is acceptable, but. . ., 187

12.2 Two new frontiers for EBM, 189

References, 193

Index, 225

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 2x3 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

The Philosophy of Evidence-Based Medicine, First Edition. Jeremy Howick.

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

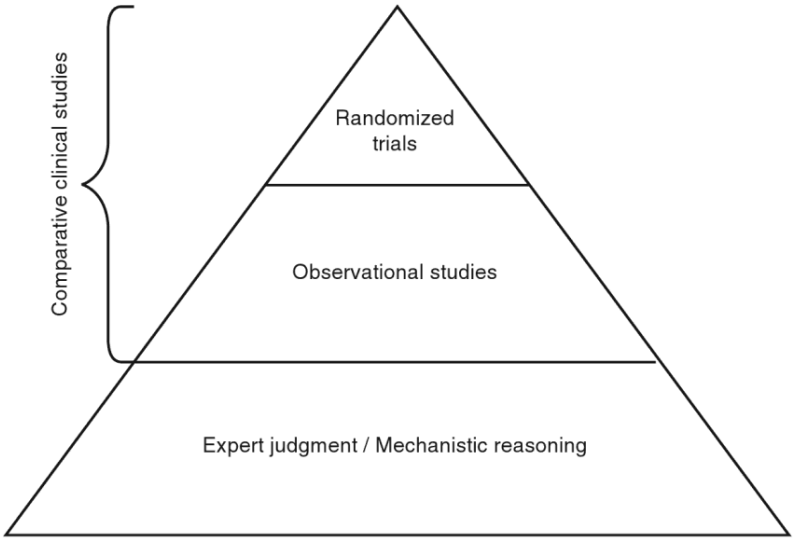

Figure 1.1 Simplified EBM hierarchy of evidence (systematic reviews of all study types is assumed to be superior to single studies).

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.

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?"

CHAPTER 2
What is EBM?

If you can believe fervently in your treatment, even though controlled tests show that it is quite useless, then your results are much better, your patients are much better, and your income is much better too. I believe this accounts for the remarkable success of some of the less gifted, but more credulous members of our profession, and also for the violent dislike of statistics and controlled tests which fashionable and successful doctors are accustomed to display.

—R. Asher [85]

The history of medicine shows many examples of forms of treatment widely considered as effective on grounds of clinical impression which have turned out to be ineffective or even harmful.

—A.B. Hill & I.D. Hill [2]

2.1 EBM as a self-proclaimed Kuhnian paradigm

The title of the paper that announced EBM to the wider community was "Evidence-based medicine: a *new* approach to teaching the practice of medicine" (my emphasis) [12]. The very first sentence of the paper reads: "A *new* paradigm for medical practice is emerging" (my emphasis) [12].

The question of whether EBM is truly new is a historical one [6,8,9,86]. While I shall provide some background to the EBM movement and recount some amusing anecdotes about early EBM advocates, a comprehensive historical analysis of the origins and genesis of EBM lies beyond the scope of this work (see Tröhler [87] for a good review of the recent historical roots of EBM). Similarly, the question of whether EBM is truly a

The Philosophy of Evidence-Based Medicine, First Edition. Jeremy Howick.

new (Kuhnian) paradigm would involve an analysis of whether Kuhnian paradigms are applicable to methodological innovations in medicine [88,89] which would take us far afield. Moreover, both questions – whether EBM is new and whether EBM is a new Kuhnian paradigm – require that we establish what EBM actually *is*. This is no straightforward task given the evolving definitions of the movement [90–93].

In this brief chapter I will contend that in spite of evolving characterizations, the EBM view that comparative clinical studies, preferably (systematic reviews of) randomized trials, provide more telling evidence for therapeutic effects than mechanistic reasoning and clinical expertise has remained constant.

I will start with a sketch of the factors that contributed to the birth of EBM movement. Then, I will review the evolving definitions of EBM and argue that its fundamental view of what counts as good *evidence* has not changed. For now, I will leave an evaluation and justifications of the EBM definition of "good" evidence for later chapters: here I will focus on charitably interpreting what the EBM system of evidence *is*.

2.2 The motivation for the birth of EBM: a sketch

The 100-year period between 1885 and 1985 brought amazing medical breakthroughs. The dramatic discovery of the rabies vaccine put an end to fear of rabid dogs, the discovery of penicillin and streptomycin suggested that infectious disease would soon be altogether eradicated, and cure for most childhood cancer was a promising sign that all cancers would soon disappear. Meanwhile, open heart surgery, hip replacements and kidney transplants indicated that we could dramatically extend our lifespans by replacing our "used" parts, and *in vitro* fertilization put an end to the misery caused by infertility [94]. Understanding the underlying *mechanisms* of health and disease appeared to drive many of these discoveries. It is difficult to see how, for example, the very idea for a rabies vaccine would have arisen without the germ theory of disease, and kidney transplants would not have been possible without understanding the immune system. The method of investigating the underlying mechanisms of disease appeared to be working well. In the century beginning in 1885 infant mortality in the USA and Europe dropped from 140 per 1000 to 5 per 1000, and life expectancy rose from under 50 to almost 80 years. It was not unreasonable to suppose, in the middle of the 20th century, that medicine would continue to advance at a furious pace and that quite soon most human suffering would all but disappear. Indeed in a 1949 article Lord Horder claimed just that: "Whither Medicine?" he asked, "Whither else than straight ahead" [95].

Eventually, however, reality set in. Infectious diseases proved more resistant than was initially envisaged, many cancers proved to be formidable

> This clinical impression was very sobering. It made me wonder whether what I had been taught at medical school might have been lethally wrong, at least in the circumstances in which I was working, and precipitated a now incurable "scepticemia" about authoritarian therapeutic prescriptions and prescriptions unsupported by trustworthy empirical evidence [98].

Meanwhile, Dave Sackett, who was the main author on many early texts on clinical epidemiology and EBM, became unpopular even as a medical student in the 1950s for questioning the apparent wisdom of his more senior colleagues.

> I was a final-year medical student on a medical ward, where a teenager with "infectious hepatitis" (now called Type A hepatitis) was admitted to my care. He presented with severe malaise, an enlarged and tender liver, and a colorful demonstration of deranged bilirubin metabolism that made me the envy of my fellow clerks. However, after a few days of total bed rest his spirits and energy returned and he asked me to let him get up and around.
>
> In the 1950s, everybody "knew" that such patients, if they were to avoid permanent liver damage, must be kept at bed rest until their enlarged liver receded and their bilirubin and enzymes returned to normal. And if, after getting up and around, their enzymes rose again, back to bed they went. This conventional wisdom formed the basis for daily confrontations between an increasingly restless and resentful patient and an increasingly adamant and doom-predicting clinical clerk.
>
> We clinical clerks were expected to read material relevant to the care of our patients. I wanted to understand (for both of us) how letting him out of bed would exacerbate his pathophysiology. After exhausting several unhelpful texts, I turned to the journals. PubMed was decades away, and the National Library of Medicine hadn't yet begun to help the Armed Forces Medical Library with its *Current List of the Medical Literature*. Nonetheless, it directed me to a citation in the *Journal of Clinical Investigation* (back in the days when it was a real clinical journal) for: "The treatment of acute infectious hepatitis. Controlled studies of the effects of diet, rest, and physical reconditioning on the acute course of the disease and on the incidence of relapses and residual abnormalities." (Chalmers et al. 1955). Reading this paper not only changed my treatment plan for my patient. It forever changed my attitude toward conventional wisdom,

clinical decisions be based on "best" evidence reached critical mass at McMaster University in Hamilton, Ontario, Canada. The group at McMaster, which included Dave Sackett, Gordon Guyatt, Brian Haynes, and Peter Tugwell, began to use the terms "clinical epidemiology" [100,101] and "critical appraisal" to describe their new approach to medicine. In 1990, Gordon Guyatt assumed the position of Residency Director of the Internal Medicine Program at McMaster, where he was charged with several tasks including justifying the innovative approach to medicine and advertising to prospective medical students. In the spring of 1990, Guyatt presented plans for changing the curriculum to the members of the Department of Medicine, many of whom were unsympathetic. Guyatt initially suggested describing the new approach as "scientific medicine." Those already hostile apparently became incensed at the implication that they had previously been "unscientific." Guyatt's second try at a name for McMaster's philosophy, "evidence-based medicine," turned out to be a catchy one. The term initially appeared in an information document aimed at prospective or new students in the autumn of 1990. The relevant passage was:

> Residents are taught to develop an attitude of "enlightened scepticism" towards the application of diagnostic, therapeutic, and prognostic technologies in their day-to-day management of patients. This approach, which has been called "evidence-based medicine". . . The goal is to be aware of the evidence on which one's practice is based, the soundness of the evidence, and the strength of inference the evidence permits. The strategy employed requires a clear delineation of the relevant question(s); a thorough search of the literature relating to the questions; a critical appraisal of the evidence, and its applicability to the clinical situation; a balanced application of the conclusions to the clinical problem [102].

2.3 Original definition of EBM

EBM was initially defined as follows:

> Evidence-based medicine de-emphasizes intuition, unsystematic clinical experience, and pathophysiological rationale as sufficient grounds for clinical decision making and stresses the examination of evidence from clinical research [12].

The terms "clinical experience," "pathophysiologic rationale" and "clinical research" require some clarification here – much more will be said in upcoming chapters.

537 Jorgensen AW, Hilden J, Gøtzsche PC. Cochrane reviews compared with industry supported meta-analyses and other meta-analyses of the same drugs: systematic review. *British Medical Journal* 2006;333:782.
538 Leopold SS, Warme WJ, Fritz Braunlich E, Shott S. Association between funding source and study outcome in orthopaedic research. *Clinical Orthopaedics and Related Research* 2003;(415):293–301.
539 Lexchin J, Bero LA, Djulbegovic B, Clark O. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *British Medical Journal* 2003;326:1167–70.
540 Yaphe J, Edman R, Knishkowy B, Herman J. The association between funding by commercial interests and study outcome in randomized controlled drug trials. *Family Practice* 2001;18:565–8.
541 Heres S, Davis J, Maino K, Jetzinger E, Kissling W, Leucht S. Why olanzapine beats risperidone, risperidone beats quetiapine, and quetiapine beats olanzapine: an exploratory analysis of head-to-head comparison studies of second-generation antipsychotics. *American Journal of Psychiatry* 2006;163:185–94.
542 Smith R. Medical journals are an extension of the marketing arm of pharmaceutical companies. *PLoS Med* 2005;2(5):e138.
543 Hopewell S, Loudon K, Clarke MJ, Oxman AD, Dickersin K. Publication bias in clinical trials due to statistical significance or direction of trial results. *Cochrane Database of Systematic Reviews* 2009;(1):MR000006.
544 Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *New England Journal of Medicine* 2008;358:252–60.
545 De Angelis C, Drazen JM, Frizelle FA *et al.* Clinical trial registration: a statement from the International Committee of Medical Journal Editors. *New England Journal of Medicine* 2004;351:1250–1.
546 Tramer MR, Reynolds DJ, Moore RA, McQuay HJ. Impact of covert duplicate publication on meta-analysis: a case study. *British Medical Journal* 1997;315:635–40.
547 Gøtzsche P. Bias in double-blind trials [thesis]. *Danish Medical Bulletin* 1990;37:329–36.
548 Gøtzsche PC. Methodology and overt and hidden bias in reports of 196 double-blind trials of nonsteroidal antiinflammatory drugs in rheumatoid arthritis. *Controlled Clinical Trials* 1989;10:31–56.
549 Gøtzsche PC. Blinding during data analysis and writing of manuscripts. *Controlled Clinical Trials* 1996;17:285–90; discussion 90–3.
550 Stelfox HT, Chua G, O'Rourke K, Detsky AS. Conflict of interest in the debate over calcium-channel antagonists. *New England Journal of Medicine* 1998;338:101–6.
551 Wilkinson R, Marmot M (eds) *Social Determinants of Health: The Solid Facts*, 2nd edn. Copenhagen: World Health Organization, 2003.