What can philosophy tell us about Evidence-Based Medicine? An assessment of Jeremy Howick’s *The Philosophy of Evidence-based Medicine*

While few doctors regard philosophical reflection on their activities as an important part of their practice, it is becoming more widely recognised that the humanities have a potentially significant role to play in medical education and research. Philosophy has the function of placing the activities and claims of researchers and clinicians into a wider critical context. Philosophy can challenge, expose, and examine the foundations and assumptions behind those activities. It can clarify questions and untangle the confusions surrounding disputes that seem not to be resolvable along the normal lines. Of the latter, the most significant and most obviously amenable to philosophical inquiry is the continuing debate over Evidence-Based Medicine (EBM). Unlike most disputes over scientific facts, which are typically (although not always) resolved once sufficient evidence has been gathered, the EBM debate does not appear to be one in which more evidence would necessarily help us see more clearly. And that is not least because the very nature and import of the evidence is one of the things that is at stake in the debate. The fact that the debate is one where the very mechanisms of evaluation are disputed lends some support to the claims of the EBM movement that it constitutes a new paradigm (Evidence-Based Medicine Working Group 1992). For a dispute over paradigms is precisely the kind of dispute where there is lacking a mechanism for evaluating and resolving the dispute. Also indicative of a dispute over paradigms is the presence of incommensurability. Incommensurability is the lack of a shared, exact measure for describing and assessing competing theories. It is often manifested in its appearing as if the various contenders are talking past each other, not engaging fully with each other’s arguments. And certainly opponents and opponents of EBM both seem to find that the other ‘side’ misunderstands their point of view and misinterprets their arguments.

The claim that EBM is a new paradigm does not in itself amount to an endorsement of EBM, for not all would-be new paradigms are advances (we may think of mesmerism, vitalism, dianetics among other such cases in the history of medicine). The fact that EBM has been widely adopted may suggest that it is widely perceived as responding to problems and anomalies of the pre-EBM period. On the other hand, there is the view that EBM is itself an expression of power, a means of control (a ‘micro-fascism’), and so its success is not a sign of its contributing to progress but rather of the effectiveness of its political strategy (allegedly aligned with the interests of capitalism) (Holmes et al. 2006). There has been philosophical discussion of EBM, although perhaps not as much as the intrinsic interest of EBM and the debates surrounding it would warrant. That maybe reflects the relative youth of philosophy of
medicine as an area of inquiry. What philosophical engagement there has been has come from both of the camps or traditions in philosophy, labelled, for largely historical reasons ‘continental’ and ‘analytic’. One example of a continental approach is the claim referred to above, that EBM is a micro-fascism, focussing on a deconstruction of EBM, concerned with its rhetoric and the power-relations in which its proponents are engaged. The concerns of analytic philosophy have been rather different. Since EBM makes claims about the nature of evidence in medicine and the quality of clinical judgment it is ripe for epistemological analysis—the careful investigation of the efficacy of the various sources of evidence encompassed by EBM (RCT, observational study, clinical expertise etc.) and evaluation of the arguments put forward by proponents of EBM regarding the relative merits of these. This analytic approach is clearly very different from the continental approach, and is indeed in many ways antithetical to it. For the analytical philosopher the intrinsic merit of an argument or viewpoint is independent of the rhetoric with which it is expressed or the political motivations of those putting it forward. Despite such differences, it is not however the case that analytic philosophers are all sympathetic to EBM. John Worrall is a well-known analytic philosopher of science who has been highly critical of what he sees as EBM’s over-emphasis on randomized controlled trials as the only satisfactory way of generating evidence in medicine (e.g. Worrall 2007). Among analytic philosophers there is a diversity of opinion on topics raised by EBM. As befits the nature of analytic philosophy, with its emphasis on precision and the analysis of particular claims and arguments, discussions tend to be focussed: there is less discussion of whether EBM is a good thing or bad; more engagement with questions such as ‘what is the epistemic function of randomization?’ or ‘can placebo controlled trials provide evidence of “absolute” effect size?’.

Jeremy Howick’s *The Philosophy of Evidence-Based Medicine* (2011) is the first book-length philosophical treatment EBM. The book is divided into four sections. The first sets the scene: what is EBM? Where does it come from? Why do we need a philosophy of evidence-based medicine? This is important. First, philosophers reading this book will need to know what its topic is. Non-medics on first hearing the phrase ‘evidence-based medicine’ are sometimes perplexed—what, they ask, might medicine be if it is not evidence-based? Perhaps the choice of this phrase reflects, some critics might suggest, the clever rhetoric of the EBM movement. Nonetheless, it is useful to be reminded of a period when clinical judgment and practice were rather less well evidenced than they are today. Medicine before the war was supported by an evidence base that is a tiny fraction of what is now available to doctors. Furthermore, the nature of that evidence was made up of individual clinical experience plus expert opinion, supplemented by reasoning based on knowledge of physiology and pathology. Expert opinion was itself a matter of the individual experience of specialists and other clinicians shared via the pages of medical journals. There was some evidence from clinical studies. The latter were almost all observational studies rather than controlled clinical trials. Rather more worrying, and a particular motivation for the nascent EBM movement was the more recent period, roughly the final three decades of the twentieth century, when much more, and better quality evidence was available, including evidence from controlled clinical trials, as well as systematic reviews and the like. Despite the growing wealth of evidence, the older habits died hard, and clinical expertise remained at the centre of both decision making and education in medicine. The consensus development conference was an advance, but only in scale, not in kind. For the consensus conference was typically a device for amalgamating the experience of many experts and was not es-
especially responsive to new kinds of evidence. Indeed, as Howick points out, such conferences often ignored or were at odds with the results of randomized controlled trials. This background is important since it reminds us why David Sackett, Gordon Guyatt, Anna Donald and colleagues promoted EBM. In so doing it points to what is central to EBM—assessing the quality of different kinds of evidence and promoting the use of the better kinds, in particular the controlled clinical trial in contrast to expert opinion. This thereby also allows us to see whether criticisms of EBM are aimed at its central concerns or engage with incidental or peripheral features of EBM.

Between the introductory part and a short concluding part, parts II and III are the heart of *The Philosophy of Evidence-based Medicine*; part II is comprised of chapters on assessing the epistemological merits of the principal features of a randomized controlled trial: randomization, double masking (blinding), and placebo controls; part III contains discussions of EBM’s relative downplaying of ‘mechanistic’ (pathophysiological) reasoning and of expert judgment.

Howick's strategy in part II is this: he starts with the well-known problems of observational studies, that they are subject to confounding, and then he moves on to ask whether the pride of place given by EBM to randomized, double-blinded, placebo-controlled clinical trials is justified by a superior ability to avoid confounding. This is the correct approach to assessing the claims of EBM. A crucial role that evidence plays in establishing the truth of a hypothesis (e.g. that an observed correlation between A and B arises because A causes B) is by ruling out hypotheses that rival the causal hypothesis. Hypotheses of confounding, e.g. that there is some common cause of both A and B, are the principal rival hypotheses needing to be eliminated. Howick considers the various components of a randomized, double-blinded, placebo-controlled clinical trial to see what they contribute to this end. Of these, randomization is the component that has hitherto been most discussed in the philosophical literature, mainly for reasons that are likely to interest philosophers more than doctors. While Howick does discuss randomization, he wisely avoids the more technical debates. The benefit provided by randomization is straightforward: it is the reason why Sir Austin Bradford Hill other others promoted randomization, viz. that randomization excludes confounding due to allocation bias. Howick deals with other criticisms of randomization. In some cases these erect a straw man, e.g. the claim that randomization can eliminate all potential confounders (this is not part of a serious EBM case for RCTs). In other cases, the criticisms are equally valid against observational studies (e.g. *some* of the criticisms that RCTs have low external validity).

Howick does lay greater emphasis on ‘Philip’s paradox’ (the ‘paradox of effectiveness’): if the ‘best evidence’ is gained from studies that are double-blind (or ‘double-masked’ in Howick’s preferred terminology) and placebo-controlled, then some of our most effective therapies are not and cannot be supported by the best evidence, for example the Heimlich manoeuvre and external defibrillation. Howick’s point is worth making when one considers how difficult true double-blinding is to achieve, even in drug trials, and practically or ethically impossible for many other therapies, including surgical interventions. As we said above, a key function of evidence is to rule out alternative hypotheses—to eliminate confounding. The best evidence does this in a number of ways, eliminating allocation bias, the placebo effect and so forth. If, as follows from this, a treatment is not, in this sense supported by the best evidence (e.g. because a placebo-control is not available), does that mean that our confidence in the intervention should be less? Not necessarily, argues Howick. Let
us say that our study design fails to eliminate one source of possible confounding, C. We should then consider the possibility that the outcome of the trial is explained by C rather than by the intervention. However, we might come to see that C cannot be the correct explanation. For the study might reveal an effect size that cannot be plausibly be explained by C. In which case we can reject C and our hypothesis that the intervention is efficacious can be confirmed. This is Howick's rule of evidence: studies provide good evidence when the effect size outweighs the combined effects of plausible confounders. This ought to be seen as an important amendment to the EBM conception of evidence. The standard version of the latter ranks studies according to their design. A study design, on this view, that delivers good evidence is one that rules out all hypotheses but two: that the intervention is effective and that the null hypothesis is true. That is what a good study design does; the results of the trial help us decide whether we should reject the null hypothesis or not. Howick is pointing out that matters are not so simple. Some study designs may not themselves rule out other hypotheses (e.g. certain confounders such as the placebo effect). But their results may be able to do that job. Penicillin was accepted as effective without a placebo-controlled trial, but it did not need the latter since the effect size is not plausibly explained by the placebo effect. Tonelli argues that this means that deciding when we can make do without controlled clinical research requires drawing on mechanistic reasoning and clinical experience—which Tonelli takes Howick to be rejecting as suitable sources of medical knowledge (and so this looks like a paradox for Howick). One ought, I think, distinguish between the ability of mechanistic thinking and clinical experience to play a role in reasoning about causal general causal relations and their being direct sources of knowledge in their own right. And it is not so much clinical experience that is relevant here as scientific judgment. Nonetheless, Tonelli is right in essence, that it is impossible to give an account of reasoning in clinical research that ignores mechanistic reasoning and judgment altogether. There is no algorithm for generating knowledge in any field, except in tightly constrained circumstances. And if medicine were to be so constrained it would fail to know much that it does indeed know.

Where does this leave the EBM hierarchy then? In my view, this teaches us that the EBM hierarchy is only a hierarchy of study designs and not really a hierarchy of evidence at all. Whether an actual study provides good evidence depends on both the study design and the results it produces (it also depends on how well the study is conducted). Consequently it is misleading to use the EBM hierarchy as a way of ranking individual studies and the evidence they produce. But that is sometimes how it is presented. Straus et al. (2005: 118), for example, tell us when we are looking for evidence to bin a study if it is not randomized and return to it only there are no studies that are randomized. They do not tell us to consider the results of a non-randomized trial before binning it. But that seems to be a mistake: one should consider both the study design and the results it delivers. On this view there is no mechanical application of the EBM hierarchy. Assessing the overall implications of the evidence requires judgment of the sort that scientists in other fields are used to employing. Of course medicine is different, in that on the one hand most doctors do not have extensive postgraduate training as scientists and on the other hand the consequences of poor judgment are serious. Nonetheless, the role of judgment in assessing evidence ought to be acknowledged rather than ignored. If we ignore it, then we will be in danger of failing to take into consideration valuable evidence from non-randomized sources, such as observational and other retrospective studies. A Cochrane collaboration meta-analysis (Taylor et al. 2011) of RCTs concerning
salt in diet and occurrence of cardiovascular disease concluded that ‘Cutting down on the amount of salt has no clear benefits in terms of likelihood of dying or experiencing cardiovascular disease.’ Yet this ignores a vast amount of epidemiological and other non-randomized sources of evidence. If a study aims to provide an authoritative view on any subject it ought to consider all the relevant evidence (this is what philosophers call the ‘requirement of total evidence’) taking into consideration both the quality of the source of the evidence and the results obtained. In this case the RCT evidence was somewhat weak due to the relatively small number of participants. When that is so, it is particularly important to look at the full range of evidence, both from prospective and retrospective studies.

At the same time, acknowledging that retrospective and non-randomized trials can produce valuable evidence, whose assessment requires judgment, can impel us to aid this judgment with research. How plausible is it then that some effect could be explained by a placebo effect? To be able to answer that question requires some notion of how powerful that effect is. Of course many EBM supporters take it that the placebo effect is considerable, since that explains the apparent efficacy of various forms of complementary and alternative medicine. There are both definitional and methodological difficulties in assessing the scale of the placebo effect. Addressing these, as Howick does with forensic skill, is important, because it is not always clear what the significance of placebo-controls are. Some authors seem to think that placebo-controlled trials measure ‘absolute’, characteristic effect size. It is assumed that the effect of an intervention is the sum of the placebo effect and a characteristic, non-placebo effect which is the effect of (for example) the pharmaceutical component alone. But as Howick points out that there is no reason why these effects should be additive rather than interact in some more complex way; and indeed there is evidence that the effects are not always additive. Howick’s preferred view of placebos is that they should not be seen as methodologically special; rather they should be considered to be one kind of treatment. So one should not see a methodological divide between placebo-controlled trials and active-controlled trials. That seems right to me: a placebo-controlled trial is a special case of an ‘active-controlled’ where one compares a (possibly) expensive new treatment with potential side effects with a very cheap treatment with limited or no side effects, the placebo (putting aside ethical issues of using a placebo as a treatment at all).

Part III raises questions concerning the evidential role of mechanistic (pathophysiological) reasoning and expert judgment. Howick also considers the problem of applying evidence to individual cases. Howick recognizes the significant role that reasoning about the processes of disease and physiology can play in generating hypotheses about potential interventions. But he also thinks that in general the complexity of these processes and the fact that our knowledge is often incomplete means that the evidential value of such reasoning is low; this conclusion, he argues, is supported by the fact that interventions (such as antiarrhythmic drugs) that are strongly suggested by pathophysiological reasoning nonetheless fail to show benefits in randomized clinical trials. Howick nonetheless takes a nuanced view—there are cases where the complexity is low and where the mechanisms are very well understood, with no gaps in our knowledge. Then in such cases, the quality of the mechanistic evidence can be high. This is a point that Howick might have emphasized more, if only to head off above-mentioned criticisms that he does not regard mechanistic reasoning as a relevant to knowledge-generation in medicine at all. Nonetheless,
there is much more to be said here; further work is necessary because it is not true that all mechanistic reasoning is of the same (low?) evidential grade. What more can we say about the evidential value of different forms of mechanistic reasoning? For example, it seems that we should distinguish between an knowledge of a particular pathophysiological mechanism being sufficient for affirming the existence of a predicted causal relation (say between treatment and benefit), on the one hand, and background knowledge of the lack of any plausible mechanism being sufficient for denying the possibility of a causal relation (e.g. between prayer and antecedent recovery, or between ultra dilute homeopathic solutions and clinical outcome). One might think that the latter is a different, evidentially more reliable category of mechanistic reasoning from the former. Howick shows only partial acknowledgment of this point, arguing that in the context of the background beliefs of the early-mid nineteenth century, there were, analogously, mechanistic grounds for denying the possibility of puerperal fever resulting from handling by medical students, and similarly, more recently, for denying the possibility of bacterial infection being a cause of peptic ulcer. (In my view the evidential positions with regard to mechanistic reasoning are quite different in these cases. Nineteenth century physicians were familiar with the idea of disease being spread by contact, even with non-living vectors. Likewise, the existence of acid-loving bacteria such as Acidithiophilus was known before the work of Marshall and Warren.) Furthermore, one ought to note that the value of modern clinical research lies not only in providing evidence of efficacy but also of safety and so of overall risk and benefit, and this is something that mechanistic reasoning may not always provide us with. Thus the value of heparin after pulmonary embolism may be held up as an example of knowledge gained from mechanistic reasoning (Tonelli 2011: 1014). But that reasoning did not reveal the (mechanistically) counterintuitive fact that heparin can *induce* arterial emboli (Kelton and Warkentin 2008).

The conclusions concerning mechanistic reasoning reinforces what I take to be one of the most important messages of this book (although Howick may not himself express things this way): the EBM hierarchy of evidence, while very roughly correct, is a blunt instrument as it stands. First, it cannot be used to rank individual studies by epistemic value: randomized trials may *tend* to be of better evidential quality than observational studies, but not every randomized study is better than every observational study. How individual studies compare also depends on the care with which they were carried out. And even if carried out as well as possible, a randomized trial may provide less good evidence than an observational study if the numbers involved in the former are small and the effects difficult to detect. The evidential quality of a study depends not only on its design and execution, but also on the actual data produced. Secondly, rationality requires that all relevant evidence be considered. For example, in criminal trials we know that eye-witness evidence is generally less reliable than fingerprint evidence. But that would be no reason to exclude eye-witness evidence from trials when fingerprint evidence is available. The jury should see all the relevant evidence; of course the jury ought also be informed of the potential limitations of the different kinds of evidence and should take this into account. Criminological research should aim at assessing and articulating these limitations; it should also aim to identify markers that distinguish the more reliable from the less reliable within each evidence type. Likewise EBM should adopt the epistemological requirement of total evidence—every piece of relevant evidence should be considered in assessing efficacy; a more sophisticated EBM would think about how different kinds of evidence can be combined in making such judgments.
Howick gives short shrift to expert clinical judgment, for the reasons with which I opened this discussion. Not only is expert opinion open to all sorts of bias and prejudice, it can be shown that in many cases it has been (and continues to be) unreliable, and can even be outperformed by mechanical rules even in the application of what is known to individual patients.

There are however two roles for expert judgment that Howick underplays. The first concerns judgment in handling and evaluating multiple and complex sources of data. There is a vast amount of information relevant to a large number of medical problems and interventions, not all of it of the same quality or of the same type; and it is often not univocal. It cannot be reasonable to expect doctors other than experts to be able to master and assess all the appropriate literature. And indeed one of the motivations of the EBM movement is to make the evidence easily available to physicians. If the process of synthesising the evidence were always a matter of carrying out a meta-analysis on a sequence of well-conducted RCTs, then perhaps this process could be carried out in a way that did not depend (much) on expert judgment. If, however, EBM were to grow in sophistication and to attempt to amalgamate different kinds and qualities of evidence (as proposed above), then judgment will be required in that process. Of course, this is a different kind of clinical expertise: it is not expertise born of experience but expertise in handling and assessing evidence.

The second role for judgment is found at the locus of much of the dissent concerning EBM—the role of evidence from research in making clinical decisions about the care of an individual patient. Silva and Wyer (2009) point out that there are two dimensions to EBM—evidence-based guidelines and evidence-based individualized care—and that proponents of EBM have tended to blur the boundaries between them. For Silva and Wyer, after the publication of (Evidence-Based Medicine Working Group 1992), the substantial debate concerns the second of these. That opinion itself might be an achievement of EBM, insofar as the initial activities of the EBM group at McMaster University were focused on the first dimension, producing evidence-based ‘Readers’ Guides’. What then can EBM achieve with regard to the care of individuals? When Rosenberg and Donald (1995: 1122) say that ‘Evidence based medicine . . . can be used to close the gulf between good clinical research and clinical practice’, Tonelli (1998: 1235) tells us that this metaphor ‘illustrates the primary philosophical limit to EBM’ since the gulf in question ‘represents an intrinsic, philosophical gap.’ Criticisms of EBM identify two such gaps. The first concerns any inference from the statistical data to predictions concerning an individual. The second concerns the role of the patient’s particular needs and values in influencing decision-making.

As regards the first kind of gap, it is salutary to be reminded that following mechanical rules is sometimes more reliable than clinical judgment in making diagnoses and recommending interventions (c.f Dawes et al. 1989, although there is disagreement about the extent of this). Howick does acknowledge that judgment will need to be used where there are no guidelines or where the guidelines are not themselves based upon strong evidence. Even putting aside the fact that such cases are still in the majority, I also wonder whether Howick underestimates the need for judgment. He is absolutely right that when a patient is typical of the group on which a treatment has been tested, then the inference to what the likely effect on this patient would be is more straightforward than critics suggest. The fact that this is an individual whereas the study was an exercise in statistics is irrelevant. The statistics are based upon effects for many individuals similar in key respects to this one patient. In the absence of relevant differentiating features, the statistical properties of
the group will translate directly into probabilities concerning this patient. To think
differently is to think that my lottery ticket has chance of winning that differs from
the probability generated by statistical calculation because it is my ticket, the ticket
of an individual. Tonelli (1998: 1236) reminds us that it is easy to image two pa-
tients, both suffering from abdominal pain and identical in all quantifiable aspects
of history, examination, and laboratory data, where one proves to have appendicitis
at surgery while the other does not. If that were the extent of the differences be-
tween the two patients, then EBM may provide us with exactly what is needed. If
it exists, research will tell us what proportion of patients like the two in the exam-
ple will indeed have appendicitis, and so what the probability is that if the patient
undergoes surgery, he will do so without any benefit but with the attendant risks of
the procedure. That information will then feed into the decision that the physician
and patient make about treatment. Tonelli uses the example to raise the question
of the importance of clinical judgment, for there may be non-quantifiable evidence
that an experienced surgeon can detect and know is relevant to the diagnosis; but
EBM has difficulty in integrating such evidence. EBM must (and a number of its ad-
vocates do) recognize the significance of this. But this is not a brick-wall limitation
to EBM. Research and technological developments may allow the factors to which
the surgeon is sensitive to become quantified and so more easily incorporated into
clinical research. For example, some of the non-quantifiable information provided
by a foetal stethoscope has been superseded by the data provided by a cardiotoco-
graph. Even in advance of such developments, the surgeon's own judgment can be
assessed by checking her predictions against what she discovers on surgery. In prin-
ciple, the quantifiable degree of the reliability of her judgments could be used to give
a more accurate assessment of the probability that the individual patient does have
appendicitis.

More generally, there can be a direct inference from population facts to (prob-
abilistic) facts concerning the individual patient, when the patient is known to be
typical of the population from which the statistical facts were gathered. There are
two kinds of case where this will not be true. In the one case, the patient does meet
the criteria that determined the population that was studied (e.g. the patient satis-
fies the inclusion criteria for the trial of a drug). Nonetheless, the patient's doctor
knows that the patient is not typical of that population. For example, a drug may
have been trialled on patients aged between 20 and 60. This patient is 59. Back-
ground knowledge may tell the doctor that older patients may be less likely to ben-
efit than the group overall. Sub-group analysis may help with this kind of case, but
not always. Another kind of case occurs when the patient fails to meet the inclu-
sion criteria. This is the problem of external validity. The therapy may have been
tested on a groups of young adults suffering solely from the one condition. But one's
patient is a child who is suffering from other conditions for which she is taking med-
ication. Furthermore, she has previously experienced unpleasant side-effects from
some (but not all) drugs of a similar (but not identical) class. In such a case it would
be irresponsible to take the statistical data from the trial to fix the chances of re-
cover or of side-effects. Again, in principle, further research could provide the in-
formation required, but in practice it is unlikely, for there may be both ethical and
practical reasons for not carrying out trials on patients in this position; and it may
not even be possible if this combination of conditions is rare. Clinical judgment will
continue to play an important role in these circumstances.

The first kind of gap concerned the use of what we know about populations in
order to come to know something (diagnosis, prognosis, likely effect of an interven-
tion etc.) about an individual. The second kind of gap concerns the use of such knowledge in making decisions about the most appropriate course of treatment for an individual. Perhaps a nagging ethical concern suggests that to use mechanical rules or to infer from the statistics to this patient is to fail to treat this patient as the individual she is. But that would be a fallacy; ethical concerns are not relevant to factual judgments. The treatment of the patient as an individual does not mean treating them as an exception to the rule unless one has good evidence that they might well be an exception. The ethical dimension of individual care comes elsewhere, in the interaction with the patient and above all in the attempt to integrate the evidence concerning prognosis and possible interventions with the needs and values of the individual patient into a programme of treatment; it is the (typically) inevitable role of needs and values that brings about the second kind of gap referred to. Howick does lay stress on the role of judgment in this process, which he regards as the integration of best external evidence with patient values and circumstances. This cannot be purely a matter of judgment, for the term ‘judgment’ implies that it is the clinician who make the decision rather than the clinician and patient together. Howick believes that the time given to GP patients in the UK is unlikely to be sufficient to achieve a good understanding of the patient’s values and circumstances. Howick also believes that it is insufficient to take full advantage of the placebo effect. A good understanding of the placebo effect and the means by which it operates suggests that an extended and empathetic interaction can have a significant therapeutic effect beyond (and very probably independently of) the sham treatment.

Howick concludes his book with two recommendations for the future of EBM that widen its remit. The first is to address the biases induced by conflicts of interest. A study might be of the highest quality according the EBM hierarchy of evidence. But research has shown that it can make a considerable difference to the published outcome if the sponsor of the research has a financial stake in that outcome. If EBM is to promote the use of best evidence, then it ought to consider how to assess studies of this sort and what practices to recommend in order to minimise the biases that can arise. The second challenge is to consider how non-medical interventions and circumstances can affect health outcomes. This would require applying EBM so that it integrates evidence from non-medical sources with medical evidence to assess a broader range of interventions that may improve health.

Anyone writing on the Philosophy of X has to deal with at least two potential audiences. On the one hand there are the practitioners of X who are interested in a philosophical reflection on what they do. On the other hand there are philosophers who are interested in the philosophical questions raised by X. These sets of interests rarely coincide exactly. In particular there are likely to be issues or ways of discussing issues that are likely to be of interest only to those trained in philosophy. Philosophy of medicine is no exception. It is easy to get fairly quickly from discussion of RCTs to philosophical intense discussions of the role of randomization or the nature of causation, articulated in a manner familiar to philosophers but not to most medics. At the same time, a genuinely philosophical contribution needs to go beyond the medicine and bring philosophical acumen to bear; it is not just abstract or grandiose generalizing about medical practice. As far as I can tell (and I bring the philosopher’s perspective to bear here), Howick does an excellent job of navigating the right path. He does not stray into territory that is likely to be only of interest to professional philosophers. On the other hand, this is clearly a philosopher’s book. Physicians may want to know more about how the more nuanced, sophisticated ac-
count of EBM that Howick presents will impact on their practice. But if we focus on the questions that can be illuminated by using the philosopher’s tools, then Howick as advanced the debate considerably; philosophically, this is exemplary philosophy of medicine. Although Howick’s book is philosophy of medicine, being a philosophy of evidence-based medicine, it does not claim to be a complete philosophy of medicine. The title of Tonelli’s (2011) critical review ‘Not a Philosophy of Clinical Medicine’ is accurate but not entirely fair. There is a lot more philosophical work to be done, not least in area that Howick identifies, such as the integration of research findings with patient values in the clinical setting, as well as in many areas unrelated to EBM, such as the nature of the patient experience of illness and medicine. Debates over EBM are heated, but one might hope for a more dispassionate assessment of which ideas associated with that name are worthy of retention development or improvement. Howick has certainly advanced matters in this direction by articulating his own view of the epistemology of EBM so clearly and with an emphasis on argument rather than rhetoric. It is to be hoped that those who take a different view will nonetheless share his approach.

References


