



# Underdetermination in evidence-based medicine

Benjamin H. Chin-Yee BSc

MD/MA Student, Faculty of Medicine, and Institute for the History and Philosophy of Science and Technology, University of Toronto, Toronto, Ontario, Canada

## Keywords

basic science, clinical research, Duhem–Quine thesis, evidence-based medicine, philosophy of science, randomized controlled trial, underdetermination

## Correspondence

Mr Benjamin Chin-Yee  
Faculty of Medicine  
University of Toronto  
152 Saint Patrick St.  
Toronto  
Ontario M5T 3J9  
Canada  
E-mail: benjamin.chinyee@mail.utoronto.ca

Accepted for publication: 18 July 2014

doi:10.1111/jep.12258

## Abstract

This article explores the philosophical implications of evidence-based medicine's (EBM's) epistemology in terms of the problem of underdetermination of theory by evidence as expounded by the Duhem–Quine thesis. EBM hierarchies of evidence privilege clinical research over basic science, exacerbating the problem of underdetermination. Because of severe underdetermination, EBM is unable to meaningfully test core medical beliefs that form the basis of our understanding of disease and therapeutics. As a result, EBM adopts an epistemic attitude that is sceptical of explanations from the basic biological sciences, and is relegated to a view of disease at a population level. EBM's epistemic attitude provides a limited research heuristic by preventing the development of a theoretical framework required for understanding disease mechanism and integrating knowledge to develop new therapies. Medical epistemology should remain pluralistic and include complementary approaches of basic science and clinical research, thus avoiding the limited epistemic attitude entailed by EBM hierarchies.

## Introduction

Evidence-based medicine (EBM) has become an influential movement in medicine but has also raised a number of controversies. This philosophical critique of EBM examines epistemological issues raised by EBM hierarchies of evidence in terms of the problem of underdetermination of theory by evidence as articulated by the Duhem–Quine thesis. EBM hierarchies of evidence privilege clinical research over basic science, favouring systematic reviews and randomized controlled trials (RCTs) while considering 'physiologic studies' and 'bench research' the lowest levels of evidence [1,2]. Scott Sehon and Donald Stanley have applied Quine's thesis to medicine to argue that EBM does not constitute a Kuhnian 'paradigm shift', but rather offers one approach for probing the Quinean 'web of belief', complementary to the methods of basic science [3]. Here, I extend this analysis, focusing on the implications of underdetermination for EBM's epistemic attitude. I argue that EBM hierarchies of evidence exacerbate the problem of underdetermination, leading to a position that eschews knowledge of pathophysiology and disease aetiology thereby providing a limited heuristic for medical research.

The Duhem–Quine thesis illustrates how clinical trials underdetermine theory to such an extent that they fail to meaning-

fully test the core beliefs of medical science. Severe underdetermination is met with conservatism, which explains EBM's scepticism of medical theory and view of disease as statistical associations at a population level. Such an epistemic attitude imposes limitations on medical research: by failing to probe core beliefs regarding the mechanisms of disease and therapeutics, EBM fails to generate a theoretical framework that is necessary for hypothesis generation and integrating knowledge to develop new treatments. Although clinical studies are an important component of medical research, clinical trials are not the sole epistemic authority in medical knowledge nor are they effective as a singular research heuristic. This discussion suggests that medicine should employ complementary approaches of clinical research and basic science to avoid the sceptical position resulting from EBM's underdetermined epistemology.

## The problem of underdetermination and the Duhem–Quine thesis

The Duhem–Quine thesis is one of the most influential concepts in the modern philosophy of science. This thesis is a combination of two separate theses proposed independently by French physicist and philosopher Pierre Duhem and American philosopher WVO

Quine [4–6]. The Duhem–Quine thesis addresses the problem of underdetermination of theory by evidence, which arises because empirical evidence alone provides insufficient grounds for acceptance or rejection of scientific theories, and therefore underdetermines our beliefs about the world. This thesis argues that a hypothesis cannot be tested by observation or experiment in isolation but rather only as part of a theoretical group; as a consequence, empirical data underdetermines theory choice. For my discussion, it will be helpful to first consider Duhem and Quine's ideas separately, to show how these concepts are useful for evaluating EBM's epistemic attitude.

The concept of a theoretical group, '*un ensemble théorique*', was proposed by Duhem [4]. Duhem argued that because experiment tests '*un ensemble théorique*' and not '*une hypothèse isolée*', a falsifying result can only call into question the whole *ensemble* and does not indicate which specific hypotheses should be revised:

Lorsque l'expérience est en désaccord avec ses prévisions, elle lui apprend que l'une au moins des hypothèses qui constituent cet ensemble est inacceptable et doit être modifiée, mais elle ne lui désigne pas celle qui doit être changée [4].

I shall illustrate an application of Duhem's thesis to medical science using a hypothetical, simplified example. A theoretical group in medical science,  $T$ , may include hypotheses:  $T_1$ , 'Disease  $X$  is caused by pathological process  $Y$ ';  $T_2$ , 'Treatment  $Z$  targets pathological process  $Y$ '; and  $T_3$ : 'Treatment  $Z$  abrogates pathological process  $Y$ '. However, there are also auxiliary hypotheses that must be included in the theoretical group, which I shall designate  $A$ . Such auxiliary hypotheses may include statements about the experimental conditions under which a result was produced.

By *modus ponens*, we can infer that if our theoretical group,  $T$ , is true then a certain experimental observation will follow:  $T \rightarrow O$ , 'Treatment  $Z$  cures disease  $X$ '. Similarly by *modus tollens*, if 'Treatment  $Z$  does not cure disease  $X$ ',  $\neg O \rightarrow \neg T$ . However, what our experiment does not tell us is what part of the theoretical group has been falsified and which theories should be modified accordingly. Therefore, given a clinical trial that produces a falsifying result,  $\neg O$ , we could reject any one of our hypotheses that make up the theoretical group,  $T_1, T_2, T_3$ , or any auxiliary hypotheses,  $A$ :  $\neg O \rightarrow \neg(T_1, T_2, T_3 \text{ or } A)$ . We are unable to identify which of the set ( $T_1, T_2, T_3$  or  $A$ ) is false. There are no logical factors to tell us which of our beliefs should be modified; thus, evidence underdetermines theory.

Duhem's thesis helps us understand why the problem of underdetermination is especially severe in EBM's epistemology. Because the number of auxiliary hypotheses is much greater in clinical studies, a clinical trial tests a much larger theoretical group, and therefore its results provide less of a logical basis for the modification of our beliefs. For example, the theoretical group tested by a clinical trial may include auxiliary hypotheses such as:  $A_1$ , 'Appropriate dosage was used in the trial';  $A_2$ , 'Patients were compliant with treatments';  $A_3$ , 'Patients did not have co-morbidities that affected the outcome'; and  $A_4$ , 'Patients in treatment and control groups had similar baseline characteristics'; etc. A falsifying result from the trial could occasion modification of our belief in any of these auxiliary hypotheses ( $A_1, A_2, A_3, A_4$ , etc.) in addition to the more core medical theories being tested.

Take for example the ISEL trial, a large RCT that tested the use of the drug gefitinib, a targeted epidermal growth factor receptor (EGFR) inhibitor, for the treatment of non-small cell lung cancer

(NSCLC) and showed that the drug did not improve survival compared with the control group [7]. Faced with this falsifying result, the authors postulated auxiliary hypotheses to explain the negative outcome:  $A_1$ , 'lack of significant survival benefit in ISEL might be [explained by] suboptimum dosing';  $A_2$ , 'geographic locations of study sites . . . could have influenced the trial results, through environmental factors such as exposure to smoking';  $A_3$ , 'population of patients with highly refractory disease who took part in ISEL could have been inherently non-responsive to any further therapy'; etc. [7]. The falsifying result could also refute more core hypotheses regarding the efficacy of gefitinib in abrogating the pathological processes driving NSCLC, for example, casting doubt on its ability to inhibit EGFR-mediated tumour progression. Indeed, the non-significant primary outcome measured in ISEL led to the belief that gefitinib was 'proven' to be 'ineffective' [8]. However, this conclusion was underdetermined by the primary end point of the trial, which measured survival in the overall population. Gefitinib was, in fact, effective in a subpopulation of patients; nonetheless, because the primary end point did not reach significance, gefitinib was deemed ineffective. Rather than refuting the core hypothesis that gefitinib is efficacious, the non-significant primary end point refutes the auxiliary hypothesis that the RCT was performed on the appropriate patient population.

Auxiliary hypotheses, although still present, are relatively fewer in laboratory research because of the considerable control that its methods enable. EBM's emphasis on RCTs as the 'gold standard' can be seen as an attempt to introduce laboratory-like control into clinical research and eliminate potential confounding variables. The existence of confounding variables increases the number of auxiliary hypotheses being tested, and thus increases the extent of underdetermination. For example, the well-known confounding variable of differences in patient medication adherence introduces a further auxiliary hypothesis that the result of a trial was due to variation in compliance between treatment groups [9]. RCTs may use adherence measurements, exclusion criteria, stratification and randomization to minimize certain confounding variables, thereby reducing the size of the theoretical group tested; however, given the complexity of the clinical setting, underdetermination remains extensive. Randomization may eliminate the auxiliary hypothesis that blinding removed selection bias; however, randomization is unable to eliminate all auxiliary hypotheses [10]. The clinical studies promoted by EBM introduce additional auxiliary hypotheses surrounding trial methodology and the complexities of the clinical setting, and are thus poorly positioned to provide a logical basis for medical theory. By contrast, basic science, using laboratory methods, is better equipped to probe the pathophysiology of disease and pharmacology of therapeutics that form the core beliefs of medical knowledge.

Duhem limited his discussion of underdetermination to the field of experimental physics. Interestingly, Duhem's philosophy was deeply influenced by Claude Bernard's experimental physiology. Bernard argued for a laboratory-based experimental method in medical science in his influential work *Introduction à l'Étude de la Médecine Expérimentale* (1865):

Je considère l'hôpital seulement comme le vestibule de la médecine scientifique; c'est le premier champ d'observation dans lequel doit entrer le médecin, mais c'est le laboratoire qui est le vrai sanctuaire de la science médicale; c'est la

seulement qu'il cherche les explications de la vie à l'état normal et pathologique au moyen de l'analyse expérimentale [11].

Duhem was impressed by Bernard and believed that '*le contrôle expérimentale n'a pas en Physique, la même simplicité logique qu'en Physiologie*' [4]. He believed that the problem of underdetermination does not arise in experimental physiology because the investigator can exert complete control over the experimental conditions to test theories in isolation. Duhem was mistaken on this point: the problem of underdetermination is ultimately insoluble and pertains even in experimental physiology. Quine's thesis of holistic underdetermination, discussed below, makes this point clear. Nonetheless, Duhem's point is salient in that it recognizes that the problem of underdetermination is less severe in well-controlled research; it is indeed much simpler to test the core hypotheses of medical science using the methods of experimental physiology.

## Response to underdetermination

Appreciating that clinical research is not equipped to determine the core beliefs of medical science helps explain EBM's epistemic attitude. For this explication, it is useful to consider Quine's thesis. Quine's view on underdetermination can be understood as a stronger version of Duhem's thesis, although it was formulated independently using a different line of reasoning. For Quine, the problem of underdetermination is not limited to physics or even to the sciences in general, but rather applies to all of our knowledge:

The totality of our so-called knowledge or beliefs, from the most casual matters of geography and history to the profoundest laws of atomic physics or even of pure mathematics and logic, is a man-made fabric, which impinges on our experience only along the edges [5].

Quine argues that any observation can occasion a corresponding modification of any aspect of our total 'field' of knowledge or beliefs [5]. According to Quine, any theory in our field of beliefs is subject to revision, while, at the same time, any theory can be held constant in light of experience and accommodated by adjustment of other beliefs. For Quine,

The total field is so underdetermined by its boundary conditions (experience) that there is much latitude of choice as to what statements to re-evaluate in light of any single contrary experience [5].

Sehon and Stanley argue that EBM operates somewhere between the periphery and the middle of our field of beliefs in medicine, which I believe accurately characterizes EBM's epistemic limitations [3]. Indeed, as illustrated by the example of the ISEL trial, the core of the field (i.e. our understanding of the pathophysiology of NSCLC) is too underdetermined by EBM for its evidence to provide any logical basis for adjustment of fundamental theories.

Holistic underdetermination leads to what is perhaps the most significant impact the Duhem–Quine thesis: it introduces a role for extralogical factors in determining our beliefs about the world. Duhem argued that in light of underdetermination, theory choice should be guided by scientific intuition [4]. Quine believed theory choice was governed by psychological conservatism, whereby we choose to revise the parts of our field of beliefs most 'germane' to experience [5]. Thus, faced with a falsifying experimental result or

observation, we are psychologically disposed to revise hypotheses nearer to the periphery of our field of beliefs, rather than refuting theories closer to the core.

Thomas Kuhn's notion of incommensurability between paradigms can be understood as a problem of underdetermination of theory by evidence. Kuhn suggested that extralogical values of the scientific community, such as 'accuracy, consistency, scope, simplicity, and fruitfulness', played an important role in theory choice between paradigms [12]. Kuhn's work – much to his own dismay – influenced a movement that looked towards the role of sociological factors, such as power, race or gender, in determining our scientific beliefs. Larry Laudan has criticized this 'sociologizing of epistemology' and maintains that epistemic considerations other than deductive logic can determine theory choice [6]. The role of sociologic factors in determining theory choice in medicine is an interesting topic to explore, especially in the context of EBM where underdetermination is particularly extensive. It is possible that because of severe underdetermination, the beliefs of EBM are more open to social influences – be they financial, gender related or cultural/racial – than are other epistemic approaches. Indeed, critiques of EBM have focused on how its knowledge claims are particularly susceptible to being influenced by special interests [13].

Following Quine, I argue that, given severe underdetermination, EBM practitioners adopt an attitude of conservatism. Proponents of EBM are confident in their research methods, such that results of a well-conducted RCT do not occasion modification of auxiliary hypotheses concerning study design. They believe that RCTs adequately control for such auxiliary hypotheses, and thus these statements are not subject to reform in the case of a falsifying result. But EBM is conservative in the sense that it will probe no deeper than the middle of the field of beliefs; clinical studies can occasion adjustment of mid-field hypotheses concerning the efficacy of a treatment in producing a specified endpoint under particular circumstances, but they cannot warrant revision of core theories about disease aetiology and therapeutic mechanisms – the *why* and the *how* of medical knowledge.

Another example from the oncology literature illustrates this point. The INTEREST trial compared treatment with the targeted EGFR inhibitor gefitinib to the cytotoxic chemotherapy docetaxel and showed that the two drugs had equivalent impact on overall survival in patients with NSCLC [14]. However, concluding based on this result that gefitinib and docetaxel are equivalent fails to recognize that the drugs work by different mechanisms [8]. Differences in the drugs' mechanisms of action are underdetermined by the results of the clinical study. One might argue that such differences are not important because the overall survival of patients was the same given either drug. Nonetheless, differences in mechanism are important for understanding *why* the drugs target different subpopulations, which *is* clinically relevant knowledge. Gefitinib targets patients with *EGFR* mutations, whereas docetaxel gives a higher response in *EGFR* wild-type patients. There are further examples in oncology where significant observations may be missed by the crude empiricist approach of RCTs, which underdetermine knowledge of pathophysiology and pharmacology [8]. As Richard Ashcroft points out, RCTs 'bracket out a whole range of scientifically and epistemologically difficult questions about why treatments work (and about why they sometimes don't)' [15]. They 'appear to settle the question of whether

a treatment works or not, without need for appeal to any deeper biological theory to explain *why* the treatment works or not' [15]. Because these core theories are so underdetermined by its preferred forms of evidence, EBM avoids such aspects of medical knowledge.

## EBM and basic science

In response to severe underdetermination, EBM adopts an empiricist attitude characterized by its scepticism of basic science and mechanistic reasoning [16]. In philosophy, empiricism is contrasted by rationalism: the empiricist derives knowledge from experience, whereas the rationalist gains knowledge by reasoning from first principles [17]. Robyn Bluhm and Kirstin Borgerson argue that in the philosophy of medicine, rationalism and empiricism are better understood as two empiricist approaches applied at different levels: rationalists seek to understand basic mechanisms of pathophysiology, which they investigate empirically, whereas empiricists operate at the population level to determine efficacy in the average patient [16]. This useful distinction helps to show how both rationalism and empiricism have important roles in medical epistemology, and highlights the deficiencies of EBM's epistemic attitude.

EBM traces its philosophical lineage to the medical empiricism that emerged in Paris during the 19th century [18]. Pierre Charles Alexandre Louis was a notable empiricist at the Parisian medical school who pioneered the 'numerical method' and conducted early clinical studies. Jules Gavarret analysed Louis's work and argued that clinical studies involving large numbers of patients were required to evaluate the efficacy of treatments, according to the statistical principle of '*la loi des grands nombres*' [19]. Nonetheless, Gavarret recognized the limitations of clinical studies; he accepted that clinical research required prior knowledge of pathological anatomy – the basic science of his day – and that statistical approaches could not be applied if disease nosology was not well established:

La maladie en expérience doit avoir un diagnostic nettement et parfaitement défini. [. . .] Tant que le diagnostic d'une affection n'aura pas atteint ce haut point de certitude, on essaierait vainement de recueillir les éléments d'une bonne statistique destinée à éclairer sa thérapeutique, un pareil travail ne pourrait conduire qu'à des conclusions erronées et d'autant plus funestes qu'elles sembleraient plus rigoureusement déduites [19].

Gavarret argued that because the ancients lacked knowledge of pathological anatomy, advanced by the Parisian school, they were unable to effectively apply statistical methods [19]. He acknowledges that clinical studies and '*la loi des grands nombres*' are limited in their ability to resolve the presence or non-presence of a cause, and that determining the cause itself is another consideration, outside of the domain of statistics:

Toutes les fois qu'il s'agit d'étiologie, les principes de la loi des grands nombres ne peuvent servir qu'à prouvé l'existence ou la non-existence d'une cause spéciale soupçonné indépendamment de toute hypothèse faite sur la nature. C'est à l'aide de considération d'un autre ordre qu'on doit chercher à déterminer la cause elle-même; cette dernière question est hors de la sphère d'activité de la statistique [19].

Although Gavarret promoted medical empiricism and was an early proponent of clinical research, he acknowledged the limitations of this approach. He emphasized that knowledge of underlying pathology is necessary for successful clinical research. Furthermore, he recognized the need for methods other than clinical studies to determine the causes of disease. EBM claims philosophical heritage among the Parisian medical empiricists but as we have seen, it has moved beyond the empiricism of its early heroes. Paul Thompson has explored this issue, showing that EBM employs mathematics as a mere tool of analysis and fails to use the language of mathematics to establish causal models of disease, which is a requirement for robust scientific understanding [20].

An example from the *Users' Guides to the Medical Literature* illustrates EBM's mistrust of basic biological science. Gordon Guyatt *et al.* cite the results of a clinical trial in which subgroup analysis suggested that aspirin was effective for stroke prevention in men and not in women [1]. A larger RCT to test this difference was subsequently conducted and refuted this result. This example is used by Guyatt *et al.* to show how post-hoc subgroup analysis can lead to erroneous conclusions. Interestingly, however, the results of the subgroup analysis in the initial clinical study stimulated animal research that confirmed the discrepancy between males and females, and showed a mechanism for it, suggesting it was caused by differences in platelet activation and pharmacokinetics between the sexes [21–23]. For Guyatt *et al.*, however, the result of the subsequent RCT rendered the underlying biology insignificant. It did not matter that there are physiologic mechanisms that could have explained the differences in the initial study; if the trial is conducted over a large enough population, there is no difference observed between sexes.

EBM's attitude is sceptical of biological explanation and limits its operations to mid-field space, going no further than to modify hypotheses about the efficacy of a drug at a population level. For EBM, the animal studies represent a mere distraction from the important result. Guyatt *et al.* are dismissive of the underlying biology, writing, 'the human mind is sufficiently fertile that there is no shortage of biologically plausible explanations in support of almost any observation' [1]. This statement reflects scepticism towards medical theories concerning disease and treatment mechanisms. However, differences in such mechanisms may have important clinical implications. There remain important differences in the effects of aspirin between sexes, as well as in subpopulations of patients who show low response to antiplatelet therapies, and these variations require further investigation [24,25]. Concluding, 'Ultimately . . . aspirin for stroke reduction was as effective in men as it was in women', and dismissing the underlying mechanism fails to appreciate nuances that may have clinical relevance [1]. For example, sex differences in pharmacokinetics can have important implications for dosing – as was recently shown to be the case for the hypnotic zolpidem – and such differences are often underdetermined by RCTs [26].

A more useful medical epistemology recognizes the importance of both rationalist and empiricist approaches. Jeremy Howick has written on the role of basic science and mechanism in EBM, acknowledging that 'high-quality mechanistic reasoning' can improve the strength of claims when coupled with clinical evidence [13]. However, Howick also emphasizes several examples where mechanistic reasoning has led to failed therapies, for instance the use of antiarrhythmic drugs to reduce sudden cardiac

death in patients post-myocardial infarction, which were shown to increase mortality. In so doing, he minimizes the importance of basic science and mechanistic reasoning in generating the theoretical framework and hypotheses tested by RCTs. Howick argues that 'observationally' derived hypotheses, based on anecdotal evidence, 'serve a practical benefit over testing hypotheses from the basic sciences' [13]. Whereas 'basic science usually yields hypotheses about therapies that are not currently in use', testing remedies from 'ready-made anecdotal evidence' would have more immediate benefit by encouraging or discouraging current usage [13]. This crude empiricist approach would severely limit our stock of therapies to those currently available and those that emerged from anecdotal case reports. Howick diminishes the importance of basic science in the development of treatments, citing a study showing that a low percentage of 'mechanistic findings' translates into clinically relevant therapies [27]. Quantifying translation of basic science research as a percentage fails to recognize the tremendous impact the development of a single therapy can have. In this study, the one case where basic science did produce clinical advantages was in the development of angiotensin-converting enzyme inhibitors, which have had a tremendous impact in the treatment of hypertension as well as other cardiovascular and renal diseases, and are the fifth most prescribed class of drugs in the United States [28]. Howick's view of basic science as 'an expensive way to generate hypotheses' fails to recognize the importance of mechanistic reasoning in medical research and reveals an inherent weakness of EBM's epistemic attitude [13].

## Implications for medical research

EBM hierarchies favour evidence that underdetermines the core of our field of medical beliefs, resulting in an epistemic attitude that is sceptical of disease pathophysiology. Such an attitude imposes limitations on medical research because it neglects the theoretical framework that is required for integrating knowledge and developing new therapies. Early proponents of EBM declared the movement a 'new paradigm', seeking to replace a 'former paradigm', which among other pitfalls was overly reliant on 'knowledge of basic science and disease mechanisms' [29]. In addition to reforming clinical decision making, EBM sought to influence the future direction of research to favour funding of RCTs over other methods of research [30]. In the field of research, EBM's choice of the term 'paradigm' carries especially troubling connotations. According to Kuhn, a paradigm is all encompassing and determines the types of questions asked by researchers, as well as the methods at their disposal [31]. EBM dictates the appropriate ways of knowing and encourages a particular epistemic viewpoint that is unlikely to provide a fruitful research heuristic.

A brief survey of medical history over the past century reveals how major medical advances came about through probing mechanisms of disease. Although epidemiologic studies have also had an important impact in medicine, progress in medical therapeutics required a basic understanding of disease aetiology. The classic example that underscores the importance of basic science in the advancement of medical knowledge is the serendipitous discovery of penicillin by Alexander Fleming, which came about through laboratory experiments on *Staphylococcus* and *Penicillium* mould [3]. More recent advances in the biological sciences of physiology, microbiology, biochemistry, molecular genetics and pharmacol-

ogy have provided a deeper understanding of mechanisms of disease and allowed for the development of many new treatments.

An example in haematology is the case of chronic myeloid leukaemia (CML) and its novel targeted therapy, the tyrosine kinase inhibitor imatinib. The clinical features of CML have been recognized by doctors since the 19th century. Chemotherapeutic agents such as hydroxyurea, busulfan, cytarabine and biologic agents such as interferon alpha had limited impact on the natural history of this ultimately fatal disease. In the 1990s, bone marrow transplant offered a potentially curative treatment for a limited number of patients with matched donors but was associated with significant treatment-related mortality

While clinical research tested the efficacy of these non-specific cytotoxic therapies, basic science was investigating the mechanisms of malignant transformation. A major advance in the understanding of CML came in 1960 when a 'minute chromosome' in cells of patients was identified, dubbed the 'Philadelphia chromosome' [32]. Over the following decades, building on advances in cytogenetics and molecular genetics, researchers showed that the Philadelphia chromosome was the product of a reciprocal translocation fusing two proto-oncogenes that produced a constitutively active tyrosine kinase, BCR-ABL [33–36]. Animal studies showed that the BCR-ABL kinase was sufficient to cause leukaemia [37]. Thus, these basic approaches successfully identified the aetiology of CML at a molecular level.

Once the pathophysiologic mechanism was understood, it was possible to develop therapies that targeted the underlying cause of CML, the aberrant BCR-ABL kinase. In 1996, researchers discovered a selective inhibitor of the BCR-ABL kinase, a drug later named imatinib [38]. Imatinib's efficacy in preventing tumour growth and specificity for BCR-ABL expressing tumours was demonstrated in leukaemic cell lines. Indeed, a phase I trial – an observational study – confirmed the safety and efficacy of imatinib, which resulted in complete remission in 98% of patients treated with the drug [39]. Imatinib underwent accelerated approval becoming the standard of care for all patients with CML. In the case of CML, understanding disease mechanism was a prerequisite for developing effective therapy. Empirical approaches that tested the effectiveness of different general chemotherapies failed to produce significant advances in the treatment of CML.

In 2003, an RCT showed that imatinib was much more effective at improving survival and slowing progression of CML than the previous standard of care, interferon alpha and cytarabine [40]. According to EBM hierarchies, one might triumphantly claim that the result of this RCT constitutes the 'best evidence' for confidence in the therapy. Although the study provides strong evidence for the clinical use of imatinib, giving precedence to the RCT fails to recognize the decades of basic science research that generated the theoretical basis for the trial. Such a claim does not acknowledge how our understanding of CML's pathophysiology also provides considerable epistemic force for belief in imatinib's efficacy.

Once the results of the RCT were available, it did not suffice to abandon our knowledge of the underlying disease mechanism; a basic scientific understanding provided the theoretical framework required to integrate knowledge and adapt to new challenges. Indeed, new challenges arose in the treatment of CML: some patients' leukaemia evolved resistance to imatinib. However, understanding the aetiology of the disease and the cause of

resistance (a resistant BCR-ABL kinase) enabled the development of a new generation of tyrosine kinase inhibitors that have been effective in this resistant subpopulation [41].

Indeed, as EBM proponents assuredly remind us, mechanistic reasoning can lead astray [13]. However, EBM fails to appreciate the importance of medical theory and mechanism in the use of current treatments and in the development of new therapies. Basic science is not only an important source of hypotheses for clinical testing, but can also occasion modifications at the core of our field of beliefs in such a way that it changes the types of questions asked and the approaches to disease treatment. CML and imatinib advanced our understanding of cancer biology, creating a new model for drug discovery and targeted therapies. This knowledge has had an impact beyond patients with CML, leading to the development of other targeted drugs. One example is gefitinib in NSCLC mentioned above, which also illustrates how RCTs often fail to accurately determine the efficacy of such targeted therapies.

## The role of clinical research

This article is not intended as a critique against clinical research; rather, it challenges the hierarchical epistemology of EBM in which RCTs are the ultimate epistemic authority and other ways of knowing are devalued. As seen above, clinical research predates the EBM movement and, indeed, such approaches have demonstrated their importance for testing the efficacy and safety of treatments. Gavarret was an early proponent of clinical studies and argued for the importance of evaluating the efficacy of therapies by systematic observation over a large number of patients. However, Gavarret recognized that clinical research alone was insufficient to determine disease aetiology and could only tell us about the presence or absence of a cause. He maintained that clinical evidence was complementary to knowledge of underlying pathology; he praised advances in pathological anatomy, the basic science of his day, as a necessary prerequisite for successful clinical research. The concept of a hierarchy, the first 'fundamental principle' of EBM, is in necessary opposition to this notion of complementarity [1].

Although evidence from clinical research and RCTs has an important role in medicine, it does not form the exclusive basis for confidence in treatments. RCTs only tell us about the use of drugs under certain controlled conditions that are often much different from the settings in which they are used, such as in patients with significant co-morbidities or in paediatric populations. In some cases, theories of disease mechanism provide greater epistemic force for evaluating the efficacy of a treatment. For example, core medical theories lead us to be sceptical of the therapeutic benefit of remote intercessory prayer in treatment despite evidence from RCTs supporting its efficacy [42]. The same applies to homeopathic remedies: basic science provides good reason to doubt their efficacy, even though some RCTs have provided evidence that they are effective [3]. By contrast, knowledge of pathophysiology provided confidence in the efficacy of rabies vaccination and penicillin for the treatment of pneumococcal pneumonia before RCTs were ever conducted [43]. CML is another case where an understanding of disease mechanism underwrites belief in the efficacy of a treatment.

Epidemiologic and clinical research are indeed useful for identifying problems and stimulating further investigation but it must

be recognized that such approaches have their limitations. The stringent design of RCTs restricts them to answering specific types of questions, primarily those comparing pharmaceutical interventions, leaving large and important domains of medicine inaccessible to its methods. RCTs are also limited by ethical considerations, such as whether or not it is ethical to randomize patients between treatment groups, explored by John Worrall in his case study of extracorporeal membrane oxygenation for persistent pulmonary hypertension of the newborn [43].

Although I argue that laboratory bench research is better positioned to probe the mechanisms of disease, this article is not making the case for an alternative hierarchy that favours basic science. Given the diversity and complexity of problems faced in medicine, an approach that favours any single methodology is unlikely to be successful. A pluralistic epistemology would appear most appropriate; however, pluralism is not compatible with EBM. Its hierarchies privilege certain ways of knowing above others, and this fact means that EBM is implemented to the exclusion of other approaches.

## Conclusion

This article has examined philosophical issues raised by EBM's epistemology in terms of the problem of underdetermination of theory by evidence and the Duhem–Quine thesis. I have argued that EBM hierarchies of evidence, in which clinical research and RCTs supersede other ways of knowing, exacerbate the problem of underdetermination and result in an epistemic attitude that is sceptical of mechanistic reasoning and explanations from the basic biological sciences. This impoverished perspective prevents the development of a theoretical framework required for a robust understanding of diseases and the development of new therapies. Medicine should employ complementary approaches of basic science and clinical research, avoiding EBM hierarchies that provide a limited research heuristic for solving the problems faced in clinical medicine.

## Acknowledgements

I would like to thank Dr. Paul Thompson, Dr. Ross Upshur and Dr. Ian Chin-Yee for their comments on earlier drafts of this essay.

## References

1. Guyatt, G., Rennie, D., Meade, M. & Cook, D. (2008) *Users' Guides to the Medical Literature: Essentials of Evidence-Based Clinical Practice*, 2nd edn. New York: McGraw Hill Professional.
2. Phillips, B., Ball, C., Sackett, D., Badenoch, D., Straus, S., Haynes, B. & Dawes, M. (1998) Oxford Centre for Evidence-based Medicine Levels of Evidence [Internet] [updated by Jeremy Howick 2009 Mar 1]. Available at: <http://www.cebm.net/index.aspx?o=1025> (last accessed 16 February 2014).
3. Sehon, S. R. & Stanley, D. E. (2003) A philosophical analysis of the evidence-based medicine debate. *BMC Health Services Research*, 3 (1), 14.
4. Duhem, P. (1906) *La théorie physique: son objet, et sa structure*. Paris: Chevalier & Rivière.
5. Quine, W. V. O. (1980) Two dogmas of empiricism. In *From a Logical Point of View: 9 Logico-Philosophical Essays*, pp. 20–46. Cambridge: Harvard University Press.

6. Curd, M. & Cover, J. A. (1998) *Philosophy of Science: The Central Issues*. London: WW Norton.
7. Thatcher, N., Chang, A., Parikh, P., *et al.* (2005) Gefitinib plus best supportive care in previously treated patients with refractory advanced non-small-cell lung cancer: results from a randomised, placebo-controlled, multicentre study (Iressa Survival Evaluation in Lung Cancer). *Lancet*, 366 (9496), 1527–1537.
8. Stewart, D. J. & Kurzrock, R. (2013) Fool's gold, lost treasures, and the randomized clinical trial. *BMC Cancer*, 13 (1), 193.
9. Gossec, L., Tubach, F., Dougados, M. & Ravaud, P. (2007) Reporting of adherence to medication in recent randomized controlled trials of 6 chronic diseases: a systematic literature review. *The American Journal of the Medical Sciences*, 334 (4), 248–254.
10. Worrall, J. (2002) What evidence in evidence-based medicine? *Philosophy of Science*, 69 (S3), S316–S330.
11. Bernard, C. (1865) *Introduction à l'étude de la médecine expérimentale*. Paris: JB Baillière et fils.
12. Kuhn, T. S. (1977) *The Essential Tension: Selected Studies in Scientific Tradition and Change*. Chicago: University of Chicago Press.
13. Howick, J. (2011) *The Philosophy of Evidence-Based Medicine*. Chichester: Wiley-Blackwell, BMJ Books.
14. Kim, E. S., Hirsh, V., Mok, T., *et al.* (2008) Gefitinib versus docetaxel in previously treated non-small-cell lung cancer (INTEREST): a randomised phase III trial. *The Lancet*, 372 (9652), 1809–1818.
15. Ashcroft, R. (2002) What is clinical effectiveness? *Studies in the History and Philosophy of Biological and Biomedical Sciences*, 33 (2), 219–233.
16. Bluhm, R. & Borgerson, K. (2011) Evidence-based medicine. In *Handbook of the Philosophy of Medicine Volume 16: Philosophy of Medicine* (eds D. M. Gabbay, P. Thagard & J. Woods), pp. 203–238. Amsterdam: Elsevier.
17. Marcum, J. A. (2008) *An Introductory Philosophy of Medicine: Humanizing Modern Medicine*. Dordrecht: Springer.
18. Sackett, D. L., Rosenberg, W. M., Gray, J. A., Haynes, R. B. & Richardson, W. S. (1996) Evidence based medicine: what it is and what it isn't. *BMJ (Clinical Research Ed.)*, 312 (7023), 71–72.
19. Gavarret, J. (1840) *Principes généraux de statistique médicale*. Paris: Béchét Jne et Labé.
20. Thompson, R. P. (2010) Causality, mathematical models and statistical association: dismantling evidence-based medicine. *Journal of Evaluation in Clinical Practice*, 16 (2), 267–275.
21. Kelton, J. G., Hirsh, J. & Carter, C. J. (1978) Sex differences in the antithrombotic effects of aspirin. *Blood*, 52 (5), 1073–1076.
22. Kelton, J. G., Carter, C. J., Santos, A. & Hirsh, J. (1982) Sex related differences in platelet function: the effect of aspirin. *Blood*, 59 (3), 625–627.
23. Buchanan, M. R., Rischke, J. A., Butt, R. & Turpie, A. (1983) The sex-related differences in aspirin pharmacokinetics in rabbits and man and its relationship to antiplatelet effects. *Thrombosis Research*, 29, 125–139.
24. Berger, J. S., Roncaglioni, M. C. & Avanzini, F. (2006) Aspirin for the primary prevention of cardiovascular events in women and men: a sex-specific meta-analysis of randomized controlled trials. *JAMA: The Journal of the American Medical Association*, 295 (3), 306–313.
25. Ben-Dor, I., Kleiman, N. S. & Lev, E. (2009) Assessment, mechanisms, and clinical implication of variability in platelet response to aspirin and clopidogrel therapy. *American Journal of Cardiology*, 104 (2), 227–233.
26. Greenblatt, D. J., Harmatz, J. S., Singh, N. N., *et al.* (2013) Gender differences in pharmacokinetics and pharmacodynamics of zolpidem following sublingual administration. *Journal of Clinical Pharmacology*, 54 (3), 282–290.
27. Contopoulos-Ioannidis, D. G., Ntzani, E. E. & Ioannidis, J. P. A. (2003) Translation of highly promising basic science research into clinical applications. *The American Journal of Medicine*, 114 (6), 477–484.
28. IMS Institute for Healthcare Informatics (2012). The uses of medicine in the United States: review of 2011 [Internet]. Available at: [http://www.imshealth.com/ims/Global/Content/Insights/IMS%20Institute%20for%20Healthcare%20Informatics/IHII\\_Medicines\\_in\\_U.S\\_Report\\_2011.pdf](http://www.imshealth.com/ims/Global/Content/Insights/IMS%20Institute%20for%20Healthcare%20Informatics/IHII_Medicines_in_U.S_Report_2011.pdf) (last accessed 16 February 2014).
29. Evidence-Based Medicine Working Group (1992) Evidence-based medicine: a new approach to teaching the practice of medicine. *JAMA: The Journal of the American Medical Association*, 268 (17), 2420–2425.
30. Sackett, D. L. (1999) Time to put the Canadian Institutes of Health Research on trial. *Canadian Medical Association Journal*, 161 (11), 1414–1415.
31. Kuhn, T. S. (2012) *The Structure of Scientific Revolutions*, 4th edn. Chicago: University of Chicago Press.
32. Nowell, P. C. (1960) A minute chromosome in human chronic granulocytic leukemia. *Science*, 132, 1497.
33. Rowley, J. D. (1973) A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. *Nature*, 243 (5405), 290–293.
34. Heisterkamp, N., Stephenson, J. R., Groffen, J. & Hansen, P. F. (1983) Localization of the c-abl oncogene adjacent to a translocation break point in chronic myelocytic leukaemia. *Nature*, 306, 239–242.
35. Shtivelman, E., Lifshitz, B., Gale, R. P. & Canaani, E. (1985) Fused transcript of abl and bcr genes in chronic myelogenous leukaemia. *Nature*, 315, 550–554.
36. Lugo, T. G., Pendergast, A. M., Muller, A. J. & Witte, O. N. (1990) Tyrosine kinase activity and transformation potency of bcr-abl oncogene products. *Science*, 247 (4946), 1079–1082.
37. Daley, G. Q., Van Etten, R. A. & Baltimore, D. (1990) Induction of chronic myelogenous leukemia in mice by the P210bcr/abl gene of the Philadelphia chromosome. *Science*, 247 (4944), 824–830.
38. Druker, B. J., Tamura, S., Buchdunger, E., *et al.* (1996) Effects of a selective inhibitor of the Abl tyrosine kinase on the growth of Bcr-Abl positive cells. *Nature Medicine*, 2 (5), 561–566.
39. Druker, B. J., Talpaz, M., Resta, D. J., *et al.* (2001) Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. *The New England Journal of Medicine*, 344 (14), 1031–1037.
40. O'Brien, S. G., Guilhot, F., Larson, R. A., *et al.* (2003) Imatinib compared with interferon and low-dose cytarabine for newly diagnosed chronic-phase chronic myeloid leukemia. *The New England Journal of Medicine*, 348 (11), 994–1004.
41. Khoury, H. J., Cortes, J. E., Kantarjian, H. M., *et al.* (2012) Bosutinib is active in chronic phase chronic myeloid leukemia after imatinib and dasatinib and/or nilotinib therapy failure. *Blood*, 119 (15), 3403–3412.
42. Giacomini, M. (2009) Theory-based medicine and the role of evidence: why the emperor needs new clothes, again. *Perspectives in Biology and Medicine*, 52 (2), 234–251.
43. Worrall, J. (2008) Evidence and ethics in medicine. *Perspectives in Biology and Medicine*, 51 (3), 418–431.