



# Mechanisms: what are they evidence for in evidence-based medicine?

Holly Andersen PhD MSc

Assistant Professor, Philosophy Department, Simon Fraser University, Burnaby, British Columbia, Canada

## Keywords

complexity, evidence, mechanisms, reasoning

## Correspondence

Assistant Professor Holly Andersen  
Philosophy Department  
Simon Fraser University  
8888 University Drive  
Burnaby, BC V5A 1S6  
Canada  
E-mail: holly\_andersen@sfu.ca

Accepted for publication: 20 July 2012

doi:10.1111/j.1365-2753.2012.01906.x

## Abstract

Even though the evidence-based medicine (EBM) movement labels mechanisms a low quality form of evidence, consideration of the mechanisms on which medicine relies, and the distinct roles that mechanisms might play in clinical practice, offers a number of insights into EBM itself. In this paper, I examine the connections between EBM and mechanisms from several angles. I diagnose what went wrong in two examples where mechanistic reasoning failed to generate accurate predictions for how a dysfunctional mechanism would respond to intervention. I then use these examples to explain why we should expect this kind of mechanistic reasoning to fail in systematic ways, by situating these failures in terms of evolved complexity of the causal system(s) in question. I argue that there is still a different role in which mechanisms continue to figure as evidence in EBM: namely, in guiding the application of population-level recommendations to individual patients. Thus, even though the evidence-based movement rejects one role in which mechanistic reasoning serves as evidence, there are other evidentiary roles for mechanistic reasoning. This renders plausible the claims of some critics of EBM who point to the ineliminable role of clinical experience. Clearly specifying the ways in which mechanisms and mechanistic reasoning can be involved in clinical practice frames the discussion about EBM and clinical experience in more fruitful terms.

## Introduction

Evidence-based medicine (EBM) has rapidly become a dominant approach to clinical medical practice. The move towards evidence-based treatment guidelines to make clinical decisions regarding patients was at least partially motivated by what medical practitioners saw as the systematic failures of other forms of reasoning in achieving the best health outcomes for patients. Importantly for philosophers of science, EBM was partly a reaction to what was perceived as a fairly widespread failure of mechanisms as evidence in clinical medical practice. Mechanistic reasoning in clinical practice often starts from a knowledge of various mechanisms in the human body and identifies a locus of dysfunction, and interventions are chosen to restore that locus to normal functioning. Its key characteristic is the reliance on mechanisms to make predictions about the outcomes of interventions. In contrast, EBM reasoning relies on treatment recommendations distilled from high-quality studies to make maximally efficacious treatment recommendations, for which no mechanistic justification may be known.

Given that mechanisms are the basis of some of the strongest accounts of explanation in the sciences currently available, it is

worth investigating, and worth questioning the legitimacy of, the tension between mechanisms as a useful account of biological explanation and their apparent unreliability in medicine. The relationship between mechanisms and EBM turns out to be illuminating for what it tells us about explanation, about reasoning practices in EBM and about the causal structure of mechanisms in the human body.

I have several goals in this paper. The main goal is to examine the role that mechanisms play in EBM: there are several distinct stages at which mechanisms might be used as evidence in diagnosis and treatment, some of which are replaced by EBM treatment guidelines, and some of which are necessitated by those same guidelines. This means that even in EBM, mechanisms cannot be entirely eliminated as evidence in clinical practice (this paper will discuss clinical practice, where treatment of patients is the main focus, rather than e.g. medical research). Rather, it means that we must be very clear on what mechanisms are to be considered potential evidence *for*: in generating predictions about potentially useful interventions, or as a means of applying EBM guidelines to individual patients. The second goal is to identify reasons why mechanisms might fail in guiding medical treatment: such failures stem from the complex, evolved, layered causal structure of the

human body, which results in mechanisms that frequently are non-modular and involve ‘hidden’ causal relationships. Finally, the third goal is to use this discussion of mechanisms as a way of making more precise some of the claims made by critics of EBM that point to an ineliminable role for clinical experience. I render this claim more specific by pointing out how mechanisms, in their second evidentiary role, are part of the knowledge critics point to as clinical experience. As such, considering the role of mechanisms allows for a more fruitful framing of the disagreement between EBM advocates and critics than simply one about evidence-based versus ‘fuzzy’ decision making.

We will see that a fundamental issue with respect to mechanisms and EBM is that there is a discrepancy between knowledge of the causal mechanisms responsible for many functions, and knowledge about what will happen under intervention on those mechanisms. One might have reliable knowledge of a given mechanism in the body, but not be able to use that knowledge to generate effective interventions to restore that mechanism once it fails. This gap supports a distinction between mechanistic *explanation*, in which mechanism models explain (sometimes very accurately) how a given mechanism works, versus mechanistic *reasoning* that involves predicting outcomes of interventions on those mechanisms, which may be confounded by a number of factors external to the mechanism itself. We can model the causal structure of some systems quite effectively in terms of mechanisms, and yet be unable to use that knowledge to predict the outcomes of interventions in those structures. Mechanisms are considered, for good reason, to be a primary form of causal explanation in many sciences; and interventions are at the heart of many contemporary approaches to causation, for example [1]. The failure of causal explanations to match up with corresponding causal interventions should thus be an issue of concern to those who are interested in mechanisms, explanation, prediction and, especially, causal methodology. The case of mechanisms in medicine renders acute and concrete the failure of causal methodology to accommodate the epistemic circumstances of sciences that study complex systems.

## Mechanistic reasoning and EBM

Mechanisms have received a great deal of attention as an account of the form taken by explanations in the so-called ‘higher-level’ sciences, such as biology or neuroscience, as compared with law-based accounts of explanation drawn from physics [2–7]. Explanations of phenomena of interest, on this new understanding of mechanisms, involve details about causal mechanisms that give rise to or support the phenomena to be explained. On a common understanding of what mechanisms are such that they can fulfil this role in explanation, a mechanism is comprised of entities that are causally connected by activities and organized in consistent ways to reliably produce some termination condition once the mechanism has been triggered; see also [8]. Details concerning the organized causal chain of entities and activities explain why some phenomenon occurs when and how it does. Mechanisms are fundamentally causal, and our models of mechanisms are models of causal relations and relata [9–11].

The role of mechanisms is not simply to model and explain what happens, but also to provide grounds for prediction about what would happen to a phenomenon of interest given specific

interventions on it; see also [12,13]. These interventions might be bottom-up [6], in that the intervention is directed at the components of the mechanism and the prediction concerns what effect this intervention will have on the overall phenomenon – how does the phenomenon change when specific components of its mechanism are altered? Interventions might also be top-down: such an intervention is directed at the overall phenomenon in question, and the effects are found at the lower level of the components – how do specific entities and activities change when the phenomenon they produce is interfered with or altered?

In this regard, mechanism models, the representations of mechanisms that figure centrally in scientific theories, are supposed to be able to serve two roles. One is as an explanation for the phenomenon in question: what are the relevant activities, entities and their organization? The second role is as a guide for predicting how mechanisms will respond to intervention on the entities and/or activities that comprise them: if we remove this entity, will the mechanism still work? Within what range must this activity take place in order for the mechanism to function?

Turning to medicine in particular, this use of mechanism models as a guide to predicting intervention outcomes is key when considering clinical medical practice. There are several identifiable stages at which clinical reasoning may involve mechanisms.<sup>1</sup> One stage is in diagnosis: mechanisms are often involved in identifying what has gone wrong with a patient such that a set of symptoms or dysfunctions is present; see also [14]. From knowledge of how the mechanisms for various human bodily functions work, a skilled practitioner can often isolate a particular mechanism that is failing to function normally, and identify a stage within that mechanism that is the problem: which entity is absent or dysfunctional, which activity is failing to take place or taking place abnormally?

Another closely related stage at which mechanisms can play a role is that of developing a treatment plan to address the identified disease or dysfunction. Having isolated a particular locus within a mechanism that is failing to operate normally, as a result of which the mechanism itself fails to operate normally, a practitioner can devise an intervention that targets the locus of failure within the mechanism. The idea is to intervene to restore the locus of failure within a mechanism to normal function with the aim of restoring the overall function of the mechanism itself, thereby alleviating the disorder or dysfunction. This use of mechanisms for prediction can also be used to disrupt mechanisms: by intervening on the mechanism that produces a set of problematic symptoms, mechanism models predict that the symptoms will thereby be eliminated or alleviated. Thus, mechanistic reasoning in clinical practice involves mechanisms in order to make predictions in addition to providing explanations.

EBM ranks the reliability provided by different kinds of evidence, where higher levels provide better evidence for or against the efficacy of a particular treatment. There are several versions of the hierarchy or levels of evidence. Some formulations of the

<sup>1</sup> Because this paper focuses on clinical practice, I am leaving out what is arguably the most important role for mechanisms in medicine, that of generating ideas for novel treatment possibilities in the first place, such that subsequent controlled trials can test these ideas. Uncovering mechanisms is often the first step in finding new treatments, but this step is prior to the kind of mechanistic reasoning in clinical practice I am discussing.

levels of evidence do not even include mechanistic reasoning as a level. The US Preventative Services Task Force lists ‘expert opinion’ as the lowest level of evidence, and does not include any category that may be plausibly construed as primarily involving mechanisms; see especially [15]. Some hierarchies rely on categories such as ‘background information’ [16]. Background information may include a wide variety of sources of information, but at least one such source of information plausibly includes the mechanisms that medical practitioners learn as part of their training. Other terms include ‘pathophysiologic rationale’, which is defined as ‘study and understanding of basic mechanisms of disease and patho-physiologic principles’ [17].

The Centre for Evidence-Based Medicine at Oxford includes mechanistic reasoning explicitly as a least reliable form of evidence. They provide the following definition of mechanism-based reasoning: ‘Involves an inference from mechanisms to claims that an intervention produces a patient-relevant outcome. Such reasoning will involve an inferential chain linking the intervention (such as antiarrhythmic drugs) with a clinical outcome (such as mortality)’ [18]. This is compatible with the two different stages I just outlined, although it does not distinguish between them. Thus, even though mechanisms are not counted as reliable evidence, they are often grouped together with other categories of evidence, and minimally characterized. This constitutes part of the motivation for this paper: a more detailed consideration of the roles mechanisms play in clinical practice reveals that much is missed by the hierarchies of evidence.

All the hierarchies rank meta-analyses of multi-site double-blinded studies as the gold standard for reliable evidence. This does not imply that mechanistic reasoning has led to inferior treatment in every case in which it has been used to generate interventions. Instead, the ordering in the evidence hierarchies reflects the fact that there are enough cases where mechanistic reasoning has failed to improve patient outcomes that the failure is more than merely an occasional exception – it is sufficiently widespread and systematic as to warrant a systematic response. In the next two sections, I will offer a more substantive case for EBM’s low estimation of mechanisms’ reliability as evidence. In addition to providing instances in which they fail, I also offer an explanation of *why*, precisely, they may fail in one major evidentiary role.

Even though the evidence in question often concerns statistical response patterns in very large populations, the ultimate goal in EBM clinical practice is finding the most effective treatment for individual patients. As such, there are two kinds of predictions that need to be distinguished. The first kind of prediction is statistical: given the treatment options evaluated in available studies, which treatment results in the best distribution of outcomes for the sample patient population? The second is singular: given the distribution of outcomes in the patient populations for several available treatments, what treatment is most likely to result in the best outcome for this individual patient? Each of these questions corresponds to a different form of mechanistic reasoning that may be used to answer it, and, as I will argue, it is only the first and not the second form that is rejected by EBM.

## Failures of mechanistic reasoning

I will examine two instances where mechanistic reasoning was supplanted with superior recommendations based on compiled

statistics about patient outcomes – in other words, two cases where mechanistic reasoning failed and EBM succeeded – and lay out likely reasons for the failure of mechanistic reasoning in these cases. These two cases illustrate distinct ways in which mechanistic reasoning may produce a poor prediction for the purposes of treatment. In the first case, a common surgery for knee problems, mechanistic reasoning failed for reasons that seem to come down to the fact that the mechanism for a healthy knee is simply different from the mechanism for an unhealthy one: the causal structure of the mechanism changes, and so restoring one locus within the mechanism is not sufficient to restore function to the original mechanism. In the second case, for infant vaccination procedures, mechanistic reasoning failed due to an interaction between two mechanisms that could not have been predicted based on our knowledge of either mechanism. In the next section, I will show how both of these failures in using mechanisms to predict intervention outcomes can be understood in terms of the evolved complexity of the mechanisms in question, and the necessarily limited information contained in the mechanism models used by practitioners to generate such predictions.

Knee lavage and debridement are common surgeries to treat knee osteoarthritis. Because osteoarthritis involves the accumulation of debris in the fluid in the knee joint, as well as a roughening texture of the interacting bone surfaces, these can be identified as loci of failure in the mechanism for normal knee function. The surgeries intervene on those loci to remove the debris and to smooth the edges of the bone in order to restore the locus to its normal condition, with the aim of thereby restoring the entire mechanism for normal knee function, of which the fluid and bone edges are components.

A recent meta-analysis performed for an EMB database of intervention recommendations found that there is ‘gold’ level evidence [19,20] that knee lavage and debridement do not improve knee function or reduce pain, and do not have any benefit over placebo surgery or non-surgical approaches like physical therapy. Research has not yet identified why these surgical interventions fail to improve knee function or alleviate pain. Debris and bone surface wear do appear to be part of the problem involved in knee dysfunction and in pain generation, yet targeting those parts of the mechanism to restore them to regular function does not restore the entire mechanism to regular functioning. This could be for several reasons, most likely that the model of the knee mechanism used as the basis for mechanistic reasoning was incomplete in a regard that was causally relevant to these interventions. This could cover a range of possibilities. There might be causal factors that come into play for damaged knees that simply are not factors in healthy knee function, such as causal relationships that are suppressed during healthy function but which are then expressed in pathological function; or there may be new causal relationships that simply do not exist in normal knee function. It could be that there are causal connections between the mechanism for knee function and some other mechanism(s) in the body. One way or another, intervening on activities or entities in the mechanism for a damaged knee changes that mechanism without thereby restoring it to its prior causal structure.

This does not mean that the mechanism models we use for normal knee function are inaccurate; no mechanism model can include all the actual, much less the potential, causal relationships in which such a mechanism may engage in a system as complex as

the human body. But it highlights how an accurate description of a normal mechanism may lack the information needed to make predictions about how that mechanism responds to various interventions when that mechanism is embedded in a complex causal environment.

The second example is the administration of prophylactic paracetamol with infant vaccination [21]. There are two mechanisms to consider in this case: the mechanism involved with vaccination for establishing immunity to a particular disease, and the mechanism(s) by which fevers of various origins can be brought down. There is nothing in these mechanisms that would lead a practitioner to suppose they would interact in a problematic way, especially since paracetamol is already used to reduce fevers when they do occur. Mechanistic reasoning suggests that the fever-reduction mechanism is causally downstream from the immunity-establishment mechanism, and thus not in a position to interfere with it. This makes it sensible to prescribe paracetamol preventatively to all infants, rather than waiting until they develop a potentially dangerous fever.

It turns out that the prophylactic administration of paracetamol interferes with the mechanism involved in vaccination, resulting in compromised disease immunity. The EBM recommendation now is that fever-reducing medication be only given if and when an infant develops a fever, not preventatively. The failure to predict the interference effect of paracetamol with antibody development is not a breakdown in either one of the mechanisms *per se*: it is not the case that the mechanism for establishing disease immunity via vaccination fails to achieve its end state when it also induces a fever. Instead, the side effect of the immunity-establishment mechanism constitutes the conditions (triggering the hypothalamus to raise body temperature) in which the second mechanism for fever reduction becomes relevant.

There is nothing in the mechanism models to indicate that these mechanisms interact in this particular way. Even using post facto ad hoc mechanistic reasoning, we cannot generate this 'prediction' of interference from reasoning concerning the two mechanisms, using the level of understanding of each mechanism that is at the disposal of practitioners making decisions with limited information in a clinical setting.

### Evolved complexity and mechanistic explanation versus prediction

In these examples, mechanistic reasoning did not result in the desired patient outcomes. And yet, we still have reason to think that the mechanism models from which practitioners are working when they engage in this kind of reasoning are legitimate. These failures highlight the gap between using a model of mechanisms to *explain* how a mechanism ordinarily works versus reasoning from that model to *generate predictions* about the outcomes of potential interventions in the mechanism, especially when it is not functioning normally. Providing a good explanation generally means having to leave out certain kinds of causal information, such as the multitude of causal interactions between various mechanisms. Explanations are often clearer, and certainly easier for practitioners to grasp and remember, when they include less of this information that can drown out information about the key entities and activities in the mechanism. Yet, this further information that complicates mechanistic explanation is often what is needed in order to

make accurate predictions about how a mechanism will behave under certain interventions. Furthermore, while medical mechanism models are generally accurate when those mechanisms are functioning, they fail to reflect what happens in the system when a new mechanism replaces the usual functional one, or when two or more mechanisms interact.

Mechanistic explanation and mechanistic prediction, in this context, come apart. These two examples illustrate a more general point: the failure of mechanistic reasoning highlighted by EBM is a result of applying that reasoning to a type of systemic complexity that may be persistently intractable to the use of mechanisms as a basis for predicting the outcomes of interventions, even though we have accurate models of mechanism subsystems of the system(s) in question.

In order to understand why mechanistic reasoning about the human body might fail in a systematic way, consider some generic structural features of causal mechanisms at work. Evolved systems in general tend to display certain kinds of causal complexity. One kind of complexity is involved in a wide range of equilibrium-maintaining systems. Another involves the layered character of many subsystems in organisms, the residue of earlier evolutionary stages that have been suppressed rather than eliminated. Two implications of this complexity are a lack of modularity of mechanisms and violations of causal faithfulness conditions.

Modularity is a property displayed by mechanisms when they can be intervened on independently of one another [1]. Modularity in causal systems means that ideal interventions, which only affect a single designated variable and leave others unchanged, can be performed. When modularity fails, only 'fat-handed' interventions can be performed. These interventions change more than one causal variable at the same time, consequently yielding much less information about which variable was responsible for the subsequent effects. In the human body and in other complex evolved systems, the mechanisms that support a variety of functions are non-modular, in that they include causal connections to other mechanisms from which they cannot be extricated without changing the causal structure of the mechanism in question, for example [22]. Some mechanisms, such as portions of the genome, are non-modular in the sense that they respond to interventions on a single locus by rearranging their causal structure elsewhere – such systems do not have the same causal structure before and after the intervention, thus violating modularity [22].

One might respond to this by claiming that the mechanisms are 'really' modular if one were to include more causal variables; the problem is that one has not fully elaborated all the entities and activities that constitute the mechanism in question. It is no doubt true that the model of a mechanism may leave out variables that only became causally relevant after the mechanism was intervened on or becomes dysfunctional. If those additional causal variables were included, the mechanism might provide more accurate predictions about how the system will behave under intervention. But here is the kicker for medicine: in many cases, if we were to include the causal variables that become relevant when intervening on a specific system but are not a part of the mechanism in normal functioning, we eventually end up including pretty much everything in the body.

The bodily mechanisms that malfunction, and on which we intervene in medicine to restore healthy function, are not modularly independent from other causal structures in the body. If we

want to add more variables to achieve modularity, then we end up in a situation where the entire organism is the first plausibly modular unit we encounter. While this might achieve modularity, it deprives mechanistic models of most if not all explanatory potential, and it renders the task of generating predictions impossibly complicated. There are solid epistemic reasons for modelling mechanisms as subcomponents of the entire organism, in spite of the fact that this is an imperfect method. This is one reason why mechanism models may be legitimately explanatory in medicine, without thereby providing the grounds from which to make accurate predictions about responses to interventions.

Consider another way in which precisely balanced mechanisms may ‘trick’ mechanistic reasoning. Causal faithfulness is the assumption, commonly made in contemporary causal methodology, for example [23], that the conditional and unconditional probabilities displayed by variables in a system reflect the underlying causal structure. This assumption is violated when, for instance, two variables exert an exactly equal but opposite causal influence on some third variable. These variables appear to exert no causal influence, but only because the influences they do exert are so precisely counterbalanced. In such cases, we do not know those additional variables are causally relevant – they are masked by the precisely balanced relationship and appear to be independent of the third variable.

It is extremely unlikely for many kinds of systems that the causal structure would be precisely balanced in just such a way as to mask causal relationships. But this kind of precise balancing is rampant in systems that are the product of evolution: violations of causal faithfulness may be the rule not the exception in complex evolved systems; see [24]. Any system, in this case an organism, capable of maintaining homeostasis against perturbation from the outside by dynamical internal responses will necessarily involve causal mechanisms that exert opposing influence of equal strength; and, these systems will often have robustly balanced causal influences: small perturbations will not ‘reveal’ that there are two opposing factors.

Once one of these mechanisms ceases to function properly, various other causal relationships will change, and previously ‘invisible’ causal relationships are suddenly revealed as important. In cases like this, intervening to restore a malfunctioning locus within the original mechanism will not restore the overall system to its original function, because the entity or activity on which the intervention is performed is now situated in a different mechanism than it was when the original mechanism was functional. Violations of causal faithfulness in a mechanism will therefore result in mechanism models that are not reliable guides to how the mechanism will behave under intervention or what the causal structure of that mechanism, or adjoining mechanisms, will be in conditions of dysfunction.

In general, in systems displaying evolved complexity, including but not limited to the human body, we should expect to find relevant systematic failures when using explanatory mechanism models as a guide for generating predictions about responses to possible interventions. This may occur because the mechanisms describing normal function fail to include additional causal influences that transpire only when the ordinary ones fail to operate, or may not include ‘hidden’ causal relationships that connect a mechanism to other mechanisms, or because the causal structure for one function is connected to other mechanisms in the body

from which it cannot be causally disconnected by intervention. We may have good models of the mechanisms at work in the human body without thereby having the knowledge necessary to use those models as a guide for treatment in medicine.

This explanation of why mechanisms should be expected to fail when used in this way for reasoning about a system like the human body renders more plausible the low status that EBM assigns to such evidence. There are biological considerations that justify the EBM claim that mechanisms are not a high-quality source of evidence in medicine, with the caveat that this applies to this particular usage of mechanisms as evidence. EBM advocates have pointed to failures of mechanistic reasoning as justification of EBM, *inter alia* [25], but have not accounted for *why* such failures occurred and, thus, why we should expect future failures of this sort. This section provides a rationale for why mechanistic reasoning fails systematically enough to warrant flagging it.

### Applying EBM guidelines: a different evidentiary role for mechanistic reasoning

We have now seen why we might have genuinely explanatory models of mechanisms in the body that nevertheless fail to provide the basis for reasoning to an effective treatment. EBM proposes replacement of mechanisms with recommendations derived from (ideally) large-scale randomized controlled trials (RCTs). However, the situation in clinical practice is more complicated than simply that of searching for EBM guidelines and then dutifully applying them. In this section, I will show that in order to apply these population-level recommendations to individual patients, mechanistic reasoning still provides the best available evidence for practitioners. The problems of reference class and of heterogeneous populations in particular necessitate the use of mechanistic reasoning in order to apply EBM recommendations in choosing a course of treatment for a patient with specific needs.

There are several implications of this. One is that even EBM should (as some advocates do) recognize that mechanistic reasoning has an important place in clinical practice, in terms of application of guidelines. Second, some of the resistance to EBM has indicated clinical expertise as playing a vital role in clinical practice in addition to guidelines for treatment. This resistance has too often been characterized as involving authority-based, ‘fuzzy’ or merely subjective approaches to treatment. I will show in the next section that the ‘clinical expertise’ defended by EBM critics can be understood in terms of the application of population-level EBM guidelines in the clinical setting using, among other techniques, mechanistic reasoning.

The problem of reference class is a long-standing issue in philosophy, and known in medicine, for example [26,27]. The general problem is that of determining the probability that a given event will occur, given that there are several different reference classes into which the event may fall, each of which assigns different probabilities to the event’s occurrence. Choosing which of several potentially conflicting reference classes an event should be construed as falling under has implications for our actions. For EBM, the problem of reference class takes the form of ascertaining the optimal treatment for a single patient that falls under several different reference classes, where distinct EBM recommendations exist for the different reference classes. This problem is

compounded by the fact that patients participating in the relevant studies are selected so as not to have other complicating factors, and are often from a comparatively narrow demographic slice of age, gender, race, etc.

Consider a hypothetical patient, an older woman with type II diabetes and breast cancer (see [28] for an example with respect to hypertension). There are recommendations for controlling diabetes, and there are recommendations for treatment options for various kinds of breast cancer at various stages. But there are not sufficient studies available to distil out EBM recommendations for patients with both a specific type of breast cancer and who are on a particular regimen for type II diabetes. The patient fits multiple reference classes, and there is no EBM-validated way to combine or bridge those classes. It is simply impractical to expect that there exist EBM-quality evidence on the best treatment methods for all possible combinations of illnesses.

How does mechanistic reasoning fill this gap? It is one way by which a practitioner can decide how to weigh potentially competing treatment recommendations to reach a decision for complicated cases, for example [29]. Between two potential breast cancer treatments, one might be more effective, but also involve a chemical pathway that a practitioner recognizes as involved in some of the problematic symptoms of type II diabetes. By comparing potential interactions between the mechanisms involved in both pathologies and potential treatments, clinicians can chart a path through the recommendations that is likely to work best for this individual patient. This is an evidentiary role for mechanisms that even some advocates of EBM have recognized as important to EBM practice:

Evidence-based medicine also involves applying traditional skills of medical training. A sound understanding of pathophysiology is necessary to interpret and apply the results of clinical research. For instance, most patients to whom we would like to generalize the results of randomized trials would, for one reason or another, not have been enrolled in the most relevant study. The patient may be too old, be too sick, have other underlying illnesses, or be uncooperative. Understanding the underlying pathophysiology allows the clinician to better judge whether the results are applicable to the patient at hand and also has a crucial role as a conceptual and memory aid. [17]

This clearly outlines several of the roles for mechanisms that I have described in this paper, both in terms of the need for simplified mechanisms used for explanation (the 'conceptual and memory aid'), as well as using mechanisms to address the problem of reference class. There is still, of course, a chance that unforeseen interactions between or structural changes to the mechanisms in question will result in suboptimal treatment. However, this remains the best *available* method of reasoning. Simply put, it may be suboptimal, but there is no better alternative at this stage of guideline application.

The second problem, that of heterogeneous populations, arises from the fact that patients may vary from one another in ways that are not tracked in the studies on which EBM relies but which do influence patient's response to a given treatment; see also [30]. It also arises when studies are performed predominantly on a narrow demographic with respect to potentially relevant factors like age, gender, ethnicity, etc. This notion of heterogeneity concerns the causal structure of internal mechanisms. Interventions on

causally heterogeneous populations mean that the patients in the population will respond differently to the same treatment; their bodies do not exhibit the same causal structure. Two patients may have distinct responses to the same medication because their bodies involve distinct mechanisms for reaction to that medication; see, for instance, [31], for heart treatment and gender differences. This results in misleading outcomes at the population level because different subpopulations will have distinct responses to the same interventions. When these subpopulations are considered as a single combined population, the overall statistics may not reflect the responses of any of the subpopulations involved, and result in misleading treatment recommendations for the entire population. For a clinical practitioner, this means that the overall statistics might not reflect the way in which this particular patient will respond. The fact that the overall population responds well to a given treatment X% of the time does not mean that any subpopulation has an X% chance of responding well to that treatment.

Heterogeneous populations can lead to Simpson's paradox cases, because the statistics from which one might infer probabilities for a given patient will reflect the compiled outcomes of distinct subpopulations with potentially very different response profiles to a given intervention. A case of Simpson's paradox was found in the treatments for kidney stones [32]. One treatment turned out to be better for both small and large stones considered separately, but when the two stone size subpopulations were grouped together, a treatment that was inferior in each subpopulation appeared to be more efficacious overall. In cases where a given treatment results in high recovery rates for one subpopulation, but worsens the condition for another subpopulation, the treatment will appear to be either mildly effective or detrimental for the overall population, depending on the relative sizes of each subpopulation.

How do mechanisms help fill this gap? While they cannot solve it completely, there are several potential avenues by which mechanistic reasoning can help guide application of EBM in cases of potentially heterogeneous patient populations. For instance, some practitioners work in areas where most of the patients come from a fairly narrow ethnic and sociocultural background, and display substantial demographic differences from the patient population(s) from which the EBM guidelines were derived. As such, experienced practitioners may come to find genuine differences in the way their patient population responds to given treatments compared with the EBM population. There may be issues like widespread food habits in a local area that a practitioner using mechanistic reasoning can infer are likely to interfere with an EBM recommended treatment. A treatment that is rated as less effective by EBM may be, for this subpopulation, more effective because it does not interact with these local complicating factors. Being aware of the potential for heterogeneous populations can also help clinicians recognize signs that a patient is not responding 'normally' (i.e. in the fashion consistent with the population-level guideline) to a given treatment and allow an earlier switch to a different treatment.

These problems of reference class and heterogeneous populations are necessarily part of the process of applying EBM guidelines derived from large populations to individuals in a clinical setting. But they have a specific upshot in terms of mechanisms: even if mechanisms are not the best available evidence for selecting possible interventions, they nevertheless serve an important

bridging function in applying EBM guidelines to patients with complicated health or demographic situations. In order to ascertain the appropriate reference class for a given patient, especially those presenting with more than one problem or who do not fit into the demographics of the population from which the EBM guideline is derived, mechanistic reasoning can be used to find at least some potential issues with complications. Reasoning about mechanisms helps bridge those gaps from population to single patient. While there will still be periodic failures, mechanisms remain the best *available* evidence.

This means that there is a distinctive role for mechanistic reasoning when applying EBM to choose an intervention from those available for a given patient, distinct from the role for mechanisms in generating possible treatment options in the first place. Thus, in spite of the fact that EBM is promoted as a preferable alternative to mechanistic reasoning, the relationship between the two is not one of genuine alternatives. EBM may supplant mechanistic reasoning when judging the population-level efficacy of particular treatments for the purposes of developing broad guidelines on treatment. But mechanistic reasoning may be required when applying those same guidelines in the clinical setting.

### Mechanisms and the role of clinical experience

Critics of EBM have resisted the formulaic aspect of EBM that they see as reducing medical practitioners to robots applying generic recommendations; see [33]. They offer clinical experience as a key element of medical practice that cannot be adequately captured in terms of EBM guidelines. Clinical experience is a very broad category that includes: the role of practitioner as social interpreter and guide of patient narrative; the role of values and personal goals of patients in choosing the right balance of risk and benefit; the use of prior personal experience with a given patient or similar patients in choosing treatments; and other forms of reasoning that are sometimes referred to in terms of cause and effect or the results of laboratory science [26,34–39].

The issues at stake here become easier to navigate when the tension can be reduced by clarifying the distinct roles for mechanistic reasoning I have argued for. As we have seen, at least some portion of what goes under the broad heading ‘clinical experience’ involves a form of mechanistic reasoning. When we clarify the role that mechanistic reasoning plays in the kinds of clinical experience that critics point to as left out of EBM guidelines, it becomes possible to get more specific about how such reasoning transpires.

Mechanistic reasoning can be one effective technique (although not necessarily the only one) for bridging the gap between population-level recommendations and individual patients. Use of this kind of reasoning is plausibly enhanced through clinical experience in ways that cannot be straightforwardly communicated via EBM-style guidelines. Consider Tanenbaum: ‘As interpreters, physicians draw on all their knowledge, including their own experience of patients and laboratory models of cause and effect’ [34]. ‘Laboratory models of cause and effects’ is what philosophers of science have so successfully construed in terms of mechanisms; essentially, Tanenbaum’s claim is that practitioners’ use of mechanisms is a major part of their clinical experience.

As another example, Tanenbaum’s ‘local knowledge’ [26] can be understood to include knowledge of how the local population of

patients responds to a given treatment, which may differ (sometimes dramatically) from the way in which a population resulting from multiple aggregated RCTs in multiple distinct geographical locations responded. Having such knowledge – that the response rates for a local population may differ from the response rate on which an EBM guideline is based – is a very important form of clinical knowledge that figures in mechanistic reasoning broadly speaking as a way of sorting patients into potential subpopulations based on mechanism differences in how they respond to various treatments.

This debate has been partially stymied by the way in which terminological boundaries have been set. The tension between EBM and clinical experience is not one of objective evidence versus subjective ‘fuzzy’ intuition. Construing the debate as pro-evidence versus anti-evidence is unfortunate for both sides: both sides are offering an account of evidence, where evidence is understood sufficiently broadly as that which provides rational warrant for a belief or plan of action. Likewise, construing this debate as ‘medicine: art or science?’ forces us to construe mechanistic reasoning in applying EBM guidelines as ‘art’. While mechanistic reasoning does not exhaust what might be grouped under the heading of ‘art’, identifying it accurately allows us to explain and assess what otherwise would have to be treated as somehow mysterious, intuitive, unteachable, etc. Mechanistic reasoning in applying guidelines can be performed better and worse; improving EBM guidelines alone will not address the skills needed for such reasoning.

Construing mechanistic reasoning in applying EBM guidelines as one strand of clinical experience shows again that critics of EBM are not advocating some kind of hopelessly subjective element in clinical decision making. Rather, they may be indicating further roles that mechanisms play beyond the guidance provided by EBM. It thus helpfully reframes the debate to consider the variety of ways in which mechanisms can be involved that fall outside the evidence hierarchies of EBM.

### Conclusion

The fact that EBM is pitched in part as a failure of mechanistic reasoning under intervention is extremely interesting [40]. Mechanisms are the subject of a great deal of contemporary research into explanations and causation; interventions are the bread and butter of the one of the most widely accepted accounts of causation. One consequence of this failure is that mechanisms, explanation and prediction come apart in potentially surprising ways. A mechanism may genuinely *explain* what is happening in a system like a healthy knee, while failing to provide the basis for *predicting* how that knee will behave under modifications to the system. The ways in which evolved complexity complicates the relationship between mechanisms, explanation and prediction – illustrated in the EBM case – have ramifications for any science that studies complex evolved systems. Even though the EBM hierarchy of evidence ranks mechanisms as low-quality evidence for the efficacy of a treatment, I have argued that reasoning based on mechanisms still has a distinct role when considered as evidence for the application of broad treatment guidelines to individual patients. Understanding clinical experience as not exhausted by, but centrally composed of, forms of mechanistic reasoning will allow for a more

fruitful investigation of the various roles that evidence plays at various stages of clinical decision making.

## References

- Woodward, J. (2003) *Making Things Happen*. New York, NY: Oxford University Press.
- Machamer, P., Darden, L. & Craver, C. (2000) Thinking about mechanisms. *Philosophy of Science*, 67 (1), 1–25.
- Glennan, S. (1996) Mechanisms and the nature of causation. *Erkenntnis*, 44, 49–71.
- Glennan, S. (2002) Rethinking mechanistic explanation. *Philosophy of Science*, 69 (S), S342–S353.
- Bechtel, W. & Abrahamsen, A. (2005) Explanation: a mechanistic alternative. *Studies in History and Philosophy of Biological and Biomedical Science*, 36, 421–441.
- Craver, C. (2007) *Explaining the Brain*. New York, NY: Oxford University Press.
- Russo, F. & Williamson, J. (2007) Interpreting causality in the health sciences. *International Studies in the Philosophy of Science*, 21 (2), 157–170.
- Andersen, H. (2011) The case for regularity in mechanistic causal explanation. *Synthese*, doi: 10.1007/s11229-011-9965-x
- Salmon, W. (1984) *Scientific Explanation and the Causal Structure of the World*. Princeton, NJ: Princeton University Press.
- Glennan, S. (2005) Modeling mechanisms. *Studies in the History and Philosophy of Science Part C*, 36 (2), 443–464.
- McKay Illari, P. & Williamson, J. (2010) Mechanisms are real and local. In *Causality in the Sciences* (eds P. McKay Illari, F. Russo & J. Williamson), pp. 818–844. New York, NY: Oxford University Press.
- Howick, J. (2011) *The Philosophy of Evidence-Based Medicine*. Oxford, NY: BMJ Books/Wiley-Blackwell.
- la Caze, A. (2011) The role of basic science in evidence-based medicine. *Biology and Philosophy*, 26 (1), 81–98.
- Nervi, M. (2010) Mechanisms, malfunctions and explanation in medicine. *Biology and Philosophy*, 25, 215–228.
- Atkins, D., Eccles, M., Flottorp, S., *et al.*; The GRADE Working Group (2004) Systems for grading the quality of evidence and the strength of recommendations I: critical appraisal of existing approaches. *BMC Health Services Research*, 4 (1), 38. doi: 10.1186/1472-6963-4-38
- Sackett, D. L., Straus, S. E., Richardson, W. S., Rosenberg, W. & Haynes, R. B. (2000) *Evidence-Based Medicine: How to Practice and Teach EBM*, 2nd edn. Edinburgh: Churchill Livingstone.
- Evidence-based Medicine Working Group (1992) Evidence-based medicine: a new approach to teaching the practice of medicine. *Journal of the American Medical Association*, 268 (17), 2420–2425.
- OCEBM Levels of Evidence Working Group (2011) *The Oxford 2011 levels of evidence*. Oxford Centre for Evidence-Based Medicine. Available at: <http://www.cebm.net/index.aspx?o=5653> (last accessed 19 June 2012).
- Reichenbach, S., Rutjes, A. W. S., Nüesch, E., Trelle, S. & Jüni, P. (2010) Joint lavage for osteoarthritis of the knee. *Cochrane Database of Systematic Reviews*, (5) CD007320.
- Laupattarakasem, W., Laopaiboon, M., Laupattarakasem, P. & Sumananont, C. (2008) Arthroscopic debridement for knee osteoarthritis. *Cochrane Database of Systematic Reviews*, (1) CD005118.
- Homme, J. H. & Fischer, P. R. (2010) Prophylactic paracetamol at the time of infant vaccination reduces the risk of fever but also reduces antibody response. *Evidence-Based Medicine*, 15, 50–51.
- Mitchell, S. D. (2008) Exporting causal knowledge in evolutionary and developmental biology. *Philosophy of Science*, 75 (5), 697–706.
- Spirtes, P., Glymour, C. & Scheines, R. (2000) *Causation, Prediction, and Search*. Cambridge, MA: MIT Press.
- Andersen, H. (2012) When to expect violations of causal faithfulness and why it matters. In *Philosophy of Science Assoc. 23rd Biennial Mtg*, pp. 1–21. San Diego, CA: PSA 2012 Contributed Papers.
- Sackett, D. L. & Rosenberg, W. M. C. (1995) The need for evidence-based medicine. *Journal of the Royal Society of Medicine*, 88, 620–624.
- Tanenbaum, S. (1995) Getting there from here: evidentiary quandaries of the US outcomes movement. *Journal of Evaluation in Clinical Practice*, 1 (2), 97–103.
- Feinstein, A. & Horwitz, R. (1997) Problems in the ‘evidence’ of ‘evidence-based medicine’. *American Journal of Medicine*, 103, 529–535.
- Tudor Hart, J. T. (1993) Hypertension guidelines: other diseases complicate management. *British Medical Journal*, 306, 1337.
- Sackett, D. L., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B. & Richardson, W. S. (1996) Evidence based medicine: what it is and what it isn’t. *British Medical Journal*, 312, 71–72.
- Williams, B. (2010) Perils of evidence-based medicine. *Perspectives in Biology and Medicine*, 53 (1), 106–120.
- Rathor, S. S., Yongfei, W. & Krumholz, H. (2002) Sex-based differences in the effect of digoxin for the treatment of heart failure. *New England Journal of Medicine*, 347, 1403–1411.
- Charig, C. R., Webb, D. R., Payne, S. R. & Wickham, O. E. (1986) Comparison of treatment of renal calculi by open surgery, percutaneous nephrolithotomy, and extracorporeal shockwave lithotripsy. *British Medical Journal (Clinical Research Edition)*, 292 (6524), 879–882.
- Cohen, A. M., Stavri, P. Z. & Hirsh, W. R. (2003) A categorization and analysis of the criticisms of evidence-based medicine. *International Journal of Medical Informatics*, 73 (1), 35–43.
- Tanenbaum, S. (1993) What physicians know. *New England Journal of Medicine*, 329 (17), 1268–1270.
- Greenhalgh, T. (1999) Narrative based medicine in an evidence based world. *British Medical Journal*, 318, 323–325.
- Goldenberg, M. (2005) On evidence and evidence-based medicine: lessons from the philosophy of science. *Social Science and Medicine*, 62, 2621–2632.
- Goldenberg, M. (2012) Innovating medical knowledge: understanding evidence-based medicine as a socio-medical phenomenon. In *Evidence-Based Medicine: Closer to Patients or Scientists?* (ed. N. M. Sitaras), pp. 11–28. Available at: InTech Open Science.
- Miles, A., Loughlin, M. & Polychronis, A. (2007) Medicine and evidence: knowledge and action in clinical practice. *Journal of Evaluation in Clinical Practice*, 13, 481–504.
- Braude, H. (2009) Clinical intuition versus statistics: different modes of tacit knowledge in clinical epidemiology and evidence-based medicine. *Theoretical Medicine and Bioethics*, 30 (3), 181–198.
- Worrall, J. (2007) Evidence in medicine and evidence-based medicine. *Philosophy Compass*, 2 (6), 981–1022.