

# A Challenge for Evidence-Based Policy

Adam La Caze<sup>1</sup> · Mark Colyvan<sup>2</sup>

Received: 11 March 2016 / Accepted: 25 April 2016 / Published online: 28 April 2016  
© Springer Science+Business Media Dordrecht 2016

**Abstract** Evidence-based policy has support in many areas of government and in public affairs more generally. In this paper we outline what evidence-based policy is, then we discuss its strengths and weaknesses. In particular, we argue that it faces a serious challenge to provide a plausible, over-arching account of evidence. We contrast evidence-based policy with evidence-based medicine, especially the role of evidence in assessing the effectiveness of medicines. The evidence required for policy decisions does not easily lend itself to randomized controlled trials (the “gold standard” in evidence-based medicine), nor, for that matter, being listed in a single all-purpose hierarchy.

**Keywords** Evidence · Evidence-based policy · Evidence-based medicine · Randomized controlled trials · Public policy

## 1 Introduction

Evidence-based policy has an attractive and reassuring ring about it. It sounds as though it should be contrasted with guesswork, ideologically driven policy, and media-reactive policy. It gestures towards accountability in government and comes with the promise of sound decisions based on scientifically respectable evidence. There has been a great deal of interest in evidence-based policy over the past few

---

✉ Adam La Caze  
a.lacaze@uq.edu.au

<sup>1</sup> School of Pharmacy, University of Queensland, Brisbane, QLD 4072, Australia

<sup>2</sup> Sydney Centre for the Foundations of Science, University of Sydney, Sydney, NSW 2006, Australia

decades, with many supporters and almost as many critics.<sup>1</sup> Much of the debate has focused on the extent to which methods employed within medicine are suitable for informing and assessing policy. Despite the considerable literature, the specific standards of evidence to which evidence-based policy subscribes, or ought to subscribe, remain unclear.

Evidence-based policy is presented as a way of deciding on policy. A viable approach to policy needs to make recommendations about how policy decisions should be carried out. We identify two recommendations that evidence-based policy makes about evidence. The first recommendation is that policy should be informed and evaluated by evidence (broadly construed). The second recommendation provides advice on what should be considered best evidence for policy decisions; ostensibly this recommendation provides standards of evidence for assessing evidence for policy. Presumably it is this second recommendation that distinguishes evidence-based policy from simply *good* policy. We evaluate the prospects of this second recommendation by explicitly comparing evidence-based policy with evidence-based medicine. We argue that the prospects of evidence-based policy adopting standards of evidence such as those employed in medicine are poor, for reasons that go beyond those frequently discussed in the literature. Evidence-based policy on this analysis is not, nor can it be, a prescriptive approach to methods for policy.

## 2 What is Evidence-Based Policy?

Medicine and policy share a focus on improving outcomes and the need for success under conditions of uncertainty. This leads to a similar desire to know what works in specific contexts. Evidence-based medicine came to prominence in the 1990s and provides a model for evidence-based policy. Early advocates of evidence-based medicine (EBM) felt that medical decision-making relied too heavily on “intuition, unsystematic clinical experience, and pathophysiologic rationale” (Evidence-based Medicine Working Group 1992). Proponents of EBM put forward the “hierarchy of evidence” as a tool for improving medical decision-making. EBM’s hierarchy of evidence is primarily a hierarchy of study designs.<sup>2</sup> The hierarchy places evidence gained through randomized controlled trials above other types of evidence, such as observational studies and the findings of basic science. The central premise of EBM is that decisions based on evidence from study designs higher up the hierarchy of evidence (e.g. randomized trials) are more reliable than decisions based on evidence from study designs lower down the hierarchy. The random allocation of participants into two experimental samples provides the principle epistemological distinction between randomized trials and observational studies. Observational studies measure outcomes in participants who are exposed to the treatment or factor of interest when

---

<sup>1</sup> Head (2010) provides a good overview (with an extensive reference list). Davies et al. (2000a) is an earlier, comprehensive treatment of evidence-based policy. Like this literature, this paper focuses on the role and merits of specific forms of *evidence* for informing policy. An example of some of the broader discussions regarding the role of science and values in policy can be found in Douglas (2009).

<sup>2</sup> EBM puts forward a number of hierarchies for different medical questions. We follow the literature on this topic and focus on the hierarchy provided for therapeutic decisions.

compared to participants who were not exposed. Observational studies are challenging to interpret because it is difficult, if not impossible, to entirely rule out the possibility that some factor that influenced the participant's exposure to the treatment of interest also influenced the participant's outcome. Random allocation eliminates this source of error.<sup>3</sup>

A common presentation of both evidence-based medicine and evidence-based policy is to focus on the problems that arise when evidence (broadly construed) is *not* incorporated into decision-making. There is no shortage of examples of medical or policy decisions based on expedience, flawed reasoning, or misplaced good will.<sup>4</sup> In this presentation, the descriptor "evidence-based" denotes the acceptance of some rather broad and somewhat uncontroversial epistemological standards with regard to the decision-making process. We label this *the first recommendation of an evidence-based approach*. According to this recommendation decisions should be based on evidence, and the outcomes of any decision should be assessed in light of further evidence. The first recommendation of evidence-based medicine/policy implies that the final decision will be taken from among the valid options. The set of valid options are those that a dispassionate observer would come to, based on the available evidence (assuming they have sufficient practical knowledge of the area and enough time to consider the evidence). On this view, adopting an evidence-based approach refers to a commitment to how deliberations will be conducted. In the context of evidence-based policy, this promise is neither contentious, nor anything particularly new or distinctive. In short, it is a commitment to *good* policy (or *informed* policy) and shares much with the rational decision-making model of the policy process (Nutley and Webb 2000, pp. 25–26).<sup>5</sup>

EBM avoids the charge of being prescriptively empty—medicine, by another name—by articulating a standard of evidence for medical decisions. To practice EBM is to adhere (at least in some sense) to the hierarchy of evidence. The move from recommending the use of evidence, broadly construed, to recommending specific standards of evidence for decision-making, we label *the second recommendation of an evidence-based approach*. There is considerable controversy over the extent to which EBM's hierarchy of evidence articulates an appropriate account of medical evidence.<sup>6</sup> Indeed, if the focus is on *all* medical decisions, then any simple application of EBM's hierarchy of evidence is difficult to justify. Medical

<sup>3</sup> As Worrall (2002) notes, randomization does *not* ensure all possible confounding factors are equally balanced between the experimental groups. Confounding due to an unlucky consequence of random allocation is, however, a different type of error to the selection bias that can occur in observational studies.

<sup>4</sup> Macintyre et al. (2001) provide examples from health policy where decisions that have ignored relevant empirical evidence have led to harm.

<sup>5</sup> Incremental or mixed incremental-rational models of the policy process emphasize the incremental and disjointed nature of policy progress. While evidence and arguments about evidence play a more diffuse role in these models, evidence (broadly construed) remains central (Nutley and Webb 2000, p. 28).

<sup>6</sup> Rawlins (2008) provides a good overview of some of the criticisms that have been raised from within medicine. For philosophical discussion, see Bluhm (2005), Cartwright (2007), Clarke et al. (2013, 2014), Grossman and Mackenzie (2005), La Caze (2008, 2009), Osimani (2014), and Worrall (2007). Cartwright (2009) and Roush (2009) discuss the use of randomized studies in policy and Montuschi (2009) discusses some of the broader problems of evidence in policy.

decisions are complex and rely on a broad range of evidence, as well as consideration of patient factors and clinical experience.<sup>7</sup> Nevertheless, EBM's hierarchy provides a suitable standard of evidence for a narrower range of medical decisions. Specifically, study designs listed higher in the hierarchy more reliably establish the *efficacy* of a treatment (see La Caze 2009 for discussion). A treatment is "efficacious" when the intended benefits of the intervention have been demonstrated in an experimental setting. Thus, EBM fulfills the second recommendation of evidence-based practice—albeit, for a more restricted set of medical decisions than is usually advertised.

The specific methodological commitments of evidence-based policy have not been so clearly articulated. Some working in health policy and education policy have explicitly adopted (or attempt to adopt) EBM's hierarchy of evidence,<sup>8</sup> but the level of adoption of EBM's hierarchy varies considerably across different policy areas. Criticisms of evidence-based policy and the (explicit or implied) adoption of medicine's hierarchy of evidence tend to focus on the feasibility and applicability of randomized studies in the policy setting.<sup>9</sup> Nevertheless, the general approach advocated in evidence-based policy often mirrors that of EBM: randomized trials are seen as the best source of evidence for establishing that an intervention works, and when randomized trials are not available or not suitable, decision-makers should move down the hierarchy to observational studies but heavily discount the evidence they provide. The field of econometrics, for instance, has developed increasingly sophisticated quantitative analyses of observational data, but the reliability of data coming from non-randomized study designs remains hotly contested (see, for instance, Leamer 1983; Hutton and Smith 2000; Angrist and Pischke 2010). In the absence of an alternative, EBM's hierarchy of evidence lurks as the *de facto* account of evidence for evidence-based policy.

### 3 Why Think that the Hierarchy of Evidence is Relevant to Policy?

First note the similarities between medicine and policy, at least at an abstract level. For example, consider the similar role played by motivating questions in medicine and policy prior to medical or policy decisions respectively. Which treatments reduce the risk of further heart disease following a heart attack? What approaches improve the educational outcomes for socially disadvantaged children? In both settings we are interested in judging the likely effects of an intervention in situations in which we neither control nor understand all the factors that may affect the outcome.

---

<sup>7</sup> The importance of attending to the human elements of health are emphasised by proponents of narrative medicine. See Greenhalgh (2012) for a discussion on the importance of avoiding the reduction medical decision-making to "mathematical estimates of the change of benefit and the risk of harm derived from high-quality research" (Greenhalgh 2012, p. 94).

<sup>8</sup> See for example, Macintyre et al. (2001) and Society for Prevention Research (2004), and for a more general discussion, see Davies et al. (2000a).

<sup>9</sup> Davies et al. (2000b) and Head (2008) discuss the difficulties of conducting randomized studies in some areas of public policy.

In medicine, positive results in well-conducted randomized trials provide a standard of evidence that the biomedical community takes to confirm the efficacy of a new treatment. Randomized trials are seen as necessary in this context because (1) there is often no shortage of plausible causal influences on patient outcomes *other* than the intervention under investigation and (2) there is a high degree of unexplained variation in the response to treatment, meaning that different outcomes may be observed in different treatment groups due to chance. Randomized trials attenuate the influence of non-intervention causal factors on the trial outcome, and allow the statistical control of chance by ensuring the study is large enough to minimize falsely attributing observed improvements to pre-defined error levels.

Cartwright (2011)<sup>10</sup> provides an argument schema for the role that a successful randomized trial plays in assessing whether an intervention is likely to work in clinical practice or a policy setting. It is tempting to think that the benefits of randomized trials in medicine will automatically be present in policy—the *structure* of the argument regarding evidence for an intervention is the same for both policy and medicine. However, the argument schema draws attention to issues that are sometimes overlooked in debates on evidence-based approaches—essential background assumptions are brought to the foreground. Cartwright’s argument schema makes the causal assumptions in arguments that the intervention will work in clinical practice or a policy setting explicit.

### Argument A

- A1.  $x$  plays a causal role in the principle that governs  $y$ ’s production [in the experimental setting].
- A2.  $x$  plays a causal role [in the practice/policy setting] as well as [in the experimental setting].
- A3. The support factors necessary for  $x$  to operate are present for some individuals [in the practice/policy setting].

Therefore,  $x$  plays a causal role [in the practice/policy setting] and the support factors necessary for it to operate are present for some individuals (Cartwright 2011, p. 222, modifications for context in square brackets).

The schema illustrates a number of important points. A strong argument for the intervention to work in the practice/policy setting requires a strong argument regarding the intervention working in the experimental setting (A1) *and* a good deal of causal knowledge to support premises A2 and A3. A strong argument that the intervention will work in the practice/policy setting requires that we know a great deal about how the intervention works, and how its effects may be influenced by factors present (or absent) in the setting of interest. The importance of this causal knowledge is often overlooked. Importantly, even when the conditions of Cartwright’s argument are met, it only follows that the intervention ( $x$ ) plays a causal role “for some individuals”. Decision-makers in policy and medicine typically want more. They want to be confident that the overall benefits of the

---

<sup>10</sup> See also Cartwright and Hardie (2012).

intervention will outweigh any costs and/or harms in the population that receive it. A good deal of causal knowledge is required to make this judgment.

While EBM's hierarchy of evidence provides a suitable standard for assessing whether new medical interventions work in the experimental setting, assessing whether the intervention works in the practice setting is more challenging. Nonetheless, the information provided by well-conducted randomized trials—in addition to the causal knowledge gained throughout the development of the intervention—help to inform these judgments. Despite the similarities between medicine and policy, we argue that some of the things that support randomized trials in medicine are not available in policy. We consider two main areas of disanalogy between policy and medicine. The first is in the sciences that underpin policy and medicine. In short, there is more causal knowledge of the sort needed to design, analyze and apply randomized trials in clinical medicine than there is in the social sciences. This difference affects our ability in the policy context to build strong arguments that the intervention will work. The difference in available causal knowledge between policy and medicine is sufficient to undermine the basis for evidence-based policy's second recommendation. The second area of disanalogy is in the questions of interest to policy and medicine. EBM's hierarchy of evidence and the associated methods are focused on comparing two to three relatively well-understood interventions. The question at the heart of many policy decisions is often much more open-ended.

## 4 Policy and Medicine: Important Differences

### 4.1 Differences in the Supporting Sciences

There are important differences in the degree to which biological and social sciences support the conduct, analysis, and interpretation of randomized studies. Moreover, these differences have been under-recognized in debates about evidence-based medicine and evidence-based policy. It is easier to decompose, develop, test and manipulate biological mechanisms than social mechanisms.<sup>11</sup> This is due to a host of factors, including: differences in the objects of the sciences, our inability to easily create meaningful experimental models of social mechanisms, and the differential resources put into social sciences as opposed to the life sciences. The outcome for policy sciences is that it is harder to get high quality causal information of the kind so important for designing and analyzing randomized trials and building good effectiveness arguments. Basic medical sciences that are focused on supporting the conduct and analysis of randomized trials are given the collective label *translational sciences*. Indeed, their importance is so well recognized that there is an international effort in the biomedical sciences to fund more and better translational science (Bornstein and Licinio 2011).

Causal knowledge in medicine comes from the basic biological sciences, such as biochemistry, immunology, pharmacology, physiology and pathophysiology. A

---

<sup>11</sup> See Steel (2004) for a discussion of social mechanisms.

critical part of many of these sciences is the use of models, often cellular or animal models, to represent or reproduce specific physiological or pathophysiological processes. Once established, these models permit isolating, testing and manipulating proposed mechanisms as well as assessing the influence of changes to background conditions. Much of what is established in the basic medical sciences is called “theory” in medicine. These theoretical claims are backed by domain-specific evidence provided by isolating, testing and manipulating the proposed mechanisms.

What is in question in clinical research is not the claims made in these basic sciences, but how the identified mechanism (or causal process) plays out in a clinical setting. Physiology tests claims about causal processes by isolating and manipulating the process. Clinical trials test claims about the outcomes of intervening on these mechanisms in patients. The science of clinical drug development is focused on bridging laboratory sciences like physiology with the clinical outcomes important to patient care. Early clinical drug development builds, tests and refines models of how the drug is distributed and removed from the body (pharmacokinetics) and how exposure is linked to clinical outcomes (pharmacodynamics).

What is learnt about the intervention during drug development feeds into the design, analysis and interpretation of randomized trials; it guides the choice of participants, the level of exposure, duration, effect measures, and size of the trial—to list just a few. Randomized trials can be conducted in the absence of this knowledge (or on the merest sketch of a causal process), but the interpretation can be difficult precisely because specific causal knowledge is absent. Importantly, all the information that is gained from basic and translational sciences also plays a role in building a strong argument that the intervention will work in the practice setting following a positive randomized trial. The better we understand how an intervention works and which factors influence “exposure” and “response”, the better an intervention can be employed in ways that promote effectiveness.

Briefly contrasting two examples might help make these ideas concrete. First, we consider the development of the drug imatinib for chronic myeloid leukaemia. Lyseng-Williamson and Jarvis (2001) summarise the key studies conducted throughout the development of imatinib. Imatinib selectively inhibits proliferation of specific leukaemic cells *in vitro*. Imatinib was also shown to eradicate tumours in animal models. These studies provide a starting point for estimating the target concentration of imatinib required to treat human patients. Clinical trials began after animal studies demonstrated a lack of significant toxicity at a wide range of doses. Druker et al. (2001) performed a dose-escalation study in which 83 patients with chronic myeloid leukaemia were assigned progressively higher doses of imatinib (25–1000 mg daily) and assessed for safety and efficacy. This study provided information on the dose–response relationship of imatinib and its relative safety. These findings, in addition to the findings of several additional clinical studies, informed the design of a large randomized clinical trial of imatinib compared to standard care for patients with chronic myeloid leukaemia. O’Brien et al. (2003) recruited 1106 patient with chronic myeloid leukaemia from 177 hospitals across 16 countries. Participants were randomized to receive imatinib or standard care. The primary endpoint for the trial was progression-free survival. At 18 months,

progression-free survival for imatinib and standard care was 92.1 and 73.5 % respectively. Imatinib is now routinely used to treat patients with chronic myeloid leukaemia.

The issue of class size reduction in education research provides a useful contrast (see Cartwright and Hardie 2012 for discussion). The Tennessee STAR study, discussed in Mosteller (1995), randomized pupils in the early grades of schooling to classes of 13–17 pupils or classes of 22–25 (with or without a teacher’s aid). The study found that smaller class sizes led to improved performance on standardized tests. Tennessee STAR was a large, well-supported study in educational policy. The benefits or otherwise of class size reduction has been studied and debated for many years. Most of the research is from smaller experimental studies and observational studies. An important impetus to the set-up and design of Tennessee STAR was a meta-analysis of this research conducted by Glass et al. (1982). Based on the success of Tennessee STAR, the state of California implemented class size reduction in 1996. Bohrnstedt and Stecher (2002) reports an evaluation of the program, which finds *no* conclusive relationship between class size reduction and student achievement.

It makes good intuitive sense that smaller class sizes may lead to improved academic outcomes, but the mechanisms by which this could occur are multiple and difficult to isolate. Unlike the situation for chronic myeloid leukaemia and imatinib, there are no well-established experimental models available for assessing the link between class size reduction and academic performance. This makes it difficult to identify and characterize the role of factors that are necessary to support the effectiveness of the intervention, as well as to characterize the “dose–response” relationship. The extent and kind of causal knowledge available to aid the transition of imatinib from experimental intervention to clinical practice was simply not available to support the transition for reduced class sizes from Tennessee STAR to the California’s class size reduction reform. The availability and type of causal knowledge underpinning Tennessee STAR and the key clinical trial for imatinib in chronic myeloid leukaemia is summarized in Table 1.

Building strong arguments that a medicine will work in practice after showing that it works in an experimental setting is not easy, and the challenges of translating promising basic science into improvements in patient care are well recognized.

**Table 1** Table of comparative features of the available causal knowledge for supporting randomized trials in medicine and policy

Feature	Imatinib	Class size reduction
Clear mechanism	Yes	No (multiple possible mechanisms)
Experimentally isolate and intervene on mechanism	Yes (in vitro and animal models)	No
Understand necessary support factors	Yes	No
Understand “dose–response” relationship	Yes	No



There are many examples in medicine in which the basic and translational sciences are less well understood than in the case of imatinib. Nonetheless the differences between what is typically available in medical contexts as opposed to policy contexts is striking, and sufficient to undermine the prospects of providing an account of evidence for policy that privileges randomized trials. Arguments for an intervention working in the policy or practice setting require a good understanding of the relevant causal processes (premises A2 and A3 in Cartwright's argument schema). The second recommendation of evidence-based policy requires this knowledge to be present, or likely to be present. While there may be examples in which such causal knowledge is available in the policy setting, the infrastructure that systematically supports the implementation of randomized-trial evidence in medicine is absent.

## 4.2 Randomized Trials Only Answer Some Questions Well

The benefits of randomized trials are clearest in testing the effects of pharmaceutical interventions. The design and analysis of confirmatory randomized trials are geared to rigorously answer a single question: Can the effects of the investigational intervention be distinguished from control? Many policy questions fall outside the narrow set of questions that are well answered by study designs listed high in EBM's hierarchy of evidence.

Consider, for example, a policy decision about importing some new agricultural product from a foreign destination. In order to properly assess the merits of such a proposal, the government considering the new importation policy needs to consider the relevant biosecurity risks: whether there is a risk of introducing agricultural diseases that will threaten or degrade the local production of the agricultural product in question. Such import risk analyses need to identify all the possible biological threats and the pathways for entry and establishment of these threats. Among other things, the risk analyses need to consider relative risks of biosecurity breaches along those pathways. That is, they need to determine whether the risks are significantly increased as a result of changing the importation policy.<sup>12</sup> This requires the development of a single policy out of a wide range of alternatives. Randomized trials are not well suited to this task, especially if (1) there are many variables of interest: randomized trials can only rigorously compare and/or control a small number of variables at any one time; and (2) some of the important risks under consideration are long-term.

There are further reasons why randomized trials are not always useful in policy. For many pharmaceutical interventions it is appropriate to assume that the mechanism by which the intervention has its effects is invariant for the duration of the intervention. It is assumed that countermoves are not possible. However, many public policy decisions involve other agents that may respond to the intervention in ways not anticipated. It is often necessary, therefore, to approach policy decisions

---

<sup>12</sup> See, for example, the recent policy decision by the Australian Government's Biosecurity Australia to allow the importation of Cavendish bananas from the Philippines. The relevant import risk analysis report and policy decisions are available from the Australian Government's Department of Agriculture Fisheries and Forestry website devoted to the matter: <http://www.daff.gov.au/ba/ira/final-plant/banana-philippines>.

not in terms of randomized trials, but in terms of a system of agents that may learn from our interventions and respond in ways not apparent in an initial trial.<sup>13</sup> Consider, a policy decision about the introduction of more stringent security measures at airports. Even if a randomized controlled trial were possible, it would not tell us anything about the potential terrorists' ability to learn about, and respond to, the new policy once it is in place. The bottom line is that many policy decisions, by their very nature, involve other agents, so require methods that take how agents may respond into account.<sup>14</sup> Here, modeling the evolution of behaviour patterns and the like might be more valuable than either randomized controlled trials or observational studies. In conservation biology, for instance, adaptive management is seen as important. It is ongoing management that is responsive to new data, including data about how the system in question responds to the new management decisions. There is continuous monitoring and constant reassessment of management strategies in light of the monitoring. This approach is explicitly dynamic and is seen as an improvement on standard static models of decision making, where the responses of either nature or other agents are not taken into account (Walters 1986).

## 5 Implications for Evidence-Based Policy

The second recommendation of evidence-based policy urges policy makers to focus their attention, where possible, on randomized trials, or other methods listed high on the hierarchy of evidence. We argue the relative lack of detailed causal knowledge undermines this recommendation. What are the implications of this argument for policy?

On the “supply-side”, there are implications for the *generation* of evidence for policy. In the right circumstances randomized trials can be very helpful for assessing whether a policy intervention is likely to work. But, among other things, the “right circumstances” require possessing a considerable amount of knowledge about the intervention and its likely effects in the setting of interest. This kind of knowledge comes from a diverse range of methods. The consequence is that the use of diverse methods answering a variety of questions should be supported for generating evidence for policy. The focus should be on the relevance of a particular question and the quality of the methods for answering that question, not whether or not the method is listed high in the hierarchy of evidence. Ironically, a move to

---

<sup>13</sup> Lindenmayer et al. (2012) provides examples of how carbon-trading policy can have unwanted downstream effects if the game-theoretic structure of the problem is not appreciated and Colyvan et al. (2011) discuss the importance of games against nature and adaptive management in conservation management decisions.

<sup>14</sup> This might not be so different from some areas of medicine where organisms are known to have responded to continued treatment in ways not anticipated by initial randomized controlled trials. Think, for example, of the way antibiotic-resistant *Staphylococci* evolved in response to clinical interventions of sequential narrow spectrum antibiotics. While game theory is usually, and most obviously, applicable to sentient and rational agents, many non-sentient systems behave as though they were agents responding in the game in question. See Skyrms (2004) for more on the evolution of cooperation in non-sentient cases via evolutionary game theory. Examples such as the antibiotic-resistant *Staphylococci* suggest that faith in randomized-controlled trials, even in medicine, is problematic.

single out randomized trials as a method worthy of more resources than alternative approaches undermines the relevance of the evidence provided by randomized trials.

On the “demand-side”, there are implications for *interpreting* the evidence available to policy makers. Policy makers, like clinicians, need to make decisions based on the *available* evidence, most often on tight timelines and under conditions of uncertainty. Unlike clinicians, policy makers do not have access to large numbers of randomized trials that are well-supported by basic and translational science. In this context, evidence-based policy’s second recommendation is poor advice. Much better advice is to develop an understanding of the policy setting, the process by which the intervention works, and the processes and factors in play in the policy setting that may support or inhibit the intervention working.<sup>15</sup> Decision-makers need to focus their attention on what is understood about the system and how an intervention might change that system. This approach requires the use of a diverse range of evidence.

## 6 Conclusion

The first recommendation of evidence-based policy is that evidence should play an explicit and central role in policy debates. This recommendation does not provide guidance on what form the evidence in question should take. This recommendation is not particularly controversial, nor particularly new, it does, however, have the virtue of being right. The first recommendation of evidence-based policy is silent on the appropriate standards of evidence and may even allow that the standards may shift according to context. The second recommendation of evidence-based policy requires a specific account of evidence. We suggest that the second recommendation should be jettisoned. Evidence-based policy cannot provide a prescriptive account of methods for policy. Evidence-based policy is better conceived of as a call for good policy: an aspiration for rational decision making rather than a blueprint for judging evidence.

**Acknowledgments** An early version of this paper was presented at the 2009 Sydney-Tilburg Philosophy of Science Conference: ‘Evidence, Science, and Public Policy’ held at The University of Sydney, 26–28 March 2009. We’d like to thank the audience at that conference for fruitful discussion and several very helpful suggestions. We are also grateful to the referees of this journal for helpful suggestions.

## References

- Angrist JD, Pischke JS (2010) The credibility revolution in empirical economics: how better research design is taking the con out of econometrics. *J Econ Perspect* 24(2):3–30
- Bluhm R (2005) From hierarchy to network: a richer view of evidence for evidence-based medicine. *Perspect Biol Med* 48(4):535–547
- Bohnstedt GW, Stecher BM (eds) (2002) What we have learned about class size reduction. California Department of Education, Sacramento

---

<sup>15</sup> This advice is exemplified in Cartwright and Hardie (2012).

- Bornstein SR, Licinio J (2011) Improving the efficacy of translational medicine by optimally integrating health care, academia and industry. *Nat Med* 12:1567–1569
- Cartwright N (2007) Are RCTs the gold standard? *BioSocieties* 2:11–20
- Cartwright N (2009) Evidence-based policy: What's to be done about relevance? *Philos Stud* 143(1):127–136
- Cartwright N (2011) Predicting what will happen when we act. What counts for warrant? *Prev Med* 53(4–5):221–224
- Cartwright N, Hardie J (2012) Evidence-based policy: a practical guide to doing it better. Oxford University Press, Oxford
- Clarke B, Gillies D, Illari P, Russo F, Williamson J (2013) The evidence that evidence-based medicine omits. *Prev Med* 57:745–747
- Clarke B, Gillies D, Illari P, Russo F, Williamson J (2014) Mechanisms and the evidence hierarchy. *Topoi* 33(2):339–360
- Colyvan M, Justus J, Regan HM (2011) The conservation game. *Biol Conserv* 144(4):1246–1253
- Davies HTO, Nutley SM, Smith PC (eds) (2000a) What works?: evidence based policy and practice in public services. Policy Press, Bristol
- Davies HTO, Nutley SM, Tilley N (2000b) Debates on the role of experimentation. In: Davies HTO, Nutley SM, Smith PC (eds) What works?: evidence based policy and practice in public services. Policy Press, Bristol, pp 251–276
- Douglas H (2009) Science, policy, and the value-free ideal. University of Pittsburgh Press, Pittsburgh
- Druker B, Talpaz M, Resta D et al (2001) Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. *New Engl J Med* 344(14):1031–1037
- Evidence-Based Medicine Working Group (1992) Evidence-based medicine: a new approach to teaching the practice of medicine. *J Am Med Assoc* 268(17):2420–2425
- Glass GV, Cahen LS, Smith ML, Filby NN (1982) School class size: research and policy. Sage Publications, Beverly Hills
- Greenhalgh T (2012) Why do we always end up here? Evidence-based medicine's conceptual cul-de-sacs and some off-road alternative routes. *J Prim Health Care* 4(2):92–97
- Grossman J, Mackenzie FJ (2005) The randomized controlled trial: gold standard, or merely standard? *Perspect Biol Med* 48(4):516–534
- Head BW (2008) Three lenses of evidence-based policy. *Aust J Public Adm* 67(1):1–11
- Head BW (2010) Reconsidering evidence-based policy: key issues and challenges. *Policy Soc* 29(2):77–94
- Hutton J, Smith PC (2000) Non-experimental quantitative methods. In: Davies HTO, Nutley SM, Smith PC (eds) What works?: evidence based policy and practice in public services. Policy Press, Bristol, pp 277–290
- La Caze A (2008) Evidence based medicine can't be.... *Soc Epistemol* 22(4):353–370
- La Caze A (2009) Evidence based medicine must be .... *J Philos Med* 34:509–527
- Leamer EE (1983) Let's take the con out of econometrics. *Am Econ Rev* 73(1):31–43
- Lindenmayer DB, Hulvey KB, Hobbs RJ, Colyvan M, Felton A, Possingham H, Steffen W, Wilson K, Youngtob K, Gibbons P (2012) Avoiding bio-perversity from carbon sequestration solutions. *Conserv Lett* 5(1):28–36
- Lyseng-Williamson K, Jarvis B (2001) Imatinib. *Drugs* 61(12):1765–1774
- Macintyre S, Chalmers I, Horton R, Smith R (2001) Using evidence to inform health policy: case study. *Br Med J* 322(7280):222–225
- Montuschi E (2009) Questions of evidence in evidence-based policy. *Axiomathes* 19(4):425–439
- Mosteller F (1995) The Tennessee study of class size in the early school grades. *Future Child* 5(2):113–127
- Nutley SM, Webb J (2000) Evidence and the policy process. In: Davies HTO, Nutley SM, Smith PC (eds) What works?: evidence based policy and practice in public services. Policy Press, Bristol, pp 13–42
- O'Brien S, Guilhot F, Larson R, Gathmann I, Baccarani M, Cervantes F et al (2003) Imatinib compared with interferon and low-dose cytarabine for newly diagnosed chronic-phase chronic myeloid leukemia. *New Engl J Med* 348(11):994–1004
- Osimani B (2014) Hunting side effects and explaining them: Should we reverse evidence hierarchies upside down? *Topoi* 33(2):295–312
- Rawlins MD (2008) De testimonio: on the use of evidence for decisions about the use of therapeutic interventions. *Lancet* 372(9656):2152–2161

- Roush S (2009) Randomized controlled trials and the flow of information: comment on Cartwright. *Philos Stud* 143(1):137–145
- Skyrms B (2004) *The stag hunt and the evolution of the social contract*. Cambridge University Press, Cambridge
- Society for Prevention Research (2004) Standards of evidence: criteria for efficacy, effectiveness and dissemination. <http://www.preventionresearch.org/advocacy/#SofE>. Accessed 25 Nov 2013
- Steel D (2004) Social mechanisms and causal inference. *Philos Soc Sci* 34(1):55–78
- Walters CJ (1986) *Adaptive management of renewable resources*. MacMillan, New York
- Worrall J (2002) What evidence in evidence-based medicine? *Philos Sci* 69:S316–S330
- Worrall J (2007) Evidence in medicine and evidence-based medicine. *Philos Compass* 2(6):981–1022