Title: **Evidence-Based Policy: Promises and Challenges**

Adam La Caze
School of Pharmacy, University of Queensland, Brisbane, Qld, 4072, Australia. Email: a.lacaze@uq.edu.au Ph: +61 (0)7 3346 1985 Fax: +61 (0)7 3346 1999

Mark Colyvan
Sydney Centre for the Foundations of Science, University of Sydney, Sydney, NSW, 2006, Australia. Email: mark.colyvan@sydney.edu.au

Abstract
Evidence-based policy is gaining 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 account of evidence. This account needs to be at least in the spirit of the hierarchy of evidence subscribed to by evidence-based medicine (from which evidence-based policy derives its name and inspiration). Yet evidence-based policy's hierarchy needs to be tailored to the kinds of evidence relevant and available to the policy arena. 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.

Key words: 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 decades, with many card-carrying supporters and almost as many critics[1]. Much of this debate has focused on the extent that methods employed within medicine are suitable for informing and assessing policy. As the debate has progressed, evidence-based policy has become more inclusive in its account of evidence, but the details of what evidence-based policy is, and what, if any, evidential standards it prescribes have become less clear.

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. In this paper, we outline what evidence-based policy must look like if it is to be a way of approaching policy in the same way that evidence-based medicine is a way of practicing medicine. We disambiguate two key recommendations that evidence-based practice makes about evidence in policy. The first recommendation is that policy should be informed and evaluated by evidence (broadly construed). And, the second recommendation provides advice on what should be considered best evidence for policy decisions; ostensibly this recommendation provides standards of evidence for assessing policy.

There are many positive reasons for accepting the first recommendation of evidence-based practice, but this recommendation alone does not distinguish what it is to practice evidence-based policy. While most agree that the standards of evidence provided by evidence-based medicine are not suitable for policy, alternative standards of evidence that provide the details for acting on the second recommendation of evidence-based practice have not been provided. Drawing analogies between evidence-based policy and evidence-based medicine, we argue that the prospects of providing a specific account of evidence for evidence-based policy are dismal (Section 4). 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?

Evidence-based practice first came to prominence in medicine,[2] and it provides a good starting point for considering the epistemological commitments of evidence-based approaches. Early advocates of evidence-based medicine (EBM) felt that medicine and medical decision-making relied too heavily on 'intuition, unsystematic clinical experience, and pathophysiologic rationale' (Evidence-based Medicine Working Group 1992), and introduced EBM as an attempt to make medical decisions more rigourous. 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.[5] 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 random allocation of participants into two experimental samples provides the

[1] Head (2010) provides a good recent overview (with an extensive reference list). Davies et al. (2000) is an earlier, comprehensive treatment of evidence-based policy.

[2] Brendan Reilly's (2004) claim illustrates the dominance of evidence-based medicine in medical decision making: 'anyone in medicine today who does not believe [in EBM] is in the wrong business'.

[5] EBM puts forward a number of hierarchies for different medical questions. We, as almost everyone else in the literature on this topic, focus on the hierarchy provided for therapeutic decisions.

principle epistemological distinction between randomized trials and observational studies. In observational studies the participants choose to receive (or undergo) the intervention or control—whether by active deliberation or serendipity. 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 and more justifiable than decisions based on evidence from study designs lower down the hierarchy.

'Evidence-based', however, has come to mean different things to different people, and as a result disagreements on the merits of evidence-based approaches often end in a stalemate. When critics of evidence-based approaches highlight the limitations of the specific account of evidence given by the evidence hierarchy, proponents reply by pointing to the benefits of basing decisions on evidence—any evidence—as opposed to ideology, power or privilege.[6] Distinguishing the two recommendations of evidence-based practice will help to avoid this cul-de-sac, and clarify the kind of contributions we might expect from evidence-based policy.

The first recommendation is a general method for decision-making, which applies to any approach to evidence-based practice (and much else besides). It has been summarized in the *Sicily statement on evidence-based practice*. The Sicily statement, written by advocates of EBM, both outlines the values of evidence-based decisions and provides a process for practitioners (Dawes et al. 2005, p. 3).

1. Translation of uncertainty to an answerable question
2. Systematic retrieval of best evidence available
3. Critical appraisal of evidence for validity, clinical relevance, and applicability
4. Application of results in practice
5. Evaluation of performance

This description of evidence-based practice is easily transferred to policy. Indeed, the first recommendation of evidence-based practice shares much with the rational decision-making model of the policy process (Nutley and Webb, 2000, pp. 25–26).[7]

While standards of evidence are implied in terms such as 'best evidence', the *Sicily Statement* does not prescribe a specific account of evidence for evidence-based approaches. In our view the *Sicily Statement* expresses a commitment to the first recommendation of evidence-based policy. According to this recommendation policy should be based on evidence, and the outcomes of any policy decision will be assessed in light of further evidence. The first recommendation of evidence-based policy implies that the final policy 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 (providing they have sufficient practical knowledge of the area and enough time to consider the evidence). The notion of 'evidence' in this recommendation is interpreted broadly.

There is no shortage of examples of policy decisions based on political expedience, false premises, flawed reasoning, or misplaced good will.[8] Use of the term 'evidence-based policy' denotes, in part, the acceptance of some rather broad and somewhat uncontroversial epistemological standards with regard to the decision-making process.

---

[6] For one example, see the exchange between Black (2001) and Donald (2001).

[7] 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).

[8] Macintyre et al. (2001) provide examples from health policy where decisions that have ignored relevant empirical evidence have led to harm.

Objective, empirical evidence should be used to underpin decisions. The available evidence may not determine policy, but policy, according to this first recommendation of evidence-based practice, will be consistent with the options suggested by the available evidence. On this view, evidence-based policy makes a recommendation about how arguments should be conducted. While this promise of evidence-based practice is not particularly contentious, nor is it anything terribly new or distinctive. In short, it is a commitment to *good* policy.

Does evidence-based policy prescribe specific methodological standards? EBM avoids the charge of being prescriptively empty—medicine, by another name—by articulating an account of evidence for medical decisions. To practice EBM is to adhere not only to the use of (any) evidence in decisions, but to adhere (at least in some sense) to EBM's hierarchy of evidence. The second recommendation—most clearly explicated in EBM, but implicit in many discussions of evidence-based policy—is that the evidence-based approach identifies and uses best evidence. In EBM, the hierarchy of evidence defines 'best' evidence. There is considerable controversy as to the extent that EBM's hierarchy of evidence articulates an appropriate account of medical evidence, and significant criticisms of the hierarchy of evidence have been raised.[9] Nevertheless, EBM, through its guidebooks and key publications, provides specific advice on how EBM is to be practiced.

The methodological commitments of evidence-based policy have not been so clearly articulated. Most advocates of evidence-based policy appear to endorse an account of evidence similar to EBM. Health policy (unsurprisingly) and education policy are two areas that have explicitly adopted (or attempt to adopt) EBM's hierarchy of evidence,[10] 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 to policy questions tend to focus on the feasibility and applicability of randomized studies.[11] EBM's hierarchy of evidence, however, provides clear advice if randomized trials are not suitable: move on to the next study design listed in the hierarchy. Most of the literature on evidence-based policy—even when recognizing the limitations of randomized studies—fails to provide any specific account of evidence for policy. In the absence of an alternative, EBM's hierarchy of evidence lurks as the *de facto* account of evidence for evidence-based policy.

3. The prospects for a single account of evidence for evidence-based policy

EBM's hierarchy of evidence and the associated methods are much more specific to the kinds of questions that arise in medicine than is typically appreciated. Moreover, even after restricting the domain to medicine, the hierarchy of evidence provides a suitable account of evidence for a remarkably narrow and qualified set of questions regarding the efficacy of medical interventions, especially pharmaceuticals (La Caze 2009). The hierarchy of evidence does not provide an account of evidence for all of medicine let alone for policy. We believe the breadth of policy decision-making, and the need to tackle many difficult and disfferent types of problems, precludes adopting any single

[9] Rawlins (2008) provides a good overview of some of the criticisms that have been raised from within medicine. For philosophical discussion, see Worrall (2007), Cartwright (2007), Grossman and Mackenzie (2005), Bluhm (2005), and La Caze (2008, 2009). 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.

[10] See for example, Macintyre et al. (2001), Society for Prevention Research (2004), and for a more general discussion, see Davies et al. (2000a).

[11] Davies et al. (2000b) and Head (2008) discuss the difficulties of conducting randomized studies in some areas of public policy.

methodological approach

We show that EBM's hierarchy of evidence is not a viable candidate for many policy questions, and we provide some arguments as to why the prospects for a specific account of evidence for evidence-based policy are poor. Rather than focusing on the feasibility of randomized trials in various policy situations (as much of the critical literature does), we clarify what it is that randomized trials and observational studies *do* and the *context* that allows them to do it. This analysis demonstrates the inapplicability of EBM's hierarchy of evidence to a great deal we are interested in with policy decisions. It also highlights difficulties for supplying any account of evidence suitable for the second recommendation of evidence-based policy. The extensive variety of methods, and the constant methodological innovation that takes place in research in the basic medical sciences (for instance medical biochemistry, physiology and immunology) provides a better model for methodology in policy.

The benefits of randomized trials are clearest in testing the effects of pharmaceutical interventions.[16] Randomized trials *confirm* the differential effects of an investigational intervention on a defined outcome when compared to standard treatment. 'Confirm' in this context should not suggest *prove*. Randomized trials provide confirmatory tests of pharmaceutical interventions in two senses: (i) randomized trials are given a regulatory role in assessing pharmaceutical interventions prior to marketing, and (ii) randomized trials have a relative superiority over alternative methods in testing the efficacy of pharmaceutical interventions (see La Caze 2009 for discussion).

Lewis B. Sheiner (1997) used the terms 'learn' and 'confirm' to distinguish between two goals in clinical drug development. Sheiner's terminology and message are helpful in the context of evidence-based policy. Sheiner uses 'learning' to capture inquiries that are focused on gaining a better understanding of the process or mechanism under investigation, and 'confirming' to capture inquiry focused on establishing that the expected outcomes of the process or mechanism eventuate.[17] Sheiner's paper was a response to a perceived over-reliance on randomized studies in the early phases of clinical drug development. Confirming studies are randomized, recruit moderate to large numbers of participants, involve a small number of comparators, measure a small number of outcomes, and are (typically) analysed using frequentist statistics. Learning studies, by contrast, may or may not be randomized, are typically smaller, involve a larger number of comparators, measure a large number of outcomes, and may benefit from alternative approaches to statistical analysis (for instance, Bayesian approaches).

Sheiner's rather simple and highly influential point—at least in terms of clinical drug development—is that due to the important role that both learning and confirming play, multiple methods and modes of analysis are required.

> Learning and confirming are quite distinct activities, implying different goals, study designs, and analysis modes. The understandable focus of commercial drug development on confirmation, as this immediately precedes and justifies regulatory approval, has led, in my view, to a parallel intellectual focus that slights learning. The predictable result […] is that clinical drug development is often inefficient and inadequate. (Sheiner 1997, p. 275)

It is important to select the method according to the question. A randomized trial *might*

[16] Successful randomized trials are the standard for approving new drugs, with new drugs requiring a significant number of randomized trials in order to receive regulatory approval. 'Success' is defined as the new intervention showing statistically (and, hopefully, clinically) significant benefits over standard treatment.

[17] For instance, in clinical pharmacology it is important to both *learn* how patient response varies with dose of a drug (i.e. the dose-response curve) and *confirm* that (on average) patients given the medication experience the change in clinical outcome that the dose-response curve suggests.

be the best design for a policy question focused on confirming the effects of an intervention. But, if the policy question requires a better understanding of the process or mechanism, then neither randomized trials nor the observational studies listed in EBM's hierarchy of evidence are likely to be much use. Indeed, important differences between policy and medicine mean that even when the policy question is focused on a confirmation-style question, features that underpin the success of randomized trials and observational studies in medicine are absent.[18]

Two factors underpin most randomized studies in clinical medicine and are frequently absent in the policy domain: (i) the depth and stability of the theoretical basis for pharmaceutical interventions, that is, our understanding of the causal process or mechanism, and (ii) the capacity to meaningfully isolate the intervention of interest when testing drugs. Randomized trials and observational studies are most useful when these factors are present. While the study designs in question may be used when one or both of these factors are absent, the utility of the studies is undermined and the focus of inquiry is often better placed on improving understanding of the processes by which the intervention works—an avenue of inquiry that requires alternative methods.

A strong theoretical focus on the medicine and how it is thought to work exists prior to conducting randomized trials. This is true in a general sense, in that more often than not much will be understood at the mechanistic level about the system the drug targets, and the kind of intervention the class of drugs under investigation will make. It is also true in a specific sense, in that much pre-clinical work has been conducted on the specific drug to show that the drug does indeed interact with the system as expected, that the drug can be formulated in a way that therapeutic concentrations can be maintained, dose ranging studies will have identified the appropriate dose to be tested, and the presence of dose limiting or serious side effects will have been investigated. This depth of knowledge is rarely available in policy contexts.

Theoretical background provides a basis for randomized trials in a number of ways. First, it permits randomized trials to be confirmatory. The randomized trial design and the statistical analysis of the results of a randomized trial are geared towards rigourously answering a single question: Can the effects of the investigational intervention be distinguished in a pair-wise comparison with control? Confirmatory questions typically only come to the fore once there is some understanding (or agreement) on questions such as how the intervention works and what kind of 'dose' or 'exposure' is required to gain the effects of the intervention. If the question is how an intervention works, or what level of exposure is required, there are better designs. Theory 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

[18] The observational study designs listed in EBM's hierarchy of evidence (cohort and case-controlled studies) don't fit neatly in to Sheiner's 'learn' or 'confirm' category—these study designs are used in epidemiology rather than clinical drug development. Calling observational studies 'confirmatory' risks too much confusion (there is a large literature discussing the relative merits of randomized versus non-randomized observational studies). That said, in the context of this discussion, the observational studies listed in EBM's hierarchy of evidence share more with randomized trials than other study designs. Observational study designs are also comparative, and while observational studies are not confirmatory in the sense the randomized studies are, they are designed to answer similar questions, namely: What are the effects of the intervention (or exposure) under investigation? Much of the statistical analysis that takes place in observational studies focuses on minimizing the bias that could have arisen due to the non-experimental nature of the study—that is, the possibility that some confounding factor influences both the likelihood that the participant exposes themselves to the intervention and the assessed outcome. The features present in testing pharmaceutical interventions underpin both randomized trials and observational studies. The following discussion focuses on randomized trails, but similar points can be made in relation to observational studies.

trials can be conducted in the absence of theory (or on the merest sketch of a theory), but the interpretation of such trials are unlikely to provide conclusive results—which, if the question is a confirmation-type question, is the point of conducting the trial in the first place.

Theory is just as important when it comes to applying the results of the trial. External validity (the ability to extrapolate the findings of the trial to non-experimental conditions) depends on theory. For example, knowing that an educational intervention had a particular effect in a set of U.S. high schools will not help educational policy development in the Australian educational system, unless there is some account of how and why the intervention worked in the U.S. system. Without an account of how the intervention brought about its effect, it is difficult to judge what influence the differences between the two systems might have on that effect.

The theoretical basis of testing pharmaceutical interventions supports the use of randomized trials in a second sense. The question at stake in tests of pharmaceutical interventions is first and foremost an empirical question, as opposed to a political or ethical question. Does the drug possess the expected benefits over harms when given to patients? Ethical and political positions are also important in medicine, but there is more consensus in medicine on what is considered a worthwhile health outcome as well as what is considered an appropriate way to bring about this outcome. In policy, the debate may well be occurring at the ethical or political level. Empirical evidence may play a role in buttressing the ethical or political argument, and evidence might be gathered to confirm that the stated objectives of the policy are achieved, but the policy is not decided by evidence alone. In some policy debates it is neither clear what the desired outcome should be nor what policies are best for bringing it about, without invoking substantial ethical and political theses.

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. Randomized trials and observational studies are not particularly good methods (and are often inferior to alternatives) when attempting to understand how an intervention might work as opposed to merely confirming that it does work. This is especially so when there is limited or highly contentious information on the process or mechanism under investigation, and when it is not possible to isolate the process under investigation.

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 importing government 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 import and establishment of these threats. Among other things, the risk analyses need to consider relative risks of biosecurity breeches along those pathways. That is, they need to determine whether the risks are significantly increased as a result of changing the importation policy.[29] It is not a comparison of two or three policy alternatives, but the development of a single policy out of a wide range of alternatives. While there is no doubt that such policy decisions are very well informed by the relevant science (feeding into the decision primarily via the import risk analysis in question), randomized trials and the other methods listed in the

[29] 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.

hierarchy of evidence are not well-suited to the decision that needs to be made.[30] The central policy question here is about risks and how they might present; assessing those risk requires a focus on the mechanisms by which biological threats may be introduced. Studies listed high in EBM's hierarchy of evidence simply do not answer the question of interest.

Other times the underlying causal process or mechanism could be well enough understood, but cannot be isolated in a way that a randomized trial could be conducted. For instance, randomized trials are unable to capture or assess situations in which counter moves are possible. 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 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.[31] Consider, for example, 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 might be more appropriately approached via game theoretic methods.[32] Here modeling the evolution of behaviour patterns and the like might be more valuable than either randomized controlled trials or observational studies.[33]

When the focus is on understanding the underlying process or mechanism, the basic medical sciences provide a better model for methodological advice. A great diversity of methods, models and experimental approaches are used within the basic medical sciences to improve our understanding of physiological and pharmacological mechanisms. A general methodology can be described, which involves testing hypotheses by systematically ruling out alternatives. But the specific methods vary considerably, and depend a lot on the question(s) under investigation and the availability of experimental models.[34] Establishing mechanisms in the social sciences presents many challenges—realistic experimental models are exceedingly difficult to come by—but work in this area is more likely to bear fruit than focusing narrowly on methods that are not fit-for-purpose.

4. Conclusion

We have distinguished two recommendations of evidence-based policy. The first

[30] For instance, a randomized trial to test the effectiveness of a proposed pathogen pathway into the country in question would involve compromising national biosecurity.

[31] Lindenmayer et al. (To appear) 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 game against nature and adaptive management in conservation management decisions.

[32] 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 if they were other 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.

[33] In conservation biology, for instance, adaptive management is seen as important. It is ongoing management that is responsive to new data. 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).

[34] Weber's (2005) discussion of the development of experimental models in relation to oxidative phosphorylation discovery provides an illustrative example.

recommendation is not particularly controversial, nor particularly new, it does, however, have the virtue of being right. The first recommendation of evidence-based policy suggests 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. On the first recommendation of evidence-based policy, the appropriate standards of evidence are open for debate and may shift according to context.

The second recommendation of evidence-based policy requires a specific account of evidence. Not only is the account provided by EBM ill-suited to policy, it is implausible that any single account of evidence will be able to cover the range of questions that arises in policy contexts. Evidence-based policy cannot provide a prescriptive account of methods for policy. Evidence-based policy is better conceived as a rallying call for good policy: an aspiration for rational decision making rather than a blueprint for judging evidence. We want the right tool for the job, not the best tool for some other job.

Acknowledgements
An earlier 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 very useful discussion, which lead to several improvements in the paper.

References

Black, N. (2001). Evidence based policy: proceed with care. *British Medical Journal*, 323: 275–278.

Bluhm, R. (2005). From heirarchy to network: A richer view of evidence for evidence-based medicine. *Perspectives in Biology and Medicine*, 48(4):535–547

Cartwright, N. (2007). Are RCT's the gold standard? *BioSocieties*, 2:11–20

Cartwright, N. (2009). Evidence-based policy: what's to be done about relevance? *Philosophical Studies*, 143(1):127–136.

Colyvan, M., Justus, J., Regan, H. M. (2011). The conservation game. *Biological Conservation*.

Davies, H. T. O., Nutley, S. M., and Smith, P. C., editors. (2000). *What Works?: Evidence Based Policy and Practice in Public Services*. The Policy Press, Bristol.

Davies, H. T. O., Nutley, S. M., and Smith, P. C. (2000a). Introducing evidence-based policy and practice in public services. In H. T. O. Davies, S. M. Nutley, and P. C. Smith (Ed.), *What Works?: Evidence Based Policy and Practice in Public Services* (pp. 1–12). Bristol: The Policy Press.

Davies, H. T. O., Nutley, S. M., and Tilley, N. (2000b). Debates on the role of experimentation. In H. T. O. Davies, S. M. Nutley, and P. C. Smith (Ed.), What Works?: Evidence Based Policy and Practice in Public Services (pp. 251–276). Bristol: The Policy Press.

Dawes, M., Summerskill, W., Glasziou, P., Cartabellotta, A., Martin, J., Hopayian, K., Porzsolt, F., Burls, A., and Osborne, J. (2005). Sicily statement on evidence-based practice. *BMC Medical Education*, 5(1):1.

Donald, A. (2001). Commentary: Research must be taken seriously. *British Medical Journal*, 323:278–279.

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.

Grossman, J. and Mackenzie, F. J. (2005). The randomized controlled trial: Gold standard, or merely standard? *Perspectives in Biology and Medicine*, 48(4):516–534.

Head, B. W. (2008). Three lenses of evidence-based policy. *The Australian Journal of Public Administration*, 67(1):1–11.
Head, B. W. (2010). Reconsidering evidence-based policy: Key issues and challenges. *Policy and Society*, 29(2):77–94.
La Caze, A. (2008). Evidence based medicine can't be.... *Social Epistemology*, 22(4):353–370.
La Caze, A. (2009). Evidence based medicine must be …. *Journal of Philosophy and Medicine*, 34:509–527.
Lindenmayer, D. B., Gibbons, P., Colyvan, M., Youngentob, K., Wilson, K., Hobbs, R. J., Possingham, H. P., Felton, A., and Steffen, W. (To appear). The potential for perverse outcomes for biodiversity from broad-scale plantation establishment in a carbon economy.
Macintyre, S., Chalmers, I., Horton, R., and Smith, R. (2001). Using evidence to inform health policy: case study. *British Medical Journal*, 322(7280):222–225.
Montuschi, E. (2009). Questions of evidence in evidence-based policy, *Axiomathes*, 19(4):425–439.
Nutley, S. M. and Webb, J. (2000). Evidence and the policy process. In H. T. O. Davies, S. M. Nutley, and P. C. Smith (Ed.), What Works?: Evidence Based Policy and Practice in Public Services (pp. 13–42). Bristol: The Policy Press.
Rawlins, M. D. (2008). De Testimonio: On the use of evidence about the use of therapeutic interventions. The Harveian Oration. http://www.rcplondon.ac.uk/pubs/brochure.aspx?e=262
Reilly, B. M. (2004). The essence of EBM. *British Medical Journal*, 329(7473):991–992.
Roush, S. (2009). Randomized controlled trials and the flow of information: comment on Cartwright. *Philosophical Studies*, 143(1):137–145.
Sheiner, L. B. (1997). Learning versus confirming in clinical drug development. *Clinical Pharmacolgy and Therapeutics,* 61(3):275–91.
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/StandardsofEvidencebook.pdf
Walters, C. J. (1986). *Adaptive Management of Renewable Resources*, MacMillan, New York.
Weber, M. (2005). *The Philosophy of Experimental Biology*. Cambridge Studies in Philosophy and Biology. Cambridge University Press, Cambridge.
Worrall, J. (2007). Evidence in medicine and evidence-based medicine. *Philosophy Compass*, 2(6):981–1022.