



What Constitutes “Good” Evidence for Public Health and Social Policy-making? From Hierarchies to Appropriateness

Justin O. Parkhurst & Sudeepa Abeysinghe

To cite this article: Justin O. Parkhurst & Sudeepa Abeysinghe (2016) What Constitutes “Good” Evidence for Public Health and Social Policy-making? From Hierarchies to Appropriateness, *Social Epistemology*, 30:5-6, 665-679, DOI: [10.1080/02691728.2016.1172365](https://doi.org/10.1080/02691728.2016.1172365)

To link to this article: <http://dx.doi.org/10.1080/02691728.2016.1172365>



© 2016 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 02 May 2016.



Submit your article to this journal [↗](#)



Article views: 712



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)

What Constitutes “Good” Evidence for Public Health and Social Policy-making? From Hierarchies to Appropriateness

Justin O. Parkhurst  and Sudeepa Abeysinghe 

Within public health, and increasingly other areas of social policy, there are widespread calls to increase or improve the use of evidence for policy-making. Often these calls rest on an assumption that increased evidence utilisation will be a more efficient or effective means of achieving social goals. Yet a clear elucidation of what can be considered “good evidence” for policy is rarely articulated. Many of the current discussions of best practise in the health policy sector derive from the evidence-based medicine (EBM) movement, embracing the “hierarchy of evidence” that places experimental trials as pre-eminent in terms of methodological quality. However, a number of problems arise if these hierarchies are used to rank or prioritise policy relevance. Challenges in applying evidence hierarchies to policy questions arise from the fact that the EBM hierarchies rank evidence of intervention effect on a specified and limited number of outcomes. Previous authors have noted that evidence forms at the top of such hierarchies typically serve the needs and realities of clinical medicine, but not necessarily public policy. We build on past insights by applying three disciplinary perspectives from political science, the philosophy of science and the sociology of knowledge to illustrate the limitations of a single evidence hierarchy to guide health policy choices, while simultaneously providing new conceptualisations suited to achieve health sector goals. In doing so, we provide an alternative approach that re-frames “good” evidence for health policy as a question of appropriateness. Rather than adhering to a single hierarchy of evidence to judge what constitutes “good” evidence for policy, it is more useful to examine evidence through the lens of appropriateness. The form of evidence, the determination of relevant categories and variables, and the weight given to any piece of evidence, must suit the policy needs at hand. A more

Correspondence to: Justin O. Parkhurst, Department of Global Health and Development, London School of Hygiene and Tropical Medicine, 15-17 Tavistock Place, London WC1H 9SH, UK. Email: justin.-parkhurst@lshtm.ac.uk

robust and critical examination of relevant and appropriate evidence can ensure that the best possible evidence of various forms is used to achieve health policy goals.

Keywords: Evidence; Health Policy; Hierarchy of Evidence; Appropriate Evidence

Background

Evidence Based Policy and the Hierarchy of Evidence

The introduction of the concept of evidence-based policy (EBP) has marked an important shift in modern political processes. While the health sector has particularly championed this idea (Berridge and Stanton 1999; Cookson 2005), similar calls to use evidence to guide policy are increasingly seen in other social policy realms as well (c.f. Davies, Nutley, and Smith 2000; MacKenzie 2000; Slavin 2008). National governments have further looked to institutionalise such approaches—with the UK, for instance, launching in 2013 a set of “What Works” centres explicitly designed around the health model (UK Government 2013).

While scholars grounded in the field of policy studies has increasingly critiqued the concept of a single or obvious “evidence base” to policy as a rhetorical device (Hammersley 2013), as ill-defined with multiple meanings (Cairney 2015), or as a “technocratic wish” in contrast to political realities (Lewis 2003), the concept still endures. It is not unusual to find in meetings, in policy discussions, or in government reports, repeated calls for more “evidence based policy”—with the aforementioned “what works centres” particularly illustrating how this discourse can be embodied in formal institutions. It is further common to hear individuals raise points such as “evidence can mean many things, like knowledge and not just research”: concepts which are now well known and have been thoroughly established and described—from Carol Weiss’ widely cited (in the academic literature at least) description of the multiple meanings of evidence use in the 1970s (Weiss 1977, 1979), to more modern comprehensive treatments of the subject such as that of Davies, Nutley and colleagues in the last decade (c.f. Davies, Nutley, and Smith 2000). In this article, we therefore recognise the importance of the policy studies work that has problematized the concept of EBP (and we have no grand ambition to pioneer a new mode of thought on evidence use in social policy), but we recognise the continued struggle to have these insights become established in policy circles. As such we propose an incremental step to construct a framework and language that might allow some of these key insights to be more understood or more easily applied by planners and programme actors who may otherwise risk continuing the same mistakes or “reinvention of the wheel” when it comes to thinking about evidence.

Changing thinking about evidence, however, requires recognition of the history of how the ideas are currently used in policy circles. As many authors have noted, the field of public health's embrace of the concept in particular evolved from the tradition of evidence-based medicine (EBM) (Berridge and Thom 1996; Petticrew 2013; Smith 2013a). Within biomedicine, EBM has generally been considered a "success" to the extent that the movement has regularised the way in which patients are assessed and treated. Partly this success has been due to the standardisation of clinical decision-making, providing physicians with a transparently objective and scientific method of choosing treatment options (Evidence-based Medicine Working Group 1992; Canadian Taskforce on the Periodic Health Examination 1994).

The answer to the question of what constitutes "good evidence" to guide *clinical practise*, is widely seen to lie in the "hierarchies of evidence". These hierarchies set out the process through which research can be evaluated, with the largest scale and most "objective" or "scientific" forms of evidence understood as inhabiting the "top" of the hierarchy. Top-level evidence is typically seen to result from research methods exhibiting key characteristics including: large and representative sample size, control for experimenter and participant bias; control for external variables; the study of a singular experimental variable; and value-neutrality (Merton 1973). It is understood that for clinical interventions, these factors are best constituted in the form of Randomised Controlled Trials (RCTs) (Chalmers et al. 1981). Non-experimental methods—such as case studies, observational data, or case-controlled studies—are seen as less useful forms of intervention research, due to their inability to control for confounding variables, and the greater potential for bias to be introduced as some stage in the research protocol (Borgerson 2009).

A number of sources have developed specific evidence hierarchies. The UK's National Institute of Health and Care Excellence (NICE), for example, produces recommendations which are awarded "grades" from "A" (recommendations being based directly on RCTs or meta-analyses of RCTs) to "D" (recommendations based upon expert opinion or inferences from upper-level studies) (NICE 2005). Similarly, the GRADE (Grading of Recommendations, Assessment, Development and Evaluation) criteria evaluates biomedical evidence, again judging evidence from RCTs as "high quality", with observational data as "low quality", and other methods as "very low quality" (Oxman and Group 2004). Other examples exist as well, including the The Strength of Recommendations Taxonomy (SORT) (Ebell et al. 2004), and the Centre for Evidence Based Medicine (CEBM) at the University of Oxford (OCEBM Levels of Evidence Working Group n.d.) While there are some variations, common to all of these approaches is the methodological superiority attributed to experimental evidence (and Annex 1 of Nutley, Powell, and Davies 2013 for further examples).

Challenges: "Good Evidence" for Policy?

EBM was seen to increase objectivity, transparency, and certainty in respect to professional practise. Since these are also ideal goals often espoused for policy-making,

it is unsurprising that the logic of EBM has underpinned current discussions of the use of evidence in policy. A natural implication of the embrace of evidence hierarchies is that “good” evidence for policy would also come from the top of these evidence hierarchies. A number of challenges arise, however, deriving from the question of whether the “good” evidence to guide clinical interventions is the same as the “good evidence” to guide policy decisions. For instance, even within health policy, economic or social factors—which may not be conducive to study via RCTs—will necessarily be implicit within policy concerns. Such issues surrounding evidence-informed policy are particularly salient when strength of evidence is discussed; which is often presented in terms of the idea of methodological quality, from a scientific (research community) perspective, rather than applicability from a policy (decision-makers) perspective (Lavis et al. 2003; Mitton et al. 2007).

Several authors in the public health community have warned against using evidence hierarchies exclusively to guide policy-making (c.f. Petticrew and Roberts 2003; Booth 2010). Writing in the *British Medical Journal*, for instance, Black has argued that EBP is “qualitatively different” to EBM, urging caution in the application of principles from clinical medicine to the realm of policy (Black 2001). It has been further noted that for most policy-making situations, the relevant considerations go beyond clinical and immediate health related issues, to involve areas of social, political or economic concern; or as Glaszou and colleagues succinctly put it, “different types of question require different types of evidence” (Glasziou, Vandembroucke, and Chalmers 2004, 39). Indeed, Petticrew and Roberts have argued that a typology based on the type of question being addressed (e.g. acceptability, effectiveness, satisfaction, etc.) is more appropriate for policy guidance than a single hierarchy (Petticrew and Roberts 2003). Indeed, drawing on political science and philosophical insights, Russel et al. further argue that: “Policy-making is the formal struggle over ideas and values” (Russell et al. 2008, 40), and criticise it as “naive rationalism” to assume evidence itself is value free and can be placed in hierarchies when it comes to decision-making. As a result, calls for methodological aptness and a context-based selection of evidence have emerged (Boaz and Ashby 2003; Petticrew and Roberts 2003; Dobrow et al. 2004).

Despite these calls, there remains a common use of language for policy to be “evidence based”, with a recurring embrace of experimental trials as “good evidence” to inform policy. Within the debates, there also can appear to be a false dichotomy between those who call for more evidence, and those who critique evidence itself as constructed and political (Krieger 1992)—with constructivist ideas often frustrating public health programme officers who typically require actionable information. As such, this paper aims to contribute to the debate in a critical, but pragmatic manner. We identify three disciplinary fields that underpin (often implicitly) many of the critical challenges levelled against evidence hierarchies. Drawing on Political Science (and policy studies, specifically), the Philosophy of Science, and Sociology (including medical sociology and the sociology of knowledge), we explore how each of these fields problematises the use of evidence

hierarchies for policy-making. However, we do this in order to develop ways forward to improve our understanding of what constitutes “good” evidence for public health goals. We share Petticrew and Robert’s concern with identifying a more appropriate use of evidence for policy, and we identify how each of the three disciplines drawn upon, in their own right, provide clues about how to ensure greater appropriateness of evidence for public health (and social) policy.

Discussion

Policy Studies: Decisions Involve More than Clinical Outcomes

The fact that there are a range of non-clinical outcomes that are often important to consider in policy debates (c.f. Petticrew and Roberts 2003; Glasziou, Vandembroucke, and Chalmers 2004; Booth 2010) has been one of the main criticisms of the idea of “evidence-based” policy, and has led some to shift to use of the term “evidence-informed” policy instead. The idea that health outcomes are but one of multiple important potential issues, however, would be a conceptual starting point for political science, and specifically for the field of policy studies, which takes it as a given that policy decisions involve choices between sets of possible outcomes (Lasswell 1990; Stone 2002) and where the allocation of social values and relative weights to multiple social, political, or economic concerns is understood to be an inherent feature of the political process (Easton 1971).

From a political perspective, then, two problems arise with the direct application of evidence hierarchies to guide policy decisions. First, as has been noted elsewhere, public health policy decisions typically involve choice between competing sets of concerns, and not just technical evaluations of effectiveness. While one social value guiding decisions will inevitably be the clinical effectiveness (or cost-effectiveness) of an intervention, other values such as social desirability and acceptability, or impact on individual liberties, human rights, and equity may all be valid considerations for public health actors (Petticrew and Roberts 2003; Clark and Weale 2012; Barnes and Parkhurst 2014). Yet for none of these are RCTs the correct form of evidence to measure their importance or scale. Prioritising evidence from experimental methods serves to obscure, rather than remove, political considerations—imposing a de facto political position that holds clinical outcomes of morbidity and mortality reduction (i.e. those things conducive to RCT evidence) above other social values.

Even when looking within health specific concerns, a second political challenge is that those interventions conducive to experimentation may not be a public health priority. Complex social or health systems interventions are often less suitable to experimentation, and as such, a focus on evidence from the top of a hierarchy may shift attention away from such issues. This may serve to medicalise public health if it prioritises treatment and individual level policies over efforts to address what are increasingly argued to be neglected public health concerns, such as the

social determinants of health (Marmot and Friel 2008; Smith 2013b), or the structural drivers of illness (see Auerbach, Parkhurst, and Cáceres 2011 in respect to HIV/AIDS).

However, the political realities of public health decision-making do not eliminate the importance of evidence. Petticrew and Roberts have noted that “different types of research question are best answered by different types of study”, (Petticrew and Roberts 2003, 528), but what becomes apparent from a political perspective, is a need to elucidate which questions are of relevance to a particular policy decision, in order to make those judgements of appropriateness. The implications for public health policy-making would be to emphasise the need to make the underlying values and competing decision criteria explicit—akin to what Schön and Rein, writing from a critical policy studies perspective, describe as a process of “frame reflection” (Schön and Rein 1994). Public health goals are not simply to increase clinical efficacy (which hierarchies of evidence are designed to assist), but must instead address multiple considerations including health equity, social acceptability, human rights, and social justice. These relevant policy concerns should be identified *ex ante*, in order to have transparency within the policy concerns at stake, and to better identify the relevant evidence bases that speak to those concerns.

Philosophy of Science: Generalisability and Evidence in Context

A second theme appearing in the literature critical of EBP is conceptually rooted in thinking about causality and generalisability within the Philosophy of Science. While some authors in this discipline share the political science concern that the technical language of hierarchies serves to obscure the political nature of policy-making (c.f. Goldenberg 2006), others have particularly noted that many public health and social policy concerns present external validity problems that experimental methods and meta-analyses are unable to address (Worrall 2010; Cartwright 2011; Cartwright and Hardie 2012).

At the core of these arguments is the recognition that the mechanisms through which an intervention works in one context may be very different, or produce different, results elsewhere; particularly when dealing with social or behavioural interventions (Pawson and Tilley 1997; Cartwright 2011). While experimental trials are designed to improve internal validity (to show the intervention actually had an effect), this says nothing about the external validity of the result. For biomedical interventions, external validity is not ensured by the trial design, but rather derives from expected similarities in human biochemistry or anatomy (Victoria, Habicht, and Bryce 2004). In social, behavioural, and health services interventions (which are increasingly the mainstay of public health planning), there fewer such guarantees, or alternative evidence is needed to justify the expectation of similar effects elsewhere.

It is true that the medical profession is increasingly seeing challenges to past assumptions of generalisability as well, such as in the growing interest in “stratified medicine”—in which individual patients are seen to respond quite differently to

treatments, thereby requiring much more tailored intervention strategies (Katikireddi 2015; Trusheim, Berndt, and Douglas 2007). Yet clinical treatments still are in seen as fundamentally different to the social realities of human interaction and behaviour that social policy (including much health policy) must address—whereby conscious individuals are continually reflecting and (re)constructing meaning around any intervention they are presented with, and where social contexts on which the mechanism of effect itself can lie may change over time and place (Pawson and Tilley 1997).

This challenge particularly affects meta-analysis, which often sit at the top of evidence hierarchies above individual RCTs, as the method can combine findings from multiple trials to evaluate intervention effect. Yet meta-analysis relies on an assumption that the same mechanism of effect exists across trial sites (and exists in the general population). Were a meta-analysis, however, to synthesise trials showing positive effects of an intervention in one setting, and negative effects in another, the conclusion might be “the intervention shows flat results”, when a more accurate (and more useful) conclusion for policy could be that “the intervention works for some groups in some contexts, and do not work for other groups in other contexts” (c.f. Pawson and Tilley 1997). An example might be an intervention of a cash transfer to prevent HIV—this could reduce HIV risk taking in a context where poverty leads people to rely on transactional sex, while increasing risk in a setting where increased wealth is associated with increased social (and sexual) networking (Parkhurst 2010). This does not mean all social interventions are unpredictable of course. Yet for many interventions in social policy and public health, there is much less certainty of predictable effect—and the changing and context-specific nature of social realities means that no amount of data gathered in a meta analysis can predict with certainty the same result in the future for many social policy issues. The health sciences often fall back on Bradford Hill’s famous criteria to judge causal effect (e.g. if it has a temporal relationship, a dose-response relationship, etc.) (Hill 1965); and these criteria can still be applied to explain whether a social intervention had an effect. But they do not say anything about mechanisms of causality, and therefore cannot answer whether we can expect the same causation elsewhere or at a later point in time. Other forms of knowledge are needed to justify such assumptions. As noted, in the clinical sciences it is the amassed knowledge of human anatomy and biochemistry that allow generalisations of causality (and which increasingly are showing limitations to generalisations as well). In economics it is evidence of market behaviour seen over centuries that point to sometimes quite predictable responses. In social and behavioural interventions, knowledge of mechanisms of effect may often have much less certainty, and these limits must be recognised.

Alternatives such as realist approaches have developed in response to the recognition that social context can determine the mechanism of effect for many interventions (Pawson and Tilley 1997; Pawson et al. 2005). In such situations, the appropriate evidence will not just be that which is measured in a trial, but also evidence of applicability or locally expected effect. Examples of such evidence (on

mechanisms in context) might include ethnographic studies, for instance, or local surveys—evidence types typically ranked particularly low in hierarchies. As Cartwright has explained “[f]or policy and practise we do not need to know ‘it works somewhere’. We need evidence for ‘it-will-work-for-us” (Cartwright 2011, 1401).

Sociology: Construction of Problems and Populations

The final discipline supporting critical reflections of evidence hierarchies is that of Sociology—particularly the traditions of medical sociology and the sociology of scientific knowledge. Sociological enquiry begins from the understanding that ill health (or good health) is not a purely biological occurrence. Patterns of health and illness are shaped through social categories of gender (Courtenay 2000; Doyal 2000), ethnicity (Krieger et al. 2003), geography (Gatrell and Elliott 2009), class and socio-economic disparities (Wilkinson 2002; Marmot and Wilkinson 2009), and other determining structures. An understanding of which kinds of evidence speaks to these issues can therefore help to improve public health outcomes (Nutley, Powell, and Davies 2012). This is increasingly recognised within the field of public health itself (Krieger 1992), but blind imposition of evidentiary hierarchies can serve to hinder, rather than enable, such a shift, by focusing the research and policy gaze on those strategies conducive to experimentation—rather than considering broader social-structural factors that are fundamental in the patterning of population health outcomes.

Sociologists also recognise that what counts as evidence—including how variables are constructed and chosen—is often an artefact of the context or culture within which it is produced (Bloor 1976). Science itself is not produced in a social vacuum, but is rather also a product of social realities and actions (Kuhn 1970; Krieger 1992). When applied to the field of health, medical sociologists have, for instance, explored how concepts like ethnicity, race or social class do or do not get adequately captured in much health research (Krieger, Williams, and Moss 1997; Collins and Williams 1999; Morrissey 2005). Critical examination of disease categories and concepts can therefore allow new ways to consider public health intervention approaches (Blaxter 1978; Imrie 2004). The current need to develop new approaches to address the social determinants of health would represent a contemporary example of this (Williams 2003; Marmot and Friel 2008).

From a sociologically informed perspective, public health actors can critically reflect on the population groups, data variables, and nature of health and illness categories utilised within bodies of evidence, to question how these constructions best serve their goals of improved population health, disease reduction, or health equity. It may be that those things technically easy to measure, quantify, or alter in experiments may not be the most appropriate constructions of health and illness to serve public health needs.

From Hierarchy to Appropriateness

The three disciplines presented above each highlight problems in applying a hierarchy of evidence to prioritise health policy decisions. They also, however, each point to ways to re-conceptualise a good use of evidence to ensure it is best aligned with the normative goals of public health. Policy makers need to identify the multiple criteria on which their decision is based, to address the contextual specificity of the interventions they aim to implement, and to consider if existing disease and population definitions suit the ultimate goals of public health improvement. An *appropriate* use of evidence, therefore, would be one which is transparent about the policy concerns at hand, which questions whether intervention effects will be expected in the target area, and which is critically aware of different ways to classify populations and health problems.

This does not mean hierarchies of evidence have no relevance. Rigour and quality will always remain important, but the measure of quality for different types of evidence will derive from the appropriate sciences that generate such evidence. Current hierarchies of evidence emphasise qualities that are appropriate for identifying intervention effect, and typically do not say anything about generalisability. Policy concerns will usually require additional types of evidence (not just evidence of intervention effect), will need to consider complex situations where simple causal relationships are not the norm, and will further require evidence of whether possible interventions will work in the desired setting. Multiple research methods—be they experiments, interviews, observations, etc.—will be needed, and each will be underpinned by its own standards of quality and validity. So for example, if public acceptability is an important policy consideration, evidence from survey research may be appropriate. Evaluation of survey quality would obviously include an assessment of statistical power, reliability, and internal and external validity (through consideration of sampling, sample size, triangulation, standardisation of delivery, etc.) (Moser and Kalton 1971). Observational or ethnographic research, on the other hand, may be useful to policy makers in understanding the cultural context that surrounds a certain policy option. These methods emphasise the importance of understanding processes through the perspective of participants (Hammersley and Atkinson 2007). Evidentiary rigour for these methods is therefore related to aspects such as researchers' immersion in the research context, validation by feeding back their findings to participants, and continued reflexivity of the researcher (Davies 2008).

Other examples abound, but ultimately, when selecting evidence, what is essential is for decision makers to firstly identify the types of information they need on which to base their decision (from their decision criteria) after which, the appropriate evidence can be judged and evaluated. Each research tradition comes with its own criteria for establishing rigour. Good evidence for policy shifts from following a single hierarchy to the question of whether that evidence is appropriate to the policy consideration and needs, with quality assessment derived from the relevant research tradition.

Summary

The calls for evidence to inform policy has been embraced in Public Health and across other social policy fields more broadly. Yet in the rush to be more evidence-based, there have been associated calls for the increased use of evidence hierarchies to guide selection of “good evidence” for policy-making. Such an approach has been widely critiqued from inside and outside the public health sphere, and remain a persistent challenge to public health planning. The use of hierarchies in this way has been described by Boaz and Ashby (2003) as focussing upon the “noise” (i.e. methodological strengths from a natural scientific perspective) produced by evidence, rather than the “signal” (message conveyed, and aims of, a particular piece of research or research field). As such, this can be counter-productive to achieving public health policy goals—particularly when those goals revolve around more than improving treatment efficacy.

To move the discussion forward, we have further developed the concept of “appropriateness” of evidence for public health policy based on insights from the fields of political science (policy studies), philosophy (of science), and sociology (of knowledge). Policy studies illustrates that appropriate evidence will be that which correctly speaks to the multiple decision criteria under consideration—and as such there is a need to render explicit the social concerns and values being considered. Hierarchies provide important ways to rank evidence in terms of intervention effect, yet intervention effect is typically only one of issues relevant to health policy decisions. The philosophy of science shows that appropriate evidence will consider the generalisability of pieces of evidence. Hierarchies typically are concerned with questions of internal validity, yet policy concerns must consider the similarity of causal mechanisms to be certain of local effect. Finally, the sociology of knowledge highlights how appropriate evidence that aims to achieve normative goals, will need to critically question the usefulness of existing classifications of populations and disease to do so.

In commenting on an earlier version of this paper, Kattikireddi makes the astute comment that promoting a framework of appropriateness (instead of hierarchies) risks substituting one normative framework for another—that of appropriateness (Kattikireddi 2015). He notes that in many cases policy goals are not pre-set, but the use of evidence—and the invocation of EBM/EBP—serves to set the goals for the policy to pursue. We acknowledge this as reflecting a broader phenomenon in which there is “co-production” through which policy values can set the boundaries of evidence, and evidence utilisation can set the boundaries of politics (c.f. Jasanoff 1987, 2004, 2011; Hoppe 2010). Yet even when uses of evidence appear to drive the delineation of policy goals, there is still a useful task to elucidate and reflect on values, as a critical constructivist position might consider. Bacchi, for instance, has promoted an approach to policy analysis that fundamentally considers the construction of problems (and possible alternatives) to explore the values and interests embedded in such constructions—asking questions around “what’s the problem represented to be” (Bacchi 2009). The idea of

evidentiary appropriateness can still be used within such an approach, however, questioning if evidence use best serves the goals that are eventually pursued in whichever way the problem is constructed.

As noted in the introduction to this article, however, the real target audience for a promotion of appropriateness are the policy-making and planning stakeholders who are principally motivated to utilise evidence for its ability to more effectively or efficiently achieve outcomes. Particular outcomes (or policy goals) are indeed political, constructed, and subject to debate. Nevertheless, the appropriateness framework requires those goals to be explicit in order to be able to judge when evidence proves useful to achieve whichever goals end up being pursued.

From a perspective of appropriateness, then, a set of strategic question can particularly be used to guide reflections on evidence by decision makers:

- (1) What are the policy concerns at hand (and is the evidence selected the most useful to address the multiple policy concerns at hand)?;
- (2) Are the data constructed in ways that best serve policy goals?;
- (3) Do we have reason to believe that the evidence is applicable to our local policy context?

These questions do not address all challenges of evidence use in policy-making, including the performative aspects of how evidence utilisation itself can dynamically delineate policy priorities or define what is seen to be policy relevant. Yet for policy makers and advocates of EBP (the champions of “what works”) evidence remains crucial. While, the quality of evidence will always be important, “good” evidence for policy from a lens of appropriateness becomes that which best serves public health needs, not that which best fits any single methodological criteria. It is argued that a better understanding of what constitutes “good” evidence for policy can allow past critical authors concerns to be incorporated, while providing a useful way forward for public health actors tasked with increasing evidence use in policy and planning.

Acknowledgements

The authors would like to acknowledge conceptual inputs from Arturo Alvarez-Rosete, Steffanie Ettelt and Benjamin Hawkins. This paper is an output from the Getting Research Into Policy in Health (GRIP-Health) research programme.

Disclosure statement

No potential conflict of interest was reported by the authors.

Authors' contributions

Both authors conceptualised the paper content. Individual sections were written and revised by both authors. Both authors reviewed approved the final manuscript.

Funding

Funding for the GRIP-Health research programme is provided by the European Research Council [Project ID# 282118]. The views expressed here are solely those of the authors and do not necessarily reflect the funding body or the host institutions.

Notes on contributors

Justin Parkhurst is a senior lecturer in Global Health Policy at the London School of Hygiene and Tropical Medicine. He manages the European Research Council supported 'GRIP-Health' programme (Getting Research Into Policy in Health), and his main areas of research relate to the political and institutional factors shaping evidence use in health policy. He has also conducted significant past work on HIV prevention, including work on HIV policy making and on the social and structural drivers of HIV risk. Correspondence to: Justin Parkhurst, Department of Global Health and Development, London School of Hygiene and Tropical Medicine, London, UK WC1H 9SH.

Sudeepa Abeyasinghe is a lecturer in Global Health Policy in the School of Social and Political Science, University of Edinburgh. Her research primarily relates to the governance of risk and uncertainty, and includes her recent book 'Pandemics, Science and Policy' (2015, Palgrave Macmillan). Correspondence to: Sudeepa Abeyasinghe, SSPS, University of Edinburgh, Edinburgh, UK EH8 9LD.

ORCID

Justin O. Parkhurst  <http://orcid.org/0000-0003-0831-6213>

Sudeepa Abeyasinghe  <http://orcid.org/0000-0001-6341-3417>

References

- Auerbach, J. D., J. O. Parkhurst, and C. Cáceres. 2011. "Addressing Social Drivers of HIV/AIDS for the Long-term Response: Conceptual and Methodological Considerations." *Global Public Health* 6: S293–309.
- Bacchi, C. L. 2009. *Analysing Policy: What's the Problem Represented to Be?* Frenchs Forest: Pearson Australia.
- Barnes, A., and J. O. Parkhurst. 2014. "Can Global Health Policy Be Depoliticised? A Critique of Global Calls for Evidence-based Policy." In *Handbook of Global Health Policy*, edited by G. Yamey and G. Brown, 157–73. Chichester: Wiley-Blackwell.
- Berridge, V., and J. Stanton. 1999. "Science and Policy: Historical Insights." *Social Science & Medicine* 49: 1133–8.
- Berridge, V., and B. Thom. 1996. "Research and Policy: What Determines the Relationship?" *Policy Studies* 17: 23–34.
- Black, N. 2001. "Evidence Based Policy: Proceed with Care Commentary: Research Must Be Taken Seriously." *BMJ* 323: 275–9.

- Blaxter, M. 1978. "Diagnosis as Category and Process: The Case of Alcoholism." *Social Science & Medicine Part A: Medical Psychology & Medical Sociology* 12: 9–17.
- Bloor, D. 1976. *Knowledge and Social Imagery*. London: Routledge & Keagan Paul.
- Boaz, A., and D. Ashby. 2003. *Fit for Purpose? Assessing Research Quality for Evidence Based Policy and Practice*. London: ESRC UK Centre for Evidence Based Policy and Practice.
- Booth, A. 2010. "On Hierarchies, Malarkeys and Anarchies of Evidence." *Health Information & Libraries Journal* 27: 84–8.
- Borgerson, K. 2009. "Valuing Evidence: Bias and the Evidence Hierarchy of Evidence-based Medicine." *Perspectives in Biology and Medicine* 52: 218–33.
- Cairney, P. 2015. *The Politics of Evidence-based Policymaking*. London: Palgrave Pivot.
- Canadian Taskforce on the Periodic Health Examination. 1994. "The Canadian Guide to Clinical Preventative Medicine." Ottawa: Minister of Supply and Services Canada.
- Cartwright, N. 2011. "A Philosopher's View of the Long Road from RCTs to Effectiveness." *The Lancet* 377: 1400–1.
- Cartwright, N., and J. Hardie. 2012. "Evidence-based Policy: A Practical Guide to Doing It Better." Oxford: Oxford University Press.
- Chalmers, T. C., H. Smith Jr, B. Blackburn, B. Silverman, B. Schroeder, D. Reitman, and A. Ambroz. 1981. "A Method for Assessing the Quality of a Randomized Control Trial." *Controlled Clinical Trials* 2: 31–49.
- Clark, S., and A. Weale. 2012. "Social Values in Health Priority Setting: A Conceptual Framework." *Journal of Health Organization and Management* 26: 293–316.
- Collins, C. A., and D. R. Williams. 1999. "Segregation and Mortality: The Deadly Effects of Racism?" In *Sociological Forum* 14 (3): 495–23.
- Cookson, R. 2005. "Evidence-based Policy Making in Health Care: What it is and What it Isn't." *Journal of Health Services Research & Policy* 10: 118–21.
- Courtenay, W. H. 2000. "Constructions of Masculinity and Their Influence on Men's Well-being: A Theory of Gender and Health." *Social Science and Medicine* 50: 1385–401.
- Davies, C. A. 2008. *Reflexive Ethnography: A Guide to Researching Selves and Others*. London: Routledge.
- Davies, H. T. O., S. M. Nutley, and P. C. Smith. 2000. *What Works? Evidence-based Policy and Practice in Public Services*. Bristol: Polity Press.
- Dobrow, M. J., V. Goel, and R. Upshur. 2004. "Evidence-based Health Policy: Context and Utilisation." *Social Science and Medicine* 58: 207–17.
- Doyal, L. 2000. "Gender Equity in Health: Debates and Dilemmas." *Social Science and Medicine* 51: 931–9.
- Easton, D. ([1971] 1953). "The Political System: An Inquiry into the State of Political Science." New York, NY: Alfred A. Knopf.
- Ebell, M. H., J. Siwek, B. D. Weiss, S. H. Woolf, J. Susman, B. Ewigman, and M. Bowman. 2004. "Strength of Recommendation Taxonomy (SORT): A Patient-centred Approach to Grading Evidence in the Medical Literature." *American Family Physician* 69: 548–56.
- Evidence-based Medicine Working Group. 1992. "Evidence-based Medicine. A New Approach to Teaching the Practice of Medicine." *JAMA* 268: 2420–5.
- Gatrell, A. C., and S. J. Elliott. 2009. *Geographies of Health: An Introduction*. Chichester: John Wiley & Sons.
- Glasziou, P., J. Vandenbroucke, and I. Chalmers. 2004. "Assessing the Quality of Research." *BMJ* 328: 39–41.
- Goldenberg, M. J. 2006. "On Evidence and Evidence-based Medicine: Lessons from the Philosophy of Science." *Social Science & Medicine* 62: 2621–32.
- Hammersley, M. 2013. *The Myth of Research-based Policy & Practice*. London: Sage.
- Hammersley, M., and P. Atkinson. 2007. *Ethnography: Principles in Practice*. London: Routledge.

- Hill, A. B. 1965. "The Environment and Disease: Association or Causation?" *Proceedings of the Royal Society of Medicine* 58: 295–300.
- Hoppe, R. 2010. "From "Knowledge Use" towards "Boundary Work": Sketch of an Emerging New Agenda for Inquiry into Science–policy Interaction." In *Knowledge Democracy Consequences for Science, Politics and Media*, edited by R. In 't Veld, 169–86. Heidelberg: Springer.
- Imrie, R. 2004. "Demystifying Disability: A Review of the International Classification of Functioning, Disability and Health." *Sociology of Health & Illness* 26: 287–305.
- Jasanoff, S. S. 1987. "Contested Boundaries in Policy-relevant Science." *Social Studies of Science* 17: 195–230.
- Jasanoff, S., ed. 2004. *States of Knowledge: The Co-production of Science and the Social Order*. London: Routledge.
- Jasanoff, S. 2011. "Constitutional Moments in Governing Science and Technology." *Science & Engineering Ethics* 17: 621–38.
- Katikireddi, S. V. 2015. "Reply to What Constitutes "Good" Evidence for Public Health and Social Policy Making? From Hierarchies to Appropriateness." *Social Epistemology Review and Reply Collective* 4: 51–5.
- Krieger, N. 1992. "The Making of Public Health Data: Paradigms, Politics, and Policy." *Journal of Public Health Policy* 13: 412–27.
- Krieger, N., J. T. Chen, P. D. Waterman, D. H. Rehkopf, and S. Subramanian. 2003. "Race/Ethnicity, Gender, and Monitoring Socioeconomic Gradients in Health: A Comparison of Area-based Socioeconomic Measures—The Public Health Disparities Geocoding Project." *Journal Information* 93: 341–78.
- Krieger, N., D. R. Williams, and N. E. Moss. 1997. "Measuring Social Class in US Public Health Research: Concepts, Methodologies, and Guidelines." *Annual Review of Public Health* 18: 341–78.
- Kuhn, T. S. 1970. *The Structure of Scientific Revolutions*. Chicago, IL: University of Chicago Press.
- Lasswell, H. D. ([1990] 1936). *Politics; Who Gets What, When, How*. Gloucester, MA: Peter Smith.
- Lavis, J. N., D. Robertson, J. M. Woodside, C. B. Mcleod, and J. Abelson. 2003. "How Can Research Organizations More Effectively Transfer Research Knowledge to Decision Makers?" *Milbank Quarterly* 81: 221–48.
- Lewis, J. M. 2003. "Evidence-based Policy: A Technocratic Wish in a Political World." In *Evidence-based Health Policy: Problems & Possibilities*, edited by V. Lin and B. Gibson, 250–9. Oxford: Oxford University Press.
- MacKenzie, D. L. 2000. "Evidence-based Corrections: Identifying What Works." *Crime & Delinquency* 46: 457–71.
- Marmot, M., and S. Friel. 2008. "Global Health Equity: Evidence for Action on the Social Determinants of Health." *Journal of Epidemiology and Community Health* 62: 1095–7.
- Marmot, M., and R. Wilkinson. 2009. *Social Determinants of Health*. Oxford: Oxford University Press.
- Merton, R. K. 1973. *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago, IL: University of Chicago press.
- Mitton, C., C. E. Adair, E. Mckenzie, S. B. Patten, and B. W. Perry. 2007. "Knowledge Transfer and Exchange: Review and Synthesis of the Literature." *Milbank Quarterly* 85: 729–68.
- Morrissey, M. 2005. "Ethnicity and Race in Drug and Alcohol Research." *Health Sociology Review* 14: 111–20.
- Moser, C. A., and G. Kalton. 1971. "Survey Methods in Social Investigation." London: Routledge.

- NICE. 2005. *Guideline Development Methods 11—Creating Guideline Recommendations*. http://www.nice.org.uk/niceMedia/pdf/GDM_Chapter11_0305.pdf
- Nutley, S., A. Powell, and H. Davies. 2013. "What Counts as Good Evidence?". London: Alliance for Useful evidence.
- OCEBM Levels of Evidence Working Group. n.d. "The Oxford Levels of Evidence 2". Oxford Centre for Evidence-Based Medicine. <http://www.cebm.net/index.aspx?o=5653>
- Oxman, A. D., and G. W. Group. 2004. "Grading Quality of Evidence and Strength of Recommendations." *BMJ* 328: 1490–4.
- Parkhurst, J. O. 2010. "Understanding the Correlations between Wealth, Poverty and Human Immunodeficiency Virus Infection in African Countries." *Bulletin of the World Health Organization* 88: 519–26.
- Pawson, R., T. Greenhalgh, G. Harvey, and K. Walshe. 2005. "Realist Review—A New Method of Systematic Review Designed for Complex Policy Interventions." *Journal of Health Services Research & Policy* 10: 21–34.
- Pawson, R., and N. Tilley. 1997. *Realistic Evaluation*. London: Sage.
- Petticrew, M. 2013. "Public Health Evaluation: Epistemological Challenges to Evidence Production and Use." *Evidence & Policy: A Journal of Research, Debate and Practice* 9: 87–95.
- Petticrew, M., and H. Roberts. 2003. "Evidence, Hierarchies, and Typologies: Horses for Courses." *Journal of Epidemiology and Community Health* 57: 527–9.
- Russell, J., T. Greenhalgh, E. Byrne, and J. McDonnell. 2008. "Recognizing Rhetoric in Health Care Policy Analysis." *Journal of Health Services Research & Policy* 13: 40–6.
- Schön, D. A., and M. Rein. 1994. *Frame Reflection: Toward the Resolution of Intractable Policy Controversies*. New York: Basic Books.
- Slavin, R. E. 2008. "Perspectives on Evidence-based Research in Education—What Works? Issues in Synthesizing Educational Program Evaluations." *Educational Researcher* 37: 5–14.
- Smith, K. 2013a. "Understanding the Influence of Evidence in Public Health Policy: What Can We Learn from the "Tobacco Wars"?" *Social Policy & Administration* 47: 382–98.
- . 2013b. *Beyond Evidence Based Policy in Public Health*. Basingstoke: Palgrave Macmillan.
- Stone, D. 2002. *Policy Paradox: The Art of Political Decision-making*. London: W.W. Norton.
- Trusheim, M. R., E. R. Berndt, and F. L. Douglas. 2007. "Stratified Medicine: Strategic and Economic Implications of Combining Drugs and Clinical Biomarkers." *Nature Reviews Drug Discovery* 6: 287–93.
- UK Government. 2013. "What Works: Evidence Centres for Social Policy." London: UK Cabinet Office.
- Victora, C. G., J.-P. Habicht, and J. Bryce. 2004. "Evidence-based Public Health: Moving beyond Randomized Trials." *American Journal of Public Health* 94: 400–5.
- Weiss, C. H. 1977. "Research for Policy's Sake: The Enlightenment Function of Social Research." *Policy Analysis* 3 (4):531–45.
- . 1979. "The Many Meanings of Research Utilization." *Public Administration Review* 39: 426–31.
- Wilkinson, R. G. 2002. *Unhealthy Societies: The Afflictions of Inequality*. London: Routledge.
- Williams, G. H. 2003. "The Determinants of Health: Structure, Context and Agency." *Sociology of Health & Illness* 25: 131–54.
- Worrall, J. 2010. "Evidence: Philosophy of Science Meets Medicine." *Journal of Evaluation in Clinical Practice* 16: 356–62.