

Bambra C. (2005)

[Reviewing the evidence: reflections from experience.](#)

Evidence & Policy: A Journal of Research, Debate and Practice, 1(2), 243-256.

Copyright:

This is a post-peer-review, pre-copy edited version of an article published in *Evidence & Policy: A Journal of Research, Debate and Practice*. The definitive publisher-authenticated version [insert complete citation information here] is available online at: <https://doi.org/10.1332/1744264053730752>

Date deposited:

24/01/2017



This work is licensed under a [Creative Commons Attribution-NonCommercial 3.0 Unported License](#)

Reviewing the evidence: reflections from experience

Dr Clare L Bamba

Abstract

Evidence-based policy (EBP), along with the ‘systematic review’, has recently emerged as a prominent strand within social science and public policy research. A number of papers have heralded this emergence but the vast majority are theoretical and concerned with how EBP research *could* be practised. This paper reflects upon the author’s experiences of identifying and reviewing evidence on the effectiveness of labour market interventions aimed at people with a disability or a chronic illness, to demonstrate some of the methodological challenges and practical problems that arise from the *actual* conduct of EBP research.

Introduction

The role of evidence in policy making has increasingly been emphasised by the current government (Davies and Smith, 1999; Walker, 2000; Boaz et al, 2002; Solesbury, 2002). For example, the ascendancy of an evidence rhetoric is apparent in government reports such as the 1999 White Paper on *Modernising government* (Cabinet Office, 1999) and the 2000 Performance and Innovation Unit report *Adding it up* (Cabinet Office, 2000), and also within more specific policy reviews such as the 1998 Acheson Independent Inquiry into Health Inequalities (Acheson et al, 1998). This interest in evidence has been reflected in an upsurge in public funding for evidence-based policy (EBP) research (Packwood, 2002) alongside the development of EBP research centres such as the EPPI centre (<http://eppi.ioe.ac.uk/>) and the ESRC’s Evidence Network (www.evidencenetwork.org/).

One of the key concerns of EBP research is discovering ‘what works’ (and also understanding why and for whom ‘it works’), through the assessment of whether policy interventions, both new and existing, are effective and efficient in achieving the desired policy goal(s) (Davies and Smith, 1999; Macintyre and Petticrew, 2000; Campbell Collaboration, 2001). The concern for effectiveness is based on the acknowledgement that resources such as time and financing are scarce, and that it is important to focus them on policies that are the most effective at producing the

desired outcome(s) (Davies and Smith, 1999; Macintyre and Petticrew, 2000; Judge and Bauld, 2001). EBP research is intended to provide the ‘rigorous, replicable, relevant, independent’ and high quality information required to form an ‘evidence base’ that can be used by policy makers to inform their decisions about the development and implementation of public policy.

BOX 1

EBP’s concern with effectiveness, and its favoured methodological approach – the systematic review (see Box 1) – have their major roots in evidence-based medicine (Klein, 2000; Macintyre and Petticrew, 2000; Campbell Collaboration, 2001; Macintyre et al, 2001; Boaz et al, 2002; Solesbury, 2002), a discipline that has long been practised in the field of clinical research as a way of gaining information about the effectiveness of different treatments and interventions. However, unlike evidence-based medicine, in which the majority of interventions under consideration tend to be singular, easily identifiable, and amenable to isolation and the experimental method (Cochrane Collaboration, 2001), the public policy interventions that EBP research addresses are likely to be multiple, hard to distinguish and immersed within a complex social, political and cultural context (Whitehead et al, 2000; Pawson, 2001b; Judge and Mackenzie, 2002). Consequently, while the original attractions of systematic review such as its rigour, comprehensiveness and replicability (Mulrow, 1994; Chalmers and Altman, 1995; Davies and Crombie, 1998; Macintyre and Petticrew, 2000; Glasziou et al, 2001) remain, or indeed become enhanced, their operationalisation within the policy field is far more problematic.

The practical issues surrounding the implementation of the systematic review method in the policy sphere have not yet been fully addressed by the emerging literature on EBP research. The majority of publications are theoretical and concerned with how EBP research *could* be practised; they do not discuss the methodological challenges and practical problems that arise from its *actual* conduct (Klein, 2000; Pawson, 2001a, 2001b; Petticrew, 2001; Boaz et al, 2002; Packwood, 2002; Sanderson, 2002; Solesbury, 2002; Young et al, 2002; Wallace et al, 2004). This has given rise to the impression that such research is straightforward and uncomplicated, requiring only the technical application of the methodology in order to get the desired results. The

reality can be very different. This paper reflects on EBP research by the author on the effectiveness of labour market interventions aimed at people with a disability or a chronic illness (see Box 2), to demonstrate some of the methodological and practical challenges posed by the analysis of public policy interventions: policy definition; policy comparability; policy isolation and evaluation context; evidence location; and evidence design.

BOX 2

Policy Definition

Traditionally, the systematic review method requires the review question to be narrow, focused, and concerned with a very limited number of interventions and outcomes (Chalmers and Altman, 1995; Davies and Crombie, 1998; Campbell Collaboration, 2001; Glasziou et al, 2001; NHS CRD, 2001; Petticrew, 2001). This is intended to draw boundaries around the review and is used to determine the selection of relevant studies for inclusion (Davies and Crombie, 1998). However, public policy questions intrinsically tend to be broad and more open in their remit. For example, they may ask ‘does a reduction in benefit payments increase the employment rate?’ or ‘what is effective in helping women with disabilities gain employment?’ They are rarely narrow and specific in their focus, and they seldom involve one intervention or the influence of only one factor. The use of a very tight or technical formulation for the review will not meet the needs of policy makers because the focus will be too restricted. For example, ‘how effective are UK anti-discrimination policies on helping disabled and chronically ill people gain employment?’ This review question only offers a perspective on the effectiveness of one policy, in one country. It does not address issues of *why* this policy is or is not effective, or whether similar policies in other countries have the same effects. What policy makers really want to know in this instance is which of a given number of policy options for helping disabled people back into work are effective, and which is the most effective (Petticrew et al, 2004; Whitehead et al, 2004). Formulating a single ‘tight’ question may be clearly desirable from a methodological viewpoint, to keep the review manageable by avoiding unnecessary relevancy screening and the assessment and comparison of multiple interventions. However, it fails to meet the policy maker’s need. In such circumstances an alternative ‘policy’ approach will need to be developed in which the

policy question is broken down into its constituent parts, some of which may be amenable to inquiry and some not.

Policy Comparability

Related to the problem of policy definition is that of policy comparability. There has been a lack of an evaluative culture surrounding social interventions, particularly in the UK, (Oakley, 1998; 2000) and evidence is not available about the impact of a variety of social policies and interventions. In order to broaden the available evidence base, EBP research often compares similar policies from different countries both in order to get a larger quantity of evidence and to see how different styles of intervention may have an impact on effectiveness. For example, both the UK and the USA have anti-discrimination legislation in place around disability (the 1995 Disability Discrimination Act and the 1990 Americans with Disabilities Act) but the UK's legislation is less rigorous and extensive. It is useful for EBP research to compare these two approaches so that a judgement can be made about whether the intervention type (anti-discrimination legislation) is effective, which intervention style is effective (lenient, strict or both) and which is the more effective (lenient, strict or neither). Comparison clearly offers a number of benefits to EBP research (not least as a substitution for the experimental method (Marsh and Mackie, 1995), but it is not without its drawbacks: the policy interventions may not be truly comparable because of differences in their approach, context (historically, culturally or geographically) and goals (Pawson, 2001a; 2001b; 2002). Furthermore, the effectiveness of a policy intervention in one setting/environment does not guarantee that it can be transferred or replicated in another. The effectiveness of a policy may be reliant as much on the context in which it operates as on its constitution, mechanism and implementation.

Policy Isolation and Evaluation Context

The isolation of the policy and its effects/effectiveness is a similarly problematic issue for the conduct of EBP research. Policy interventions are deeply contextualised by their historical, cultural and geographic background, and by the wider policy environment in which they are implemented. Policies do not work in a vacuum; they are put in place within a pre-existing policy context and work alongside, either in a supporting or oppositional role, other policies. It is therefore extremely difficult to untangle the impact of individual interventions. For example, in the UK, the New

Deal for Disabled People and the Disabled Person's Tax Credit were introduced at a similar time (1998 and 1999 respectively), with the shared goal of improving the assistance offered to unemployed disabled people. It is hard to determine with any certainty that perceived outcomes for this population were the result of the New Deal intervention (X) or the tax credit (Y) or an entirely different, unnamed intervention (Z). The introduction of similar policies targeted at the same groups, in the same geographical areas, in the same time period makes the evaluation of the results of any one intervention particularly difficult to identify (Walker, 2001). Furthermore, it is difficult to assess whether the claimed effects of a policy intervention would have happened anyway, even if it had not been implemented (Judge, 2001). For example, those New Deal participants who found work might have done so even without the help of the scheme through, for example, a general rise in the overall employment rate (Bambra et al, in press). It is very hard to isolate cause and effect when examining the effectiveness of policies at this macro level.

In a commentary on the rise of EBP under the post-1997 New Labour government, Walker (2001) highlights the evaluation context as a further barrier to establishing the effects of specific social policy interventions. While New Labour were initially keen to pilot and evaluate policies and programmes ahead of implementation, the evaluations were often undermined, incomplete or rushed (Walker 2001). This led to a lack of clarity in establishing their effectiveness. For example, the New Deal for Lone Parents (prototype pilots) evaluation was undermined because the policy went nationwide during the evaluation period, leading to the possible contamination of the original control groups. Of more direct relevance to the social intervention focused upon in this paper, the evaluations of the New Deal for Disabled People (personal adviser pilots) were incomplete and somewhat rushed. Initially, these evaluations were planned to last for a period of two years, and were intended to guide the decision on whether the policy should be enacted nationally. However, in the event, the decision was taken on the basis of interim results, and the policy expanded after less than one year. Walker comments that the evaluations were further restricted in their ability to establish the employment and employability outcomes of the intervention because of a smaller than anticipated sample size and a lower than expected programme participation rate.

Evidence Location

Evidence-based policy seeks to provide evidence about the operation, impact, outcomes and effectiveness of public policy interventions. The rationale is that this evidence will then be used to inform public policy formulation and implementation for the better in both the present and the future. It is therefore vital that the policy interventions that are researched are policy-relevant and as current as possible: EBP research needs to address the policy ‘questions of the day’ such as ‘how effective is the New Deal for Disabled People?’ or ‘how effective is the government’s health inequalities programme?’ or ‘do wage subsidies increase the employment rate of people with disabilities?’ Herein lies much of the attraction of EBP research, but it also poses a number of problems for the practical conduct of that research.

The demand by the policy process for answers to current questions means that the evidence base available to the EBP researcher is likely to be partial, limited and mainly ‘grey’, as the primary policy evaluations may not yet have been carried out, let alone published in the peer reviewed literature and indexed in electronic databases. For example, in the review of the effectiveness of recent labour market interventions aimed at people with a disability or a chronic illness (see Box 2), the contemporary nature of the review question meant that there were almost no relevant papers within the published and indexed literature, particularly for the employment programmes and work incentive interventions. The majority of the data was located in greyer sources such as government or policy institute reports, and accessed via web trawling, personal communication and citation follow-up. A similar pattern has also been experienced in the conduct of EBP research with a wider timeframe. For example, in a systematic review on the health effects of housing improvement only six of the 18 sources of evidence finally included in the synthesis were located by electronic means (Thomson et al, 2001). Similarly, in a systematic review of the health effects of new roads only a minority (eight) of the included studies were retrieved from databases, with the vast majority (24) identified by other means (Egan et al, 2003).

The prevalence of grey sources has an impact on the amount of time needed to conduct a review: searching websites, contacting experts in the field, and locating and accessing other sources of grey literature is much more time consuming than conventional electronic database literature searches. It has the added difficulty of

being a near endless task that the researcher must at some point decide to stop both for resource reasons and to ensure the policy relevancy of the end product – is the review of the evidence still contemporary, is it still a ‘debate of the day’ or has that debate now moved on?

Another problem facing the EBP researcher in identifying policy-relevant evidence is the language barrier. Locating grey literature that relates to UK policies is problematic at the levels already mentioned, but searching for non-UK grey literature is a near impossible task unless the researcher is fluent in relevant languages, has a firm grasp of likely evidence locations or has the services of an information technologist with both good general searching skills (internet as well as databases) and a good knowledge of the subject in hand. Non-UK grey literature, unlike the majority of non-English language published journal literature that is indexed in mainstream databases, is less likely to be written in the English language and EBP research budgets seldom stretch to affording translation services. Reviews are unable to include much of this sort of (relevant) material, introducing an English language bias and reducing the comprehensiveness of the synthesis. This bias is often extended by a lack of knowledge among English language reviewers of foreign language databases (some of which include grey literature, and some of which at least can be searched in English), and – equally (if not more) importantly – a lack of access to what are often subscription-based services.

For example, in the review of the effectiveness of recent labour market interventions aimed at people with a disability or a chronic illness (see Box 2), grey literature searching was conducted with relative ease for the two English language countries (UK and USA) but with less ease for Sweden (partial English coverage), while searching the German grey literature was an unfeasible task. Even when relevant non-English language grey literature documents were retrieved, mainly through contacts with country experts, the research budget could not afford the services of a translator. This was particularly the case when the documents in question were lengthy, with perhaps only a small part that was relevant to the research question, and which may well have been excluded anyway at the quality stage of the review process.

Evidence Design

In evidence based medicine, the systematic review method traditionally relies upon a rigid 'hierarchy of evidence' which is used to determine which pieces of research are included in the final synthesis, and which are excluded. Typically, only experimental designs are included (Cochrane Collaboration, 2001; NHS CRD 2001). The hierarchy is seen as particularly important in solving disputes between the findings of different research papers. For example, if the findings of paper A suggest that vocational rehabilitation is an effective way to help people with disabilities into work, while paper B suggests the opposite, the two research designs will be compared and the one that is higher up the hierarchy will be preferred, given more weight in the final synthesis, and be more apparent in the review's recommendations. However, the validity and utility of this traditional evidence hierarchy is increasingly being questioned within evidence-based medicine on the grounds that it over-simplifies the complex nature of research, underplays the value of non-experimental research, and ignores the contribution of qualitative research (Black, 1996; Sackett and Wennberg, 1997; Popay et al, 1998; Bigby, 2001; Dixon-Woods and Fitzpatrick, 2001; Malterud, 2001; Rychetnik, 2002).

These criticisms are particularly relevant to EBP research. Broader policy interventions are, by virtue of their operation within open systems in which evaluations are more problematic and expensive to conduct, less likely to be the subject of experimental research designs. They are much more likely to be covered by observational research designs (considered as 'lower' down the traditional evidence hierarchy) or those, most notably qualitative research, that do not figure in the hierarchy at all. The evidence hierarchy, with its tight methodological exclusion criteria, cannot therefore be transferred to EBP research without the danger that every review will conclude that there is no (experimental) evidence on this topic (Davies, 2000; Whitehead et al, 2000; Croucher et al, 2003; Egan et al, 2003; Petticrew, 2003; Petticrew and Roberts, 2003). Alternative approaches are thus being developed and used in which any idea of an evidence hierarchy is abandoned: while the focus on quality is maintained, the emphasis on which research design is better able to produce it is reduced (Croucher et al, 2003; Pawson et al, 2003; Petticrew and Roberts, 2003; Thomas et al, 2003). The issue of study quality is not whether one research design is 'better' than another but rather how does the study in question match up to the 'ideal

type' quality criteria for that particular design (Petticrew and Roberts, 2003), or how do studies with differing research designs match up to more generic standards such as transparency, accuracy or propriety (see Pawson et al, 2003). These more inclusive and realistic approaches to establishing study quality acknowledge the diversity of the policy-relevant evidence base and facilitate cross-study synthesis of evidence from different types of research design and evaluation (Thomas et al, 2003). They intelligently acknowledge that large parts of the evidence jigsaw (Whitehead et al, 2004) are, and can only ever be, covered by non-experimental or qualitative research evaluation, and thus enhance the utility of EBP research to policy makers by enabling the synthesis of the 'best available' evidence.

Conclusion

This paper has highlighted a few of the methodological challenges and practical problems that the conduct of a systematic review of a broad policy intervention(s) can produce. In theory, the systematic review process of 'locating, appraising, and synthesising evidence' (Petticrew, 2001) seems a simple and promising – if arduous – task, but numerous difficulties emerge while trying to conduct it (Wallace et al, 2004). EBP research is more problematic in practice because the interventions, such as labour market policies, are themselves hard to define, compare, isolate and evaluate; and because policy-relevant evidence, especially from non-English speaking countries, is difficult to locate and based on a wide variety of study design.

However, in pointing out the struggles that the review process can encounter, this paper is not saying that conducting EBP research is pointless or impossible. It simply points out that the practical side of reviewing is a lot less easy and a lot more complicated than the theory suggests, perhaps echoing Walker's comment that "social science cannot deliver to the policy maker all that is expected" (Walker, 2001: 325). EBP research was perhaps initially presented to both policy makers and researchers as a straightforward approach to locating, synthesising and highlighting research evidence. However, as this paper (and others such as Wallace et al, 2004) has tried to point out, it is far more than the rigid application of a methodological technique.

Policy makers (and perhaps even some researchers) need to be reminded to limit their expectations, especially of the evaluation of macro-level policies, in light of the

fragile and situated nature of social research. However, we as social researchers nonetheless need to continue our commitment to the evolution of EBP research methods, so that they are best able to deal with the limitations outlined and deliver as much as possible of the evidence base that policy makers desire. Already, much methodological development has occurred within the EBP movement. For example, theoretical work (Klein, 2000; Boaz et al, 2002; Packwood, 2002; Solesbury, 2002; Young et al, 2002); more practice-based developments such as rapid reviews or 'evidence nuggets' (Gibbs et al, 2003; Lucas, 2003; Lucas and Liabo, 2003; Brocklehurst and Liabo, 2004); alternative assessments of evidence design and quality (Croucher et al, 2003; Pawson et al, 2003; Petticrew and Roberts, 2003; Thomas et al, 2003); new means to appraise and synthesise qualitative research (McDermot and Graham, 2003), and new generic approaches to narrative synthesis (Pawson et al, 2003; Thomas et al, 2003); and even technical developments on issues such as database searching methods (Grayson and Gomersall, 2003). This work should be continued and expanded, so that EBP research emerges and thrives as a flexible, essential and prominent feature of the policy making process, regardless of who has power in Westminster.

Acknowledgements

Thanks to two anonymous referees and colleagues from the ESRC Evidence Network for their helpful comments on earlier drafts.

Box 1: Systematic review methodology

The systematic review is “a review of a clearly formulated question that uses systematic and explicit methods to identify, select and critically appraise relevant research, and to collect and analyse data from the studies that are included in the review” (Oakley and Fullerton, 1998). Put more simply, it is a method of locating, appraising, and synthesising evidence.

The review attempts systematically to locate all available research, both published and unpublished, and critically evaluates it on grounds of relevancy and a predetermined methodological criteria. Only research that is judged to be both relevant to the review question, and which fulfils the methodological inclusion criteria, is combined into the final review analysis. The systematic review combines the results of these studies and thus provides a summary of the ‘best available evidence’ on a given question (Chalmers, 1995; Greenhalgh, 1997; Cochrane Collaboration, 2001; NHS CRD, 2001).

Good quality systematic reviews follow a formal procedure which starts with the formulation of a precise question that “includes a definition of the participants, the intervention to be assessed and the outcomes to be measured” (Davies and Crombie, 1998; Hamer and Collinson, 1999). This is followed by an extensive literature search that is intended to find all research that relates to the review question. The search covers all relevant published and unpublished work (grey literature) which is identified through a combination of electronic database searches, searching the references of identified studies, hand searching relevant journals, and contacting experts in the field (Davies and Crombie, 1998; Hamer and Collinson, 1999; Cochrane Collaboration 2001; Glasziou et al, 2001; NHS CRD, 2001).

Box 2: A systematic review of the effectiveness of different labour market interventions, from Germany, Sweden, UK and USA, in helping disabled and chronically ill people gain employment

The high numbers of people unemployed and claiming benefits due to chronic illness or disability has become an increasingly salient political concern. This is especially so in the UK where the number and costs of the main incapacity and disability benefit claimants increased rapidly from 2.6 million claimants in 1987 at a cost of around £7 billion, to 5.3 million claimants in 1996 at a cost of about £21 billion (Brindle and McAskill, 1997). There are currently about 2.7 million people claiming incapacity and disability benefits in the UK, around 40% of whom will still be in receipt of these benefits six months later. Of this 40%, about 80% will still be claiming after five years (Blair, 2002). Incapacity and disability benefits are often claimed for a longer duration than unemployment benefits. For example, in 2002, 76% of claimants had been in receipt of their benefit for over two years compared to 9% for unemployment benefit claimants (DWP, 2002a). These benefits also currently account for the largest chunk of income maintenance programmes with 8.4% of the working age population claiming, compared to 2.6% for unemployment benefit and 2.3% for lone parent benefit (DWP, 2002b).

Tackling the number of claimants and level of expenditure on disability and invalidity benefits is not an easy task for policy makers. The UK government has therefore had to take a relatively incremental approach, gradually increasing the number of active labour market measures, for example the training, employment and work incentive measures incorporated in the New Deal for Disabled People package, and the employment rights incorporated in the Disability Discrimination Act. This review was initiated to assess the effectiveness of these types of labour market measures both in the UK and elsewhere.

- Question: How effective are different types of labour market interventions on helping disabled and chronically ill people gain employment?
- Method: Systematic review of all available types of study
- Population: Disabled/chronically ill
- Geographic coverage: Germany, Sweden, UK and USA.
- Intervention: Labour market measures legislation (for example, anti-discrimination law, or workplace quotas), employment programmes (for example, work experience, or work placements), work incentives (for example, wage subsidies, or tax credits), vocational rehabilitation (for example, job placement assistance, or training)
- Outcomes: Effectiveness – in terms of employment rate and duration, income, hours worked.

(Bambra, 2003; Bambra et al, in press)

References

- Acheson, D., Barker, D., Chambers, J., Graham, H. and Whitehead, M. (1998) *Independent Inquiry into Inequalities in Health*, London: Stationery Office.
- Bambra, C. (2003) *What types of 'welfare to work' and for whom? Protocol (abridged version)*, ESRC Centre for Evidence-Based Public Health Policy. Downloadable from: <http://www.msoc-mrc.gla.ac.uk/Evidence/Research/Review%2005/Review5-Protocol.pdf>
- Bambra, C. Whitehead, M. and Hamilton, V. 'Does "welfare to work" work? A systematic review of the effectiveness of welfare to work programmes for people with a chronic illness or disability', *Social Science and Medicine*, in press.
- Bigby, M. (2001) 'Challenges to the hierarchy of evidence: Does the emperor have no clothes?', *Archives of Dermatology*, vol 137, no 3, pp 345-6.
- Black, N. (1996) 'Why we need observational studies to evaluate the effectiveness of health care', *British Medical Journal*, vol 312, no 7040, pp 1215-18.
- Blair, T. (2002) 'Speech on welfare reform', *The Guardian*, 10 June.
- Boaz, A., Ashby, D. and Young, K. (2002) *Systematic reviews: What have they got to offer evidence based policy and practice?*, Working Paper no 2, London: ESRC UK Centre for Evidence Based Policy and Practice. Downloadable from: www.evidencenetwork.org/
- Brindle, D. and MacAskill, E. (1997) 'Analysis: disability cuts', *The Guardian*, 16 December.
- Brocklehurst, N. and Liabo, K. (2004) 'Evidence nuggets: promoting evidence based practice', *Community Practitioner*, vol 77, no 10, pp 371-5.
- Cabinet Office (1999) *Modernising government*, Cm 4310, London: Stationery Office. Downloadable from: <http://www.archive.official-documents.co.uk/document/cm43/4310/4310.htm>
- Cabinet Office (2000) *Adding it up: Improving analysis and modelling in central government*, London: Stationery Office.
- Campbell Collaboration (2001) *Guidelines for the preparation of review protocols: Version 1.0: January 2001* Downloadable from: http://www.campbellcollaboration.org/c2_protocol_guidelines%20doc.pdf
- Chalmers, I. and Altman, D.G. (eds) (1995) *Systematic reviews*, London: BMJ Publishing.

Cochrane Collaboration (2001) *Cochrane reviewers' handbook*, 4.1.4 edn, Oxford: Cochrane Collaboration. Downloadable from:
<http://www.cochrane.org/resources/handbook/index.htm>

Croucher, K., Quilgars, D., Wallace, A.; Baldwin, S. and Mather, L. (2003) *Paying the mortgage? A systematic literature review of safety nets for homeowners*, York: York University Department of Social Policy and Social Work. Summary downloadable from:
<http://www.york.ac.uk/chp/mppireview.pdf>

Davies, H. and Crombie, I. (1998) *What is a systematic review?* West Malling: Hayward Medical Publications.

Davies, H. and Smith, P. (1999) 'Editorial: What works? The role of evidence in public sector policy and practice', *Public Money and Management*, vol 19, no 1, pp 3-5.

Davies, P. (2000) 'Contributions from qualitative research', in H. Davies, S. Nutley, and P. Smith (eds) *What works? Evidence based policy and practice in public services*, Bristol: The Policy Press, Bristol, pp 291-316.

Dixon-Woods, M. and Fitzpatrick, R. (2001) 'Qualitative research in systematic reviews', *British Medical Journal*, vol 323, no 7316, pp 765-6.

DWP (2002a) *Incapacity benefit and severe disablement allowance: Quarterly summary statistics*, London: Department for Work and Pensions, Analytical Services Directorate. Downloadable from:
http://www.dwp.gov.uk/asd/ib_sda.asp

DWP (Department for Work and Pensions) (2002b) *Client group analysis: Quarterly bulletin on the population of working age on key benefits*, London: Department for Work and Pensions, Analytical Services Directorate. Downloadable from:
<http://www.dwp.gov.uk/asd/cga.asp>

Egan, M., Petticrew, M., Ogilvie, D. and Hamilton, V. (2003) 'New roads and human health: a systematic review', *American Journal of Public Health*, vol 93, no 9, pp 1463-71.

Gibbs, J., Underdown, A. and Liabo, K. (2003) *Group-based parenting programmes can reduce behaviour problems of children aged 3-10 years*, Evidence Nugget, London: What Works for Children, City University. Downloadable from:
http://www.whatworksforchildren.org.uk/online_reading/parenting/parenting_read.htm

Glasziou, P., Irwig, L., Bain, C. and Colditz, G. (2001) *Systematic reviews in healthcare: a practical guide*, Cambridge: Cambridge University Press.

Grayson, L. and Gomersall, A. (2003) *A difficult business: Finding the evidence for social science reviews*, Working Paper no 19, London: ESRC UK Centre for Evidence Based Policy and Practice. Downloadable from:
www.evidencenetwork.org/

- Greenalgh, T. (1997) 'How to read a paper: Papers that summarise other papers (systematic reviews and meta-analyses)', *British Medical Journal*, vol 315, no 7109, pp 672-5.
- Hamer, S. and Collinson, G. (eds) (1999) *Achieving evidence based practice: A handbook for practitioners* Edinburgh: Baillière Tindall in association with the Royal College of Nursing.
- Judge, K. (2001) 'Evaluating welfare to work in the UK, in N. Gilbert and R. van Voorhis, (eds) *Activating the unemployed: A comparative appraisal of work-oriented policies*, London: Transaction.
- Judge, K. and Bauld, L. (2001) 'Strong theory, flexible methods: Evaluating complex community-based initiatives', *Critical Public Health*, vol 11, no 1, pp 19-38.
- Judge, K. and Mackenzie, M. (2002) 'Theory-based evaluation: New approaches to evaluating complex community-based initiatives', in J. P. Mackenbach and M. Bakker (eds) *Reducing inequalities in health: A European perspective*, London: Routledge, pp 300-13.
- Klein, R. (2000) 'From evidence based medicine to evidence based policy?', *Journal of Health Services Research and Policy*, vol 5, no 2, pp 65-6.
- Lehman, A., Goldberg, R., Dixon, L., McNary, S., Postrado, L., Hackman, A., and McDonnell, K. (2002) 'Improving employment outcomes for people with severe mental illness', *Archives of General Psychiatry*, vol 59, no 2, pp 165-72.
- Lucas, P. (2003) *Breakfast clubs and school fruit schemes: Promising practice*, Evidence Nugget, London: What Works for Children, City University. Downloadable from:
<http://www.whatworksforchildren.org.uk/docs/Nuggets/pdfs/Breakfastclubs090104.pdf>
- Lucas, P. and Liabo, K. (2003) *One-to-one, non-directive mentoring programmes have not been shown to improve behaviour in young people involved in offending or anti-social activities*, Evidence Nugget, London: What Works for Children, City University. Downloadable from:
http://www.whatworksforchildren.org.uk/online_reading/mentoring/mentoring_read.htm
- Macintyre, S., Chalmers, I., Horton, R., and Smith, R. (2001) 'Using evidence to inform health policy: A case study', *British Medical Journal*, vol 322, no 7280, pp 222-5.
- Macintyre, S. and Petticrew, M. (2000) 'Good intentions and received wisdom are not enough', *Journal of Epidemiology and Community Health*, vol 54, no 11, pp 802-3.
- Malterud, K. (2001) 'The art and science of clinical knowledge: Evidence beyond measures and numbers', *The Lancet*, vol 358, no 9279, pp 397-400.
- Marsh, D. and Mackie, T. 1995, 'The comparative method', in D. Marsh and G. Stoker (eds) *Theories and methods in political science*, London: Macmillan.

- McDermott, E. and Graham, H. (2004) *Experiences of being a teenage mother in the UK: A report of a systematic review of qualitative studies*. Downloadable from: http://www.msoc-mrc.gla.ac.uk/evidence/Research/Research_MAIN.html
- Mulrow, C. D. (1994) 'Systematic reviews: rationale for systematic reviews', *British Medical Journal*, vol 309, no 6954, pp 597-9.
- NHS CRD (2001) *Undertaking systematic reviews of research on effectiveness: CRD's guidance for those carrying out or commissioning reviews*, CRD Report 4: 2nd edn, York: NHS Centre for Reviews and Dissemination. Downloadable from: <http://www.york.ac.uk/inst/crd/report4.htm>
- Oakley, A. (1998) 'Public policy experimentation: Lessons from America', *Policy Studies*, vol 19, no 2, pp 93-114.
- Oakley, A. and Fullerton, D. (1998) 'The lamp-post of research: support or illumination?', in A. Oakley and H. Roberts (eds) *Evaluating social interventions: a report of two workshops funded by the Economic and Social Research Council*, Ilford: Barnardos, pp 4-38.
- Oakley, A. (2000) *Experiments in knowing: Gender and method in the social sciences*, Cambridge: Polity Press.
- Packwood, A. (2002) 'Evidence-based policy: Rhetoric and reality', *Social Policy and Society*, vol 1, no 3, pp 267-74.
- Pawson, R. (2001a) 'Evidence based policy: In search of a method', *Evaluation*, vol 8, no 2, pp 157-81.
- Pawson, R. (2001b) 'Evidence based policy: The promise of realist synthesis', *Evaluation*, vol 8, no 3, pp 340-58.
- Pawson, R. (2002) 'Evidence and policy and naming and shaming', *Policy Studies*, vol 23, no 3/4, pp 211-30.
- Pawson, R., Boaz, A., Grayson, L., Long, A., and Barnes, C. (2003) *Types and quality of knowledge in social care*, Knowledge Review no 3, London: Social Care Institute for Excellence. Downloadable from: <http://www.scie.org.uk/publications/list.asp>
- Petticrew, M. (2001) 'Systematic reviews from astronomy to zoology: Myths and misconceptions', *British Medical Journal*, vol 322, no 7278, pp 98-101.
- Petticrew, M. (2003) 'Why certain systematic reviews reach uncertain conclusions', *British Medical Journal*, vol 326, no 7392, pp 756-8.
- Petticrew, M. and Roberts, H. (2003) 'Evidence, hierarchies, and typologies: Horses for courses', *Journal of Epidemiology and Community Health*, vol 57, no 7, pp 527-9.
- Popay, J., Rogers, A., and Williams, G. (1998) 'Rationale and standards for the systematic review of qualitative literature in health services research', *Qualitative Health Research*, vol 8, no 3, pp 341-51.

Rychetnik, L., Frommer, M., Hawe, P. et al. (2002) 'Criteria for evaluating evidence on public health interventions', *Journal of Epidemiology and Community Health*, vol 56, no 2, pp 119-27.

Sackett, D.L. and Wennberg, J.E. (1997) 'Choosing the best research design for each question', *British Medical Journal*, vol 315, no 7123, p 1636.

Sanderson, I. (2002) 'Evaluation, policy learning and evidence-based policy making', *Public Administration*, vol 80, no 1, pp 1-22.

Solesbury, W. (2002) 'The ascendancy of evidence', *Planning Theory and Practice*, vol 3, no 1, pp 90-6.

Thomas, J., Sutcliffe, K., Harden, A., Oakley, A., Oliver, S., Rees, S., Brunton, G. and Kavanagh, J. (2003) *Children and healthy eating: A systematic review of barriers and facilitators*, London: EPPI Centre, Social Science Research Centre, Institute of Education, University of London.

Thomson, H., Petticrew, M. and Morrison, D. (2001) 'Health effects of housing improvement: systematic review of intervention studies', *British Medical Journal*, vol 323, no 7306, pp 187-90.

Walker, R. (2000) 'Learning if policy will work: The case of the New Deal for Disabled People', *Policy Studies*, vol 21, no 4, pp 313-32.

Walker, R. (2001) 'Great expectations: Can social science evaluate New Labour's policies?', *Evaluation*, vol 7, no 3, pp 305-30.

Wallace, A., Croucher, K., Quilgars, D. and Baldwin, S. (2004) 'Meeting the challenge: Developing systematic reviewing in social policy', *Policy and Politics*, vol 32, no 4, pp 455-70.

Whitehead, M., Diderichsen, F., and Burstrom, B. (2000) 'Researching the impact of public policy on inequalities in health', in H. Graham (ed) *Understanding health inequalities*, Buckingham: Open University Press, pp 203-18.

Whitehead, M., Petticrew, M., Macintyre, S., Graham, H., and Egan, M. (2004) 'Evidence for public health policy on inequalities I: The reality according to policy makers', *Journal of Epidemiology and Community Health*, vol 58, no 10, pp 811-16.

Whitehead, M., Petticrew, M., Macintyre, S., Graham, H., Bambra, C. and Egan, M. (2004) 'Evidence for public health policy on inequalities II: Assembling the evidence jigsaw', *Journal of Epidemiology and Community Health*, vol 58, no 10, pp 817-21.

Young, K., Ashby, D., Boaz, A., and Grayson, L. (2002) 'Social science and the evidence-based policy movement', *Social Policy and Society*, vol 1, no 3, pp 215-24.