Journal of Development Effectiveness
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/rjde20

Narrative approaches to systematic review and synthesis of evidence for international development policy and practice

Birte Snilstveit [a b], Sandy Oliver [c] & Martina Vojtkova [a]

[a] International Initiative for Impact Evaluation (3ie), London, UK

[b] London International Development Centre, 36 Gordon Square, WC1H OPD, London, UK

[c] EPPI - Centre, Institute of Education, London, UK

Version of record first published: 18 Sep 2012.

To cite this article: Birte Snilstveit, Sandy Oliver & Martina Vojtkova (2012): Narrative approaches to systematic review and synthesis of evidence for international development policy and practice, Journal of Development Effectiveness, 4:3, 409-429

To link to this article: http://dx.doi.org/10.1080/19439342.2012.710641

Journal of Development Effectiveness
Vol. 4, No. 3, September 2012, 409–429

Routledge
Taylor & Francis Group

Narrative approaches to systematic review and synthesis of evidence for international development policy and practice

Birte Snilstveit[a,b]*, Sandy Oliver[c] and Martina Vojtkova[a]

[a]*International Initiative for Impact Evaluation (3ie), London, UK;* [b]*London International Development Centre, 36 Gordon Square, WC1H OPD London, UK;* [c]*EPPI – Centre, Institute of Education, London, UK*

Thus far, most systematic reviews commissioned to inform international development policy have focused on questions of 'what works', drawing on experimental and quasi–experimental studies of the effects of interventions. This article argues that systematic review methods can be applied to answer a range of different questions for international development and pays particular attention to methods of synthesising qualitative evidence that apply the key principles of systematic reviewing of being comprehensive, systematic and transparent. The article introduces readers to the types of questions for which reviews of qualitative evidence might be appropriate, the types of evidence such reviews might include and the range of methods available for their synthesis.

Keywords: systematic review; qualitative synthesis; evidence-based policy; narrative synthesis; development effectiveness

1. Introduction

Over the last decade, there has been an increased focus on enhancing the use of evidence to inform international development policy and practice. Donors, implementing agencies and governments are under increasing pressures to demonstrate that their policies are informed by evidence. This has led to an increase in the demand for systematic reviews of the evidence pertaining to a range of different areas of policy with relevance to low- and middle-income countries.

Thus far, most systematic reviews commissioned to inform international development policy have focused on questions of 'what works', drawing on experimental and quasi-experimental studies of the effects of interventions. However, such reviews are increasingly criticised for being too narrow and 'rigid', failing to address other important questions such as, for instance, why a particular interventions work or not (Mallett *et al.* 2012). Moreover, a number of reviews of effectiveness have been inconclusive due to a lack of evidence on effects and there are rising concerns that useful evidence on other questions is being overlooked (Snilstveit 2012).

A narrow focus on effectiveness when policy-makers need answers to a range of questions beyond 'what works' is rightly criticised, although this is not a limitation of systematic review methodology per se. The principles underpinning systematic review methodology can be applied to answer a range of different questions, and include the synthesis of a range of different types of evidence, including qualitative evidence (Petticrew and Roberts 2006,

*Corresponding author. Email: bsnilstveit@3ieimpact.org

ISSN 1943-9342 print/ISSN 1943-9407 online

http://dx.doi.org/10.1080/19439342.2012.710641
http://www.tandfonline.com

Gough *et al.* 2012). To maintain the momentum around systematic reviews in international development and to utilise existing research and support the development of evidence-informed polices, there is a need to synthesise evidence across the broad range of questions asked by policy-makers and practitioners in the field.

This article argues that systematic review methods can be applied to answer a range of different questions for international development and pays particular attention to methods for synthesising qualitative evidence that apply the key principles of systematic reviewing of being comprehensive, systematic and transparent. Applying systematic review methodology to a broader range of evidence has implications for every step of the review process, from defining the review question to developing inclusion criteria, searching for studies, critically appraising studies, synthesising and presenting findings. It is beyond the scope of this article to offer guidance on all of these aspects (see Snilstveit (2012) for a review of some of these issues).

Rather, this article introduces readers to the types of questions for which reviews of qualitative evidence might be appropriate, the types of evidence such reviews might include and the range of methods available for their synthesis. The remainder of this article is structured as follows. The next section outlines a range of different questions of relevance to international development policy and practice, and the evidence that systematic reviews might draw on to answer these different types of questions. It then goes on to discuss the range of existing methods of synthesis, providing some relevant examples of their applications to date. The fourth section discusses the application of qualitative synthesis in international development and reflects on the methods that are likely to be most appropriate to international development reviews, and for which questions. The final section draws the conclusion that systematic reviews are an important tool for making better use of existing evidence to support policy development.

2. Different types of reviews for addressing different types of questions

What is the rate of maternal mortality in low- and middle-income countries? What are the policy options for reducing the gender gap in education? What are the potential factors affecting the successful implementation of interventions to improve food security? What are the alternatives to assist people vulnerable to climate change adapt to a changing climate, and what are the effects of these options? These questions are examples of the range and type of questions that are of relevance to evidence-informed policy in international development. It highlights that policy-makers and practitioners need answers to a range of different questions to inform the different stages of the policy-making process (Lavis 2009, Lewin *et al*. 2012).

These broader policy questions can be divided into more clearly defined research questions that might be appropriate and manageable for a systematic review (Petticrew and Roberts 2006). Such research questions can pertain to the scale of a problem, the factors underlying a problem, barriers and facilitators to implementing a policy option, acceptability of proposed solutions to users, the positive and negative effects of an intervention and the cost-effectiveness of different interventions (Petticrew and Roberts 2003, Lavis 2009, NICE 2009). The demand for answers to such a range of questions has led to the development of different types of systematic reviews (Petticrew and Roberts 2006, Pope *et al.* 2007, NICE 2009, Snilstveit 2012).

Snilstveit (2012) argues for the use of theory as a way of enhancing the policy relevance of systematic reviews of effectiveness. Similarly, drawing on a theoretical framework for a particular issue or interventions can be useful in identifying relevant and researchable

questions from a broader policy question. If the broader policy question relates to defining and framing an issue, a broad conceptual framework can be useful. Although if the broader policy question is concerned with identifying implementation considerations, a programme theory for how the intervention is supposed to work, and for whom, can be useful for identifying relevant research questions (Petticrew and Roberts 2006, Tugwell *et al*. 2010).

Table 1 provides an example applied to education of how the broader questions that might be relevant for the different stages of the policy-making process can be divided into more answerable research questions, and indicates the sources of evidence appropriate for addressing these questions.

Different types of evidence are appropriate for answering different types of questions, and once the review question(s) is clarified, the next step is to determine the inclusion criteria for the focus of study and study design (Petticrew and Roberts 2006). To answer questions about the effectiveness of interventions, studies should be able to demonstrate causality through attribution analysis for which experimental and quasi-experimental studies are generally considered most appropriate (White 2009). However, 'hierachies of evidence' for other questions would be different (Petticrew and Roberts 2006, Saini 2012). To answer questions such as 'why' an intervention works (or not), or 'how' something works, qualitative research and surveys would be more appropriate than experimental and quasi-experimental studies (Petticrew and Roberts 2006). Although quantitative research is generally concerned with counting and measuring, qualitative research is commonly applied to addressing questions related to the meanings, concepts, definitions, characteristics and descriptions of issues (Berg 2009). Qualitative research encompasses a range of different research methods employed to address such questions, including ethnography/participant observation, qualitative interviewing, focus groups, discourse analysis, qualitative analysis of text and case studies (Bryman 2001).

There is a debate in the literature as to what constitutes the quality in qualitative research (Spencer *et al*. 2003), and this debate extends to the methodological literature on qualitative evidence synthesis (Hannes 2011; see also Snilstveit (2012)). Just as Hannes (2011) suggests 'qualitative research as a scientific process needs to be "rigorous" and "trustworthy" to be considered as a valuable component of Cochrane systematic review', we argue this also extends to other types of evidence synthesis that aim to inform policy. Following Spencer *et al.* (2003), we suggest that qualitative research should be (1) contributory, (2) defensible in design (3) rigorous in conduct and (4) credible in claim (p. 20). Although there is a lot of high-quality research in the field of international development, there is also a relatively large volume of poor-quality research that fails to adopt a rigorous approach. The boundaries between research and advocacy are often blurred, and such material needs to be treated with caution.

Complementary application of quantitative, qualitative and mixed methods research is well recognised for primary studies (Cresswell 2003) and this approach is now appearing in systematic reviews in international development. For instance, rich, qualitative studies contributed to more understanding in a review of adherence to tuberculosis treatment (Munro *et al*. 2007). Thinner qualitative findings resulting from recognised (not necessarily naturalistic) methods for data collection and analysis were also useful in reviews comparing the quality of health services in low- and middle-income countries (Berendes *et al*. 2011, Basu *et al*. 2012). However, too little qualitative research was found directly linked to trials to explain differences in the impact of lay health workers (Glenton *et al.* 2011).

Syntheses of qualitative evidence can answer a different (and often complementary) set of questions from quantitative syntheses of effectiveness (Saini 2012). Such an approach starts with the research question and adopts the method(s) that is best suited for answering

Table 1. Examples of different types of reviews needed for policy making.

Steps in the policy-making process	Policy question	Examples of research question	Sources of appropriate evidence to address the question
Defining and framing the problem.	What is the need for intervention, in terms of the nature, magnitude and framing of the problem? Example: What is the education deficit in low- and middle-income countries and what are the factors associated with it?	What are the school enrolment rates in low- and middle-income countries? What is the relationship between school enrolment rates and distance to school, type of school, quality of teaching, class size and school meal provision? What are the barriers and facilitators to increasing enrolment rates/improving education outcomes? What factors are associated with increasing enrolment rates/reducing drop-out rates/improving education outcomes? How do the above differ for the most disadvantaged sub-groups? How do parents and children perceive/experience public and private school education?	Reviews of observational studies addressing the nature and magnitude of the problem (for example, surveys, studies based on routine data to assess prevalence or burden of an issue, including data disaggregated along dimensions of disadvantage). Reviews of qualitative studies of views and experiences regarding a problem.
Assessing potential policy options.	What is the appropriate set of policy options to address the problem and what are the effects of these options? Example: What are the most effective approaches to improve education outcomes in low- and middle-income countries and how effective/cost-effective are these?	Do interventions to increase enrolment rates improve education outcomes? Does increasing access to private schools work better than teacher training?	Reviews of studies of the effectiveness of interventions – experimental, non-experimental. Reviews of economic evaluations of interventions.

		Does providing vouchers for private schools do more harm than good? Is it worth investing in school vouchers? Are school vouchers acceptable to beneficiaries and other stakeholders? Will beneficiaries be willing to or want to send their children to private schools?	Reviews of qualitative studies of views and experiences regarding interventions.
Identifying implementation considerations for selected policy options.	What are the potential factors affecting the successful implementation of the policy options? Example: What are the potential factors affecting the successful implementation of education interventions?	How do vouchers for private schools work? Are children, parents and teachers satisfied with the education services provided? What are the barriers and facilitators to effective implementation/up-take?	Reviews of effectiveness studies of implementation strategies and of equity across the most disadvantaged – least disadvantaged groups. Reviews of qualitative studies of the acceptability of interventions. Reviews of process evaluations of implementation strategies, which may use mixed methods.

Source: Adapted from Lewin *et al*. (2012), NICE (2009) and Petticrew and Roberts (2003).

it (ibid). If a systematic review aims to answer several questions, researchers might need to draw on a range of different types of evidence.

3. Narrative approaches to synthesis of qualitative evidence

Narrative methods of synthesis can be used to synthesise both quantitative and qualitative studies and have been used when the experimental and quasi-experimental studies included in a systematic review are not sufficiently similar for a meta-analysis to be appropriate (Mays *et al.* 2005a). Narrative synthesis is used in different ways. In this article, we use it as an overarching term to describe a family of methods for synthesising data narratively, focusing particularly on the application of narrative approaches to the synthesis of qualitative evidence.

Synthesising empirical qualitative evidence is analogous to synthesising quantitative evidence in that 'there is a shared interest in synthesising empirical studies' (Noblit and Hare 1988, p. 10). However, unlike quantitative synthesis that converts information into a common metric and synthesises these data to test a theory using statistical meta-analysis, qualitative synthesis aims to synthesise qualitative data, which is commonly text-based. Such reviews adopt a narrative, as opposed to statistical, approach to research synthesis and seek to generate new insights and recommendations by going beyond the summary of findings from different studies as in traditional narrative reviews.

Some of the weaknesses of narrative synthesis noted in the literature are the lack of transparency (Dixon-Woods *et al*. 2005) and the lack of clarity on methods and formal guidance on how to conduct such a synthesis (Mays *et al.* 2005b). Popay *et al.* (2006) provide a guidance on how narrative synthesis can be conducted in a more systematic and transparent way, focusing on synthesis of evidence on effectiveness of interventions and factors determining the implementation of interventions. The guidance document provides specific suggestions as to which tools and techniques are appropriate for narrative synthesis (see Table 2 for an overview), and two worked examples which are helpful in illustrating how these can be used in practice. These techniques can be usefully applied for enhancing the transparency of all types of narrative synthesis, including the specific methods of synthesis reviewed below. Other tools such as evidence tables, specialised software (such as Nvivo and EPPI reviewer) and matrices can also be used to enhance the transparency of most methods of narrative syntheses.

As narrative reviews have become increasingly systematic, their methods have diversified and the terms to describe them have proliferated. The methodological literature describes a range of methods available for integrating qualitative evidence, as well as methods for synthesising both quantitative and qualitative evidence. A number of articles and books reviewing these methods have been published in recent years (see, for example, Dixon-Woods *et al*. 2005, Mays *et al.* 2005a, Pope *et al.* 2007, Barnett-Page and Thomas 2009, Noyes *et al.* 2011, Saini and Shlonsky 2012). Methods include content analysis, critical interpretive synthesis, thematic synthesis, realist synthesis, grounded theory, case survey, qualitative comparative synthesis, meta-summary and framework synthesis.

Reviewing this literature and examples of applications of these methods suggests there are more labels describing the methods of synthesis than there are genuine differences between their different methods (Barnett-Page and Thomas 2009). Most of these methods involve some form of structured synthesis of studies, with a distinguishing feature being the extent to which the various methods aim to test, explore or generate theories and the extent to which they interpret the evidence from the included studies in the synthesis (Thomas *et al.* 2012). Methodologists variously conceptualise methods as being

Table 2. Tools and techniques for narrative synthesis.

Element of synthesis	Suggested tools and techniques
Developing a theory of how the intervention works, why and for whom?	No specific tools or techniques identified. However, it is noted that tools and techniques suggested for other elements of the synthesis can contribute to developing the theory of change.
Developing a preliminary synthesis of findings of included studies.	Textual description of studies, groupings and clusters, tabulation, transforming data into a common rubric, vote counting as a descriptive tool, thematic and content analysis for translating data.
Exploring relationships in the data.	Graphs, frequency distributions, funnel plots, forest plots and L'Abbe plots; moderator variables and sub-group analyses; idea webbing and conceptual mapping; reciprocal and refutational translation; qualitative case descriptions; investigator/moderator triangulation; conceptual triangulation.
Assessing the robustness of the synthesis.	Weight of evidence (for example, Harden and Gough 2012); best evidence synthesis; validity assessment (for example, the CDC approach); reflecting critically on the synthesis process; checking the synthesis with authors of primary studies.

Source: Popay *et al*. (2006).

on a continuum from aggregative approaches at one end to interpretive (or configurative) synthesis methods at the other (Noyes and Lewin 2011, Saini and Shlonsky 2012, Thomas *et al.* 2012). Figure 1 illustrates this schematically, placing the different methods on this continuum.

The choice of method of synthesis depends on a number of factors, including the question and purpose of the synthesis, the nature of the evidence as well as time and resources (Noyes *et al*. 2011). For questions aiming to test a hypothesis or theory, aggregative approaches are most appropriate. Answering questions that aim to explore or conceptualise an issue might be best addressed through a configurative/interpretive synthesis (Thomas *et al.* 2012). Moreover, the nature of the evidence to be synthesised will also influence what is the appropriate synthesis method (Voils *et al*. 2008, Noyes and Lewin 2011). Aggregative methods might be more appropriate for synthesising descriptive, or 'thin' qualitative evidence, whereas configurative/interpretive methods may require studies containing richer data (Noyes and Lewin 2011). Finally, the choice of synthesis method also depends on the time, resources and skills available to the review team.

Below we outline some of the approaches that have been developed for syntheses that include qualitative evidence, including some examples. The strengths and weaknesses of these methods are summarised in Table 3.

3.1. Content analysis

There are several quantitative approaches to synthesis that involve transforming all data into quantitative measures such as frequencies (Mays *et al.* 2005a). Content analysis, first developed as a method for primary research, is one such approach. It involves developing

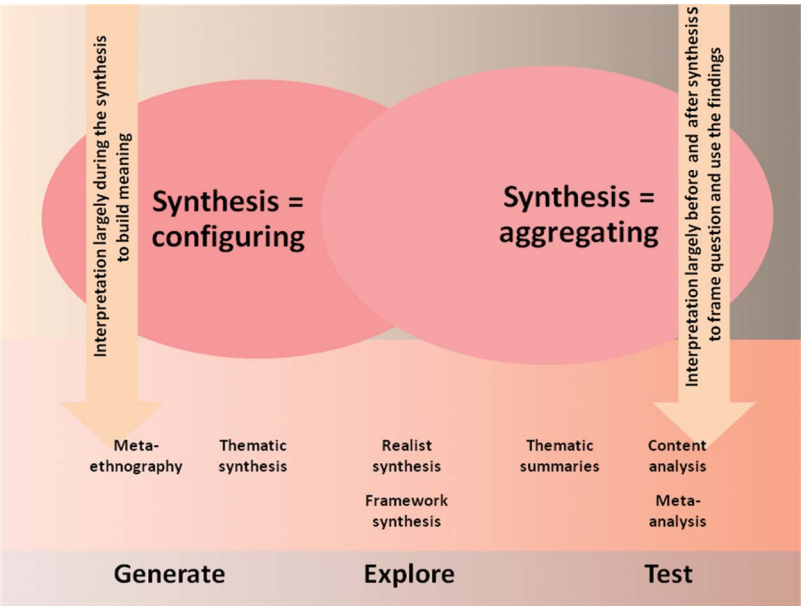

Figure 1. Methodological continuum of synthesis approaches and methods.
Source: Adapted from Thomas *et al*. (2012).

themes or categories a priori, for instance by drawing on a programme theory, coding data according to these categories and creating tabulations of frequency counts to identify key findings (Mays *et al.* 2005b). Content analysis can be used to synthesise data from both quantitative and qualitative studies (Pope *et al.* 2007). The categories for coding may be determined by drawing on a theoretical framework, for instance the programme theory underlying a particular intervention. To enable several people to code papers in a consistent way, it requires categories to be adequately defined and mutually exclusive (Dixon-Woods *et al*. 2005). Content analysis is a well-developed method and there are a range of software packages to assist with conducting the analysis (Dixon-Woods *et al*. 2005).

Although content analysis is a well-established method in the social sciences, Dixon-Woods *et al.* (2005) suggest there are limitations to its usefulness for synthesis, including that 'it is inherently reductive and tends to diminish complexity and context' (p. 50), as well as the risk of treating an issue as unimportant due to lack of available data. Nevertheless, content analysis does not have to be an exclusively quantitative and reductionist method (Berg 2009), and if applied thoughtfully, keeping these issues in mind, it can be a useful method for synthesising a large number of studies. As an aggregative synthesis method, it might be most appropriately applied to descriptive, 'thin' data, for instance from project documents or surveys.

An example of a systematic review using content analysis is a review examining the reasons for restraining patients in care (Evans and Fitzgerald 2002). The authors coded categories and themes reported as reasons for restraining patients from the findings of the 23 included studies, including ensuring safety, understaffing and maintaining the social environment. They used four categories and four sub-categories and report the percentage of reports citing these various categories to determine common findings.

Table 3. Strengths and weaknesses of methods for narrative synthesis.

Method	Strengths	Weaknesses	Guidelines and examples
Content analysis: Data categorised and frequency of each theme determined to identify key findings. Categories/themes identified a priori.	Widely used, transparent and replicable. Method well developed. Software to assist with the analysis is available.	Described as reductive approach – concerns that importance of context and complexity of primary studies reduced. 'Danger that absence of evidence (non-reporting) could be treated as evidence of absence (not important)' (Dixon-Woods *et al*. 2005, p. 50).	Evans and Fitzgerald (2002)
Thematic summaries: Summarises findings of included studies within salient themes.	Offers detailed description of included studies. Can accommodate variations in context and 'thin' findings.	Can lead to 'vote counting' (see Waddington *et al*. (2012) for a discussion of the potential pitfalls of vote counting).	Thomas *et al*. (2012)
Framework synthesis: Identify key concepts in advance and supplement with themes emerging from included studies. Configure and aggregate findings from quantitative/qualitative studies. Themes are sometimes counted and displayed in tables.	Structured by a transparent and evolving framework that is open to policy, practitioner and service user perspectives. Can be applied to research and other literatures, and conducted by novice reviewers. Framework can accommodate variations in context and 'thin' findings. Quicker than only identifying themes grounded in the data.	Application of tentative theory early in the procedure may lead to overlooking some themes grounded in the data.	Kiwanuka *et al*. (2011), Koehlmoos *et al*. (2011)
Thematic synthesis: Coding of text to identify main themes emerging from the literature. Synthesis of findings organised around these key themes and concepts. Can incorporate both quantitative and qualitative findings. Themes are sometimes counted and displayed in tables.	Offers a structured way to organise the literature and identify key themes. Allows the synthesis of both quantitative and qualitative data.	Can be either data-driven or theory-driven. Diversity of approaches gives rise to uncertainty about how synthesis developed.	Harden *et al*. (2006), Thomas *et al*. (2007), Thomas and Harden (2008)

(*Continued*)

Table 3. (Continued).

Method	Strengths	Weaknesses	Guidelines and examples
Realist synthesis: Theory-driven approach, focused on the underlying programme theory and mechanisms driving an intervention. Allows the inclusion of a wide range of evidence, from RCTs to newspaper reports.	Focus on informing policy-makers of how, why and where interventions work. Focus on context–mechanisms–outcomes enables causal chain analysis. Allows the synthesis of a range of different types of evidence.	Concerns over feasibility and the inclusion of weaker study designs in the synthesis. Lack of methodological guidance (although see RAMESES). Resource intensive.	Greenhalgh *et al*. (2007), van der Knapp *et al*. (2008)
Meta-ethnography: Key themes, concepts and metaphors identified in the primary studies. Three types of synthesis: (1) translate themes across studies (reciprocal translation); (2) identify, characterise and explain contradictory findings (refutational synthesis), and (3) use findings to develop a general interpretation (lines-of-argument synthesis).	Goes further in developing new theory than other approaches. Useful for policy-makers in that it can assist in explaining different research findings.	Has not yet been used to synthesise findings from both quantitative and qualitative studies. Resource-intensive approach. Issues related to transparency and offers no guidance on quality appraisal. Requires considerable experience of qualitative analysis.	Britten *et al*. (2002), Munro *et al*. (2007), Williamson *et al*. (2009)

Source: Dixon-Woods *et al*. (2005) and Mays *et al*. (2005a, 2005b).

3.2. Thematic summaries

As outlined by Thomas *et al.* (2012), thematic summaries draws on the conceptual framework of the review to categorise studies into thematic groups that are relevant for the intended reader. Examples of categories that can be used to create the different thematic groups include the type of intervention, participants, study design and/or study quality, outcomes and programme theory used in the included study. A detailed assessment of the characteristics of the included studies allows the allocation of each study into the correct thematic group. Findings of the studies in each thematic group are then analysed and synthesised separately using an aggregative or interpretive narrative synthesis method. Common ways of synthesis include a tabulation of the findings into a thematic framework based on the a priori identified categories, analysis of the differences between studies in each group, identification of any divergent findings and a synthesis of the findings under each theme. The method shares the weaknesses of narrative synthesis, in particular if vote counting is used as a method to synthesise findings. Thematic summaries use an entirely deductive approach to generate the themes they contain, and therefore heavily rely on the quality of the initial conceptual framework of the review to provide an organising structure.

3.3. Framework synthesis

Framework synthesis has evolved from framework analysis employed for policy relevant primary research where urgency requires rapid results that can be delivered by large teams (Ritchie and Spencer 1994, Pope *et al.* 2000). It begins with a tentative framework that is either borrowed from elsewhere (Carroll *et al.* 2011) or constructed from key concepts and dimensions of difference seen in interventions and their contexts (Oliver *et al*. 2008). Reading abstracts and subsequently full texts reveals new themes grounded in the included studies which may then be used to amend the framework. Studies are coded according to the developing framework in an iterative process until the body of evidence can be presented coherently in tables or images matching the framework. Both the framework and the evidence extracted from the included studies are open to visual inspection and debate by a large review team, a wider group convened to guide the review, and readers of the final report. Patterns in the data can be recognised by reviewers and readers comparing the findings of studies positioned differently within the framework, allowing conclusions to be drawn about relationships between study findings and variations in populations, interventions and their theoretical basis and context.

A framework's key concepts and dimensions may be introduced by policy priorities or constraints, practitioner knowledge, service user perspectives, public interests or the review team's familiarity with the issues. The range of perspectives employed to build the framework maximises the relevance of subsequent findings. A framework that accommodates a large number of different types of studies is currently being employed for two parallel reviews of the impact of protected terrestrial areas (such as national parks and national forests) on the well-being of local residents. Having a single framework will allow the outcomes of the synthesised quantitative evidence to be compared with the views of people talking about different categories of protected areas, from different socio-economic positions and in terms of livelihood strategies, social capital, empowerment, human rights and access to ecosystem resources (Pullin *et al.* 2012).

Framework synthesis can draw on 'thin' research findings and other types of literature such as policy documents or think pieces either to inform the framework or present collective learning. This was valuable for a review of policies addressing health professionals holding two paid roles (dual practice); in the absence of impact evaluations, much was

learnt from policy analyses, country case studies, cross-sectional surveys and economic models (Kiwanuka *et al.* 2011).

These applications show how framework synthesis can build and consolidate knowledge by describing large numbers of studies across broad sweeps of literature even where the research is underdeveloped. It has additional benefits of accommodating complexity in programmes, variations in context and allowing the focus and interpretation to be guided by different perspectives.

Box 1. Social franchising evaluations: a scoping review

Koehlmoos and colleagues (2011) employed framework synthesis to review the scope and nature of existing evidence on social franchising interventions in health service delivery, including a focus on reach, implementation, equity and sustainability. The review was conducted to complement a review of effectiveness of social franchising interventions which did not find any rigorous impact evaluation evidence. The authors adopted a framework synthesis approach in order to cover a broader scope addressing a range of complementary policy-relevant questions, have greater flexibility of analysis (by allowing emerging findings to inform analysis) and facilitate context and intervention-specific learning.

The authors constructed a conceptual framework based on existing empirical studies, 'think pieces', interrogation of researchers familiar with social franchising and consultation with various global and local social franchising stakeholders. The resulting framework addressed a broad range of complementary policy-relevant questions about the intervention beyond effectiveness, including questions on reach (adoption by franchisees and service users), implementation and quality of care (adherence and integrity), equity and sustainability. User engagement was a continuous process throughout the review, with stakeholders being asked to comment on ongoing work, assist in identification of eligible studies and reflect on the findings and their implications.

The authors included all types of evaluation evidence (excluding opinion pieces, policy documents and non-systematic reviews). They drew on randomised and quasi-experimental evidence to answer effectiveness questions, and observational studies such as surveys, cohort studies, case–control studies and case studies to answer questions on reach, implementation, equity and sustainability. The quality of all studies was assessed using relevant quality assessment tools and the quality rating was used to weigh the strength of the evidence when drawing conclusions and recommendations for policy and practice.

Although the review did not find sufficient evidence on the effectiveness of social franchising, the framework synthesis approach contributed to our knowledge about the reach, quality of care, client satisfaction and equity of access and targeting of social franchising across a variety of settings.

The review found that across settings, social franchising was not associated with an increase in client volumes or use of treatment for sexually transmitted diseases. However, the authors found mixed outcomes for changes in unmet need for family planning services. Clients of franchises appeared to be satisfied with the quality of service received and expressed intent to return for future health services. Franchise providers were more likely to have received training compared with non-franchise private providers, although this was due to associations with government service in the past rather than the franchise. The review found mixed results about equitable access

to social franchises, especially with regards to social franchise services reaching the young, poor and/or illiterate parts of the population and clinics in low-income urban areas did not always benefit the targeted low-income populations.

In addition to the need to conduct more rigorous effectiveness studies of social franchising, the review indentifies additional dimensions for future research, including a better coverage of geographic regions (particularly Central Asia and South America), service delivery areas (especially family planning), implementation and equity considerations and the effects of different models of social franchising.

Source: Koehlmoos *et al*. (2011).

3.4. Thematic synthesis

A team of researchers at the EPPI Centre at the University of London have developed and applied thematic synthesis to the analysis of qualitative research (Thomas and Harden 2008). This method draws on 'thematic analysis' methods used for primary research and involves three steps: coding of text, developing descriptive themes and generating analytical themes. The first step involves coding of the findings of included studies 'line-by-line'. Codes are created inductively and can result in a large number of codes; for example, Thomas and Harden (2008) generate 36 different codes from eight studies. The second step involves grouping together similar codes and creating new codes or descriptive themes that cover several of the initial codes. Although the first two steps remain close to the original studies, the third step, generating the analytical themes, means going beyond this. Thomas and Harden (2008) compare this stage to the third-order interpretations of meta-ethnography (see below). In the example, the included studies focused on children's views about food and healthy eating and the review questions focused on barriers to, and facilitators of, healthy eating. At this stage, the reviewers used the descriptive themes from the analysis of children's views to infer barriers and facilitators to healthy eating and implications for interventions. This was first done independently by the reviewers and then as a group; through discussions and re-examination, six analytical themes emerged, which were translated into recommendations for interventions.

This approach to qualitative synthesis has been applied to a number of systematic reviews (Thomas *et al*. 2003, 2007, Harden *et al*. 2006) of both intervention and non-intervention studies and provides an example of a systematic review that includes both thematic synthesis and a meta-analysis in the same article.[1]

3.5. Meta-ethnography

Meta-ethnography was first developed as an approach to synthesis of qualitative, ethnographic studies by Noblit and Hare (1988), in their review of school desegregation in the United States. Meta-ethnography is an interpretive synthesis approach that re-conceptualises key themes and synthesises and extends the findings of individual studies (Williamson *et al*. 2009). Both the methodological and applied literature on meta-ethnography have been expanding in recent years, making it one of the most widely applied approaches to qualitative synthesis (for example, Britten *et al*. 2002, Munro *et al*. 2007, Atkins *et al*. 2008, Williamson *et al*. 2009). It provides a method for explaining different findings and it has therefore been suggested it might be particularly useful in informing policy (Mays *et al.* 2005b, pp. 9–10).

Meta-ethnography starts with identifying a research question that can be informed by qualitative research (for example, seeking to explain treatment adherence Munro *et al*.

2007) and deciding on the focus of the synthesis (inclusion criteria). The searching and selection of included studies should be systematic but may not need to be comprehensive or exhaustive since the aim of meta-ethnography is to produce theoretical generalisations rather than statistical generalisations (Pope *et al.* 2007). The included studies are quality-assessed and repeatedly read to extract details on study setting, population and intervention, and to identify primary themes (or first-order constructs; reported by participants) and secondary themes and concepts (or second-order constructs; interpretations by study authors) (Britten *et al*. 2002).

The synthesis process then involves d*etermining how the studies are related* by examining primary and secondary themes and concepts across studies and organising them into relevant tertiary categories (third-order constructs; interpretations by reviewers). In order to do this, the studies are translated *into one another* based on similar or re-occurring concepts (Atkins *et al*. 2008). These translations are then synthesised using one of three alternative methods.

Reciprocal translation is used when the concepts in multiple studies agree and correspond with one another and can be synthesised into common overarching themes (Campbell *et al*. 2002, Pope *et al.* 2007, Thomas *et al.* 2012). Refutational synthesis is used when the identified concepts refute each other or adopt competing ideological or disciplinary perspectives (Campbell *et al*. 2002, Pope *et al.* 2007, Thomas *et al.* 2012). The 'lines of argument' synthesis is about inference and involves an additional interpretive step that links the translations and interpretations and integrates the similarities and differences between concepts into a holistic scheme (Pope *et al.* 2007). According to Noblit and Hare, 'the goal of the lines-of-argument synthesis is to discover a "whole" among a set of parts' (1988, p. 63).

The synthesis then needs to be expressed in an accessible (and transparent) format, for example diagrammatically. Noblit and Hare (1988) highlight the subjective nature of the synthesis process, noting that reviewer's values and understanding of the studies influence synthesis, which is inevitably only one of several possible interpretations of the evidence. The above outlined process should not necessarily be seen as a linear process, and some repeated iterations of steps may need to be followed. Box 2 provides an example of a systematic review using meta-ethnography. For more details and examples of the three synthesis approaches (reciprocal translation, refutational synthesis and lines-of-argument synthesis), see Noblit and Hare (1988).

Box 2. Limits to modern contraceptive use among young women in developing countries: a systematic review of qualitative research

The review by Williamson and colleagues (2009) uses meta-ethnography to synthesise qualitative evidence on the factors that limit the adoption of modern contraceptive methods reported by young women in low- and middle-income countries.

The authors included qualitative studies of young women's own reports on the uptake, use or choice of discontinuation of modern contraceptive methods such as condoms and oral/injectable hormonal methods. They assessed the quality of the included studies with regards to reporting and appropriateness of analysis and interpretation.

The authors then identified, extracted and compared key themes from each included study, identifying similarities and contradictions across studies. This translation of the identified themes across studies was then used to arrive at a 'lines-of-evidence synthesis', which developed broader understandings while maintaining the specificity of

the individual included studies. The authors illustrate the identified lines-of-argument categories with representative quotes from the included studies.

The synthesis identified four main categories of limitations to contraception use by young women in low- and middle-income countries: (1) lack of education and knowledge about pregnancy risk, prevention and access to modern contraceptives, (2) concerns about hormonal contraceptive side-effects, menstrual disruption and fertility fears, (3) relationship considerations and partner pressure and control, particularly regarding contraception use and pregnancy and (4) desire to protect reputations and social status in the face of social disapproval of premarital sex and pregnancy. Use of hormonal contraceptives was mostly limited by lack of knowledge, limited access and fears of side-effects. Condoms were often viewed as a more accessible and acceptable method, although their use was limited by their association with disease, promiscuity and commercial sex. As a result of the identified difficulties associated with the use of modern contraceptive methods, young women were more likely to rely on traditional methods such as periodic abstinence, withdrawal or a range of herbal mixtures from traditional healers.

The authors conclude that interventions to promote the use of modern contraceptives need to counter the negative perceptions of modern contraceptive methods through the use of multi-faceted approaches that target all levels of society. Interventions need to challenge inaccurate beliefs and cultural norms around fertility, and provide more targeted promotion of life skills, support and access to youth-friendly services for adolescents. The involvement of both sexes in sexual and reproductive health interventions is important (particularly for condom use), and the dual role of condoms as birth control and prevention from sexually transmitted diseases should be capitalised on.

The synthesis advances our understanding of why adolescent reproductive health interventions in developing countries have had only limited effects on contraception use to date.

Source: Williamson *et al*. (2009).

3.6. Realist synthesis

Realist synthesis has been proposed as an approach to systematic reviews of evidence on complex interventions (Pawson *et al*. 2005). It is based on realist evaluation, which 'seeks to unpack the mechanism of *how* complex programmes work (or *why* they fail) in particular contexts and settings' (p. 21). The realist approach differs from the clinical trials model of establishing causality through experiments by focusing on context (the economic, social, legal, administrative and political circumstances in which the intervention takes place) and mechanisms (the programme theory of change), which determine impact on outcomes. Realist synthesis is theory-driven and the first step includes articulating the key theories of how interventions work, which are then explored in the review (Pawson 2006). It proposes to focus on programme mechanisms, rather than on the types of programmes as the unit of analysis and argues that similar mechanisms are likely to be operating in interventions across different sectors.

Realist synthesis has received increased attention in recent years, including in the context of systematic reviews in international development. It is worth noting that the term 'realist synthesis' has been used in different ways, including to explore a theory and its application in different contexts (Pawson 2002), to understand which aspects of interventions shape impacts in different contexts in the framework of effectiveness

reviews (Greenhalgh *et al.* 2007) and to develop a theory by exploring the dimensions of interventions with similar goals (Carr *et al*. 2011).

Pawson's original work took as its starting point a single theory (for instance, naming and shaming for encouraging appropriate behaviour) and explored the evidence available from different contexts (for example, protection of children against paedophiles and cars against thieves) to explore and refine the theory (Pawson 2002). A subsequent 'realist' review of school feeding programmes took as its starting point the evidence synthesised by a review of effectiveness in order to improve understanding of programme effectiveness (Greenhalgh *et al.* 2007; see also van der Knapp *et al*. (2006) for a similar application).

It is worth noting that none of these studies adopted the inclusive approach to evidence proposed by Pawson *et al.* (2005) but limited the realist review to collecting additional data on context and mechanisms from studies included in the systematic review of effectiveness. Their emphasis on the question of 'what works' raised concerns for these authors about including study designs not considered appropriate for answering causal questions (van der Knapp *et al*. 2008). The idea of focusing on programme mechanisms, rather than types of programmes as the unit of analysis is indeed attractive, especially in an area where evidence on effectiveness is still limited, although combining the approach with standards for traditional effectiveness reviews is advisable if applied to systematic reviews that also aim to answer causal questions.

3.7. Meta-narrative

The methods considered above offer a choice of approaches to synthesis, where that choice is determined by the nature of the question, the initial clarity of the concepts and relationships and whether the purpose is to generate, explore or test a mid-range programme theory. Working at a deeper, theoretical level, meta-narrative reviews address more critically the assumptions underpinning studies, not just the theories underpinning interventions (Greenhalgh 2011). This approach has been developed to synthesise complex evidence, and applied to reviews of diffusion and sustainability of innovations in health service delivery and organisation (Greenhalgh *et al*. 2005a, 2005b), the meaning of community within and across research traditions (Bertotti *et al.* 2011) and eliciting patient's perspective in patient-centred outcomes research (Garces *et al*. 2012). The approach has many parallels with realist synthesis and includes examining storylines from different research traditions. The reviewers then focus on identifying relevant concepts from each tradition, creating the meta-narrative for each tradition first, and then bringing them together in a synthesis around the key dimensions, enabling a systematic analysis of contradictions. The key drawback of this approach is the substantial resources and cross-disciplinary expertise that need to be devoted to this type of synthesis. Conversely, international development can be seen as an inherently cross-disciplinary field, and two of the reviews mentioned above have included research traditions rooted in developing countries (Greenhalgh *et al*. 2005a, Garces *et al*. 2012), thereby offering the opportunity to make use of learning across economic divides.

4. Discussion

Systematic review methodology is a powerful tool for bringing together evidence from a range of different studies in a transparent and systematic manner. Critiques of systematic reviews as being too narrow (Mallett 2012) are problematic, because they are based on the assumption that the methodology can only be applied to reviews of effects. But just as primary research can be applied to answer a range of different questions, the logic of

systematic reviews can also be applied to syntheses of evidence addressing a range of different questions (Gough *et al.* 2012).

Whatever research questions authors aim to answer, they should apply the key principles of systematic reviewing, by being comprehensive, systematic and transparent. Approaching the literature systematically means adopting explicit inclusion criteria, a systematic search, selecting studies for inclusion based on predetermined inclusion criteria, critically appraising included studies and collecting and synthesising data from the included studies in a transparent and systematic manner (Saini and Shlonsky 2012).

In addition to adopting different inclusion criteria for the types of studies and determining the most appropriate methods of synthesis, applying the logic of systematic review methodology to qualitative evidence also has implications for the other stages of the review process, especially the search and critical appraisal of included studies (Noyes *et al*. 2011).[2] The latter is a particularly contentious issue, with a lack of agreement on the best approach for assessing the quality of qualitative evidence (Noyes *et al*. 2011; see also Snilstveit, (2012) for more details on this debate).

The aim of systematic reviews of qualitative evidence is not to assess the effects of interventions, although they can complement systematic reviews of effectiveness studies and help explain and interpret results of syntheses of evidence on 'what works', as well as answering other questions, which can be useful in informing policy and practice (Noyes *et al*. 2011, Saini and Shlonsky 2012). Such syntheses can be conducted as stand-alone reviews or as part of multi-component reviews, including synthesis of evidence on effects (Noyes *et al*. 2011, Gough and Thomas 2012, Snilstveit 2012).

Moreover, in some instances, policy-makers are asking questions of 'what works' with the intention of this informing a policy decision, but it is known that the evidence on effectiveness is limited. If evidence about effects is weak, then systematic reviews might more usefully focus on research that can lead to developing/refining interventions, for instance by synthesising feasibility and acceptability studies. If evidence about effects is heterogeneous, then work needs to focus on identifying the active ingredients or synergistic or antagonistic components of interventions (for example, Candy *et al*. 2011). If interventions do not exist then work reviewing epidemiological studies of risk factors as well as studies of people's views and experiences might be useful in developing a programme theory and designing interventions.

There is less agreement on the best approaches to systematic reviews that go beyond the questions of effects. Therefore, in addition to informing policy and practice, systematic reviews of qualitative evidence to answer questions pertinent to international development can also contribute to informing these debates and moving the field of qualitative synthesis forward.

5. Conclusion

Systematic review methodology can be applied to a range of questions to offer understanding of problems facing low- and middle-income countries and provide greater direction for developing evidence-informed solutions. Although systematic reviews are an important *tool* for evidence-informed decision-making, they often provide partial answers to broader policy questions and 'are not themselves decisions' (Petticrew and Roberts, p. 45). Nevertheless, broader application of syntheses addressing a range of questions allow for better utilisation of the existing evidence base and can support the development of more evidence-informed polices.

Notes

1. It includes both intervention and non-intervention studies and incorporates three main syntheses: (1) meta-analysis of quantitative evaluations of interventions attempting to increase consumption of fruit and vegetables among children; (2) thematic qualitative synthesis of 'non-intervention' research focusing on children's views of healthy eating; and (3) 'cross-study synthesis' using findings from the thematic synthesis to interpret the quantitative results.
2. Pope *et al*. (2007) and Saini and Shlonsky (2012) offer comprehensive textbooks on qualitative synthesis. Hannes and Lockewood (2012) provide a range of worked examples of a number of qualitative synthesis methods. Snilstveit (2012) reviews some of the debates around methods for qualitative synthesis and provides guidance on resources for authors undertaking qualitative synthesis.

References

Atkins, S., *et al*., 2008. Conducting a meta-ethnography of qualitative literature: lessons learnt. *BMC medical research methodology*, 8 (21), doi: 10.1186/1471-2288-8-21.

Barnett-Page, E. and Thomas, J., 2009. Methods for the synthesis of qualitative research: a critical review. *BMC medical research methodology*, 9 (59), doi: 10.1186/1471-2288-9-59.

Basu, S., *et al*., 2012. Comparative performance of private and public healthcare systems in low- and middle-income countries: a systematic review. *PLoS medicine*, 9 (6), e1001244, doi: 10.1371/journal.pmed.1001244.

Berendes, S., *et al*., 2011. Quality of private and public ambulatory health care in low and middle income countries: systematic review of comparative studies. *PLoS medicine* [online], 8 (4), e1000433. Available from: http://www.plosmedicine.org/article/info%3Adoi%2F10.1371%2Fjournal.pmed.1000433 [Accessed 23 June 2012].

Berg, B.L., 2009. *Qualitative research methods for the social sciences*. Boston, MA: Allyn & Bacon.

Bertotti, M., Jamal, F., and Harden, A., 2011. *Connected communities: a review of conceptualisations and meanings of* '*community*' *within and across research traditions: a metanarrative approach* [online]. Available from: http://www.ahrc.ac.uk/FundingOpportunities/Documents/CC%20scoping%20studies/Bertotti.pdf [Accessed 5 July 2012].

Britten, N., *et al*., 2002. Using meta ethnography to synthesise qualitative research: a worked example. *Journal of health services research and policy*, 7 (4), 209–215.

Bryman, A., 2001. *Social research methods*. Oxford: Oxford University Press.

Campbell, R., *et al*., 2002. Evaluating meta-ethnography: a synthesis of qualitative research on lay experiences of diabetes and diabetes care. *Social science and medicine*, 56, 671–684.

Candy, B., *et al*., 2011. Using qualitative synthesis to explore heterogeneity of complex interventions. *BMC medical research methodology*, 11 (124). doi:10.1186/1471-2288-11-124. Available from: http://www.biomedcentral.com/1471-2288/11/124 [Accessed 25 June 2012].

Carr, S., *et al*., 2011. An evidence synthesis of qualitative and quantitative research on component intervention techniques, effectiveness, cost-effectiveness, equity and acceptability of different versions of health-related lifestyle advisor role in improving health. *Health technology assessment*, 15 (09).

Carroll, C., Booth, A., and Cooper, K., 2011. A worked example of 'best fit' framework synthesis: a systematic review of views concerning the taking of some potential chemopreventive agents. *BMC medical research methodology*, 11, 29. doi: 10.1186/1471-2288-11-29.

Cresswell, J., 2003. *Research design: qualitative, quantitative and mixed methods approaches*. London: Sage.

Dixon-Woods, M., *et al*., 2005. Synthesising qualitative and quantitative evidence: a review of possible methods. *Journal of health service research and policy*, 10 (1), 45–53.

Evans, D. and FitzGerald, M., 2002. Reasons for physically restraining patients and residents: a systematic review and content analysis. *International journal of nursing studies*, 39 (7), 735–743.

Garces, J.P.D., *et al*., 2012. *Eliciting patient perspective in patient-centered outcomes research: a meta narrative systematic review* [online]. Available from: http://www.pcori.org/assets/Eliciting-Patient-Perspective-in-Patient-Centered-Outcomes-Research-A-Meta-Narrative-Systematic-Review.pdf [Accessed 5 July 2012].

Glenton, C., Lewin, S., and Scheel, I.B., 2011. Still too little qualitative research to shed light on results from reviews of effectiveness trials: a case study of a Cochrane review on the use of lay health workers. *Implementation science*, 6, 53.

Gough, D. and Thomas, J., 2012. Commonality and diversity in reviews. *In*: D. Gough, S. Oliver, and J. Thomas, eds. *An introduction to systematic reviews*. London: Sage, 38–65.
Gough, D., Oliver, S., and Thomas, J., eds., 2012. Introducing systematic reviews. *In*: *An introduction to systematic reviews*. London: Sage, 1–16.
Greenhalgh, T., 2011. Why do we always end up here? Evidence-based medicine's conceptual cul-de-sacs and some off-road alternative routes. *Journal of primary health care*, 4 (2), 92–97.
Greenhalgh, T., *et al*., 2005a. *Diffusion of innovations in health service organisations: a systematic literature review*. Oxford: Blackwell Publishing.
Greenhalgh, T., *et al*., 2005b. Storylines of research in diffusion of innovation: a meta-narrative approach to systematic review. *Social science and medicine*, 61, 417–430.
Greenhalgh, T., Kristiansson, E., and Robinson, V., 2007. Realist review to understand the efficacy of school feeding programmes. *British medical journal*, 335, 858–861.
Hannes, K., 2011. Critical appraisal of qualitative research. *In*: J. Noyes, *et al.*, eds. *Supplementary guidance for inclusion of qualitative research in Cochrane systematic reviews of interventions 570* [online], Version 1. The Cochrane Collaboration Qualitative Methods Group. Available from: http://cqrmg.cochrane.org/supplemental-handbook-guidance [Accessed 9 November 2011].
Hannes, K. and Lockwood, C., 2012. *Synthesizing qualitative research: choosing the right approach*. Oxford: Wiley-Blackwell.
Harden, A. and Gough, D., 2012. Quality and relevance appraisal. *In*: D. Gough, S. Oliver, and J. Thomas, eds. *An introduction to systematic reviews*. London: Sage, 153–178.
Harden, A., *et al*., 2006. *Young people, pregnancy and social exclusion: a systematic synthesis of research evidence to identify effective, appropriate and promising approaches for prevention and support*. London: EPPI Centre, Social Science Research Unit, Institute of Education, University of London.
Kiwanuka, S.N., *et al*., 2011. *Dual practice regulatory mechanisms in the health sector: a systematic review of approaches and implementation*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
Koehlmoos, T., *et al*., 2011. *Social franchising evaluations: a scoping review*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
Lavis, J., 2009. How can we support the use of systematic reviews in policymaking. *PLoS medicine*, 6 (11), e1000141. doi:10.1371/journal.pmed.1000141.
Lewin, S., *et al*., 2012. Guidance for evidence-informed policies about health systems: assessing how much confidence to place in the research evidence. *PLoS medicine*, 9 (3), e1001187. doi:10.1371/journal.pmed.1001187.
Mallett, R., *et al*., 2012. The benefits and challenges of using systematic reviews in international development research. *Journal of development effectiveness*, 4 (3), 445–455.
Mays, N., Pope, C., and Popay, J., 2005a. Systematically reviewing qualitative and quantitative evidence to inform management and policy-making in the health field. *Journal of health service research and policy*, 10 (Suppl 1), 6–20.
Mays, N., Pope, C., and Popay, J., 2005b. *Details of approaches to synthesis. A methodological appendix to the paper: systematically reviewing qualitative and quantitative evidence to inform management and policy making in the health field* [online]. Available from: http://www.google.co.uk/url?a=t&rct=j&q=&esrc=s&source=web&cd=2&ved=0CEoQFjAB&url=http%3A%2F%2Fciteseerx.ist.psu.edu%2Fviewdoc%2Fdownload%3Fdoi%3D10.1.1.113.2530%26rep%3Drep1%26type%3Dpdf&ei=6FkOUPWlEuHT0QWx4YCQDA&usg=AFQjCNHRxbH1y4UNKCMGS8GjdL5jcfZz3A [Accessed 24 July 2012].
Munro, S.A., *et al*., 2007. Patient adherence to tuberculosis treatment: a systematic review of qualitative research. *PLoS medicine*, 4 (7), 1230–1245.
NICE, 2009. *Methods for the development of NICE public health guidance*. 2nd ed. London: NICE.
Noblit, G.W. and Hare, R.D., 1988. *Meta-ethnography: synthesizing qualitatives Studies*. Qualitative Research Methods (Vol. 11). Newbury Park: Sage Publications.
Noyes, J. and Lewin, S., 2011. Supplemental guidance on selecting a method of qualitative evidence synthesis, and integrating qualitative evidence with Cochrane intervention reviews. *In*: J. Noyes, *et al*., eds. *Supplementary guidance for inclusion of qualitative research in Cochrane systematic reviews of interventions* [online], Version 1, The Cochrane Collaboration Qualitative Methods Group. Available from: http://cqrmg.cochrane.org/supplemental-handbook-guidance [Accessed 7 January 2012].

Noyes, J., *et al*., 2011. Qualitative research and Cochrane reviews. *In*: J.P.T. Higgins and S. Green, eds. *Cochrane handbook for systematic reviews of interventions* [online], Version 5.0.1. The Cochrane Collaboration. Available from: http://www.cochrane-handbook.org [Accessed 29 November 2011].

Oliver, S., *et al*., 2008. A multidimensional conceptual framework for analyzing public involvement in health services research. *Health expectations*, 11 (1), 72–84.

Pawson, R., 2002. Evidence and policy and naming and shaming. *Policy studies*, 23 (3/4), 211–230.

Pawson, R., 2006. *Realist synthesis: supplementary reading 6: digging for nuggets: how 'bad' research can yield 'good' evidence* [online]. Available from: http://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CEgQFjAA&url=http%3A%2F%2Fwww.leeds.ac.uk%2Frealistsynthesis%2FSupread6.doc&ei=EIX0T-mwGOHK0QWPxoGVBw&usg=AFQjCNGPBe7nPxACjj9JTZK3YJl1xyGNjw [Accessed 4 July 2012].

Pawson, R., *et al*., 2005. Realist review – a new method of systematic review designed for complex policy interventions. *Journal of health policy and research*, 10 (Suppl 1), 21–34.

Petticrew, M. and Roberts, H., 2003. Evidence, hierarchies, and typologies: horses for courses. *Journal of epidemiology and community health*, 57, 527–529.

Petticrew, M. and Roberts, H., 2006. *Systematic reviews in the social sciences – a practical guide*. Oxford: Blackwell Publishing.

Popay, J., *et al*., 2006. *Guidance on the conduct of narrative synthesis in systematic reviews* [online]. Version 1: April 2006, ESRC Methods Programme. Available from: http://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=6&ved=0CF4QFjAF&url=http%3A%2F%2Fciteseerx.ist.psu.edu%2Fviewdoc%2Fdownload%3Fdoi%3D10.1.1.178.3100%26rep%3Drep1%26type%3Dpdf&ei=3QMQUIqxK-GK0AXfgoGoBA&usg=AFQjCNGrW1dB7FBraR9t3ITS0FjiWqqg-A [Accessed 25 July 2012].

Pope, C., Mays, N., and Popay, J., 2007. *Synthesising qualitative and quantitative health evidence:a guide to methods*. Maidenhead: Open University Press, 79–89.

Pope, C., Ziebland, S., and Mays, N., 2000. Qualitative research in health care: analysing qualitative data. *British medical journal*, 320, 114–116.

Pullin, A.S., *et al*., 2012. Human well-being impacts of terrestrial protected areas? *CEE protocol 11-009* [online]. Collaboration for Environmental Evidence. Available from: www.environmentalevidence.org/SR11009.html [Accessed 24 June 2012].

Ritchie, J. and Spencer, L., 1994. Qualitative data analysis for applied policy research. *In*: A. Bryman and R.G. Burgess, eds. *Analyzing qualitative data*. London: Routledge.

Saini, M., 2012. *Qualitative synthesis to help explore complex interventions: an evolving approach within systematic reviews*, presented at the Campbell Colloquium, Copenhagen, 29–31 May 2012. Available from: http://www.sfi.dk/Default.aspx?ID=11043 [Accessed 26 June 2012].

Saini, M. and Shlonsky, A., 2012. *Systematic synthesis of qualitative research*. Oxford: Oxford University Press.

Snilstveit, B., Oliver, S., and Vojtkova, M., 2012. Narrative approaches to systematic review and synthesis of evidence for international development policy and practice. *Journal of development effectiveness*, 4 (3), 409–429.

Spencer, L., *et al*., 2003. *Quality in qualitative evaluation: a framework for assessing research evidence*. London: Government Chief Social Researcher's Office.

Thomas, J., *et al*., 2003. *Children and healthy eating: a systematic review of barriers and facilitators*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Thomas, J., *et al*., 2007. *Accidental injury, risk-taking behaviour and the social circumstances in which young people live*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Thomas, J. and Harden, A., 2008. Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC medical research methodology*, 8 (45), 1–10.

Thomas, J., Harden, A., and Newman, M., 2012. Synthesis: combining results systematically and appropriately. *In*: D. Gough, S. Oliver, and J. Thomas, eds. *An introduction to systematic reviews*. London: Sage, 179–226.

Tugwell, P., *et al*., 2010. Assessing equity in systematic reviews: realising the recommendations of the commission on social determinants of health. *British medical journal*, 341, 873–877.

van der Knapp, L.M., *et al*., 2008. Combining Campbell standards and the realist evaluation approach – the best of two worlds? *American journal of evaluation*, 29 (1), 48–57.

van der Knapp, L.M., Nijssen, L.T.J., and Bogaerts, S., 2006. *Violence defied? a review of prevention of violence in the public and semi-public domain*. The Hague: Research and Documentation Centre, Ministry of Justice.

Voils, C.I., *et al*., 2008. Making sense of qualitative and quantitative findings in mixed research synthesis studies. *Field methods*, 20 (1), 3–25.

Waddington, H., *et al*., 2012. How to do a good systematic review of effects in international development: a tool kit. *Journal of development effectiveness*, 4 (3), 359–387.

White, H., 2009. Theory-based impact evaluation: principles and practice. *Journal of development effectiveness*, 1 (3), 271–284.

Williamson, L.M., *et al*., 2009. Limits to modern contraceptive use among women in developing countries: a systematic review of qualitative research. *Reproductive health*, 6 (3), 1–12.