Qualitative research synthesis for health policy analysis: what does it entail and what does it offer?

Lucy Gilson1,2

1Health Policy and Systems Division, School of Public Health and Family Medicine, University of Cape Town, Anzio Road, Observatory 7925, Cape Town, South Africa and 2Health Economics and Systems Analysis Group, Department of Global Health and Development, London School of Hygiene and Tropical Medicine, 15-17 Tavistock Place, London WC1H 9SH, UK

This set of articles has its roots in the 2008 special edition of Health Policy and Planning, titled ‘Future directions for health policy analysis: a tribute to the work of Professor Gill Walt’, which among other things called for ‘better use of the existing but often descriptive body of policy analysis through synthesis of existing case study material’ (Gilson et al. 2008, p. 292). Picking up that challenge, this edition presents a set of five articles that, through synthesis of available research, seek to consolidate and develop the body of health policy analysis (HPA) work in low- and middle-income countries (LMICs). This work is currently fragmented across geographic settings and policy issues, is more descriptive than analytic and is weakly theorized (Gilson and Raphaely 2008; Walt et al. 2008).

Policy analysis starts from the understanding that policy making is a process of continuing interaction among institutions (the structure and rules which shape how decisions are made), interests (groups and individuals who stand to gain or lose from change) and ideas (including arguments and evidence) (John 1998) (Gilson et al. 2008, p. 291). Such analysis is a legitimate area of academic inquiry and has practical importance for health system development.

Qualitative synthesis, meanwhile, can be defined as ‘as any methodology whereby study findings are systematically interpreted through a series of expert judgements to represent the meaning of the collected work. In a qualitative synthesis, the findings of qualitative studies—and sometimes mixed-methods and quantitative research—are pooled. Judgement-based qualitative methodologies are used to draw conclusions regarding the collective meanings of this pool of research’ (Bearman and Dawson 2013, p. 253).

Although gaining attention in health research, qualitative synthesis remains a new area of work. Set against the traditions of systematic review, the interpretive judgements involved raise concerns about bias, and there are also questions about whether it is appropriate to draw any form of generalization from qualitative data generated in specific places and times (Wallace et al. 2006; Thomas and Harden 2008).

Starting broadly from what Barnett-Page and Thomas (2009) identify as an objective idealist or critical realist epistemological position, we were, nonetheless, emboldened to explore the application of qualitative synthesis to HPA research. These articles are broadly based on the understandings that:

- qualitative studies can be systematically compared between countries and contexts, and across time;
- such comparison highlights empirical and conceptual patterns that hold across individual experiences—such as patterns of behaviour or features of context that influence policy change experience and
- these patterns represent analytic generalizations, i.e. empirically based conceptual conclusions that can be considered in other settings and times.

In systematic review terminology, we aimed to conduct knowledge support rather than decision support synthesis—using interpretive rather than integrative synthesis approaches (Dixon-Woods et al. 2005; Pope et al. 2007). Together the articles seek, first, both to map the terrain of some areas of HPA work and to generate theoretical and conceptual insights that can inform future empirical research. Second, as a methodological contribution, they demonstrate and test several different synthesis approaches: aspects of meta-study, narrative synthesis, framework synthesis, thematic synthesis and meta-ethnography.

Implementing the syntheses: common approaches and key issues

Process of working

The five articles were all developed through a collaborative, team-based process. Two full-team workshops, in 2009 and 2010, allowed us jointly to develop our overall approach and to share experience of implementing the early steps of review. On the basis of initial reflection on qualitative synthesis methodology, we together discussed the selection of questions and appropriate synthesis approaches, agreed to adopt a common literature search approach and reviewed experience of quality appraisal and data extraction. Early drafts of synthesis reports were then tested with a wider audience of policy analysts and
synthesis experts at a workshop held in Oxford in September 2010. The authors of each article then undertook the subsequent steps of article finalization, submission for publication and revision after review, with back up from the project co-ordinator (Gilson). These various processes allowed triangulation of various key judgements made within the team and with broader groups. The length of time from the Oxford workshop to the date of publication reflects the time demands of synthesis processes—especially when set against the parallel commitments of all team members.

Identifying the synthesis questions and selecting synthesis approaches

Drawing on our own prior knowledge of the field, as a team of experienced policy analysts, we sought to identify questions that would, through synthesis of current work, deepen understanding of policy processes. Because of its enduring relevance, despite its simplicity, we were guided by the ‘stages heuristic’ framework, which identifies delineated policy stages, each of which can be the subject of inquiry (Sabatier 2007). We also sought to identify questions that would allow us to explore different approaches to synthesis.

We addressed both broad questions allowing description of the terrain of work for some policy stages, and more specific questions focused on deeper exploration of some dimensions of these stages. Two articles (Berlan et al. and Erasmus et al.) focus on mapping a full set of literature addressing, respectively, policy formulation and adoption and policy implementation, as well as using different approaches (narrative synthesis and elements of meta-study), synthesizing the substantive concerns addressed within this literature. Berlan et al. derive conceptual ideas about the seven individual steps entailed in policy formulation, whereas Erasmus et al. tease out the types of issues addressed within implementation work and explore the disciplinary and theoretical influences over it. Walt and Gilson also review a full set of literature in applying framework synthesis to test and develop a pre-existing conceptual framework of agenda setting. The other two articles adopt inductive synthesis approaches to consider specific implementation questions with a set of deliberately sampled articles. Erasmus uses meta-ethnography to offer a synthetic reflection on how the theory of street-level bureaucracy has been used in the relevant LMIC literature, and Gilson et al. apply thematic synthesis in considering the practice of power in implementation.

Common review steps

Although qualitative synthesis ultimately entails interpretive judgement, it begins with a structured and systematic review process. In line with the commonly accepted precepts for this work (see for example, Gough et al. 2012), all five articles presented here were developed through a set of common steps, as outlined later.

Searching for literature

Simply identifying research articles addressing health policy processes is a challenge because there are diverse bodies of potentially relevant literature, drawn from different disciplinary traditions, located both in specialist health journals and broader social policy and development journals (Gilson and Raphaely 2008). An additional challenge for the articles presented here is that the boundaries of different stages of the policy process are not tightly specified or always reflected in key words, so abstracts sometimes offered promises of relevance not revealed in the full article (Berlan et al. and Walt and Gilson). Judgement was, therefore, needed just to identify articles for consideration. Similarly, it is hard to find qualitative health research through literature review, and some argue that it may be inappropriate to use completely reproducible search strategies (Dixon-Woods et al. 2006).

At the start of our work, we, therefore, made a key decision: for all syntheses, we would work primarily from the foundation of the first ever review of LMIC HPA literature (Gilson and Raphaely 2008), updated until 2009 (so covering the period 1994–2009) by using the same databases (PubMed and the International Bibliography of the Social Sciences). We also decided to work only with published literature.

Although more articles of relevance to our questions will by now have been published, we judge that the articles we considered provide an adequate base for the work presented here. Even when mapping work within a policy stage (Berlan et al., Erasmus et al. and Walt and Gilson), the syntheses do not primarily seek to combine and amalgamate data from a fully comprehensive set of articles (as is expected in the dominant forms of systematic review). Instead, all the articles presented here seek, through synthesis, to generate novel insights that hold beyond the particular body of work reviewed.

Appraising articles

We appraised article quality before finalizing the selection of articles included in our syntheses. Erasmus et al., for example, used a quality assessment checklist adapted from Wallace et al. (2006) together with review of the richness of the implementation description (Roen et al. 2006). Walt and Gilson, meanwhile, assessed quality by examining the extent to which each article outlined a clear framework and described its methods; how far the issue was contextualized; how reflective the authors were (acknowledging weaknesses or limitations) and how far they paid attention to bias.

However, applying a set of quality criteria often proved difficult because details on methods were limited, and across disciplinary traditions and journals, there are quite different expectations of what methodological details authors should provide (Walt and Gilson). Ultimately, as is fairly common in qualitative synthesis (Dixon Woods et al. 2007), we found holistic judgements of quality more persuasive than either overarching study design or specific criteria, as commonly used in systematic review. We follow Popay et al. (1998), therefore, in judging that a key marker of good quality is a description detailed enough to allow the reader to make sense of the experience presented. The team approach, meanwhile, allowed triangulation of quality judgements among reviewers.

Relevance to the review question was, moreover, a critical appraisal criterion. Indeed, for Erasmus, this was the main criterion applied. As noted earlier, simply identifying relevant literature is a challenge where the phenomena of focus are complex and emergent. In qualitative synthesis, it is anyway quite common to appraise articles, or even particular elements within them (Pawson 2006), on the basis of their credibility.
and contribution to the synthesis, rather than using a priori quality judgments (Dixon Woods et al. 2007; Pearson 2007).

**Sampling articles**

For three of the syntheses presented here, we used a comprehensive sample of articles from the broadly defined pool of good quality and relevant literature (Berlan et al., Erasmus et al. and Walt and Gilson).

However, sampling articles for interpretative synthesis is primarily a purposive process. Purposive or theoretical sampling might entail deliberately selecting a set of heterogeneous articles to allow comparison and contrast among them (Thomas and Harden 2008), for example, or because of their relevance to conceptual development (Dixon-Woods et al. 2005). Gilson et al., therefore, selected, two specific sets of articles because they presented implementation experience in two contrasting areas of policy change (system change: decentralization and service delivery change: reproductive health) that were judged to offer potentially different insights on the practice of power in implementation. Erasmus, meanwhile, only selected the four articles that explicitly drew on street level bureaucracy concepts in examining policy implementation experience,judging that all other identified articles considered street level bureaucracy concepts too superficially.

We recognize that purposive sampling has the limitation of perhaps excluding articles with important and relevant insights. However, even in a comprehensive set of articles, the absence of evidence on an issue cannot definitively be said to reflect an absence of experience: syntheses always work with an incomplete knowledge base (Gilson et al.). Ultimately, in any case, sampling always has to reflect the resources of the project (Dixon-Woods et al. 2006).

**Data extraction and presentation**

All systematic reviews require a systematic process of data extraction to allow subsequent review.

The first step is to identify which data to extract, e.g. findings or concepts? and then it is necessary to decide how to represent them in ways that allow comparison and contrast across articles or other categories of relevance. In this edition, Erasmus, following meta-ethnography practice, focused on extracting concepts and conceptual understandings, whereas the other articles focused primarily on extracting findings. All syntheses, however, used data from all sections of the articles they reviewed, rather than only from the formally defined findings section. Gilson et al. also extracted authorial judgements on the grounds that these context-rich insights were valuable for interpreting the data presented in the articles reviewed. Walt and Gilson, meanwhile, uses a pre-existing conceptual framework to help select the data they extracted, as they sought to test that framework against the experience reported in the articles. Notwithstanding their inductive synthesis approach, pre-existing understandings were also acknowledged to shape data extraction processes in all other articles—helping to define the boundaries of agenda setting and implementation (Berlan et al. and Erasmus et al.), to identify practices of power (Gilson et al.) and to identify the concepts of focus in the meta-ethnography of street level bureaucracy (Erasmus).

Different approaches were then applied in the practice of data extraction. Berlan et al. and Erasmus et al. used specifically developed, but quite open, templates, and Walt and Gilson used a template based on the conceptual framework they tested deductively. Erasmus, meanwhile, made detailed, handwritten notes on the articles reviewed, and Gilson et al. used Atlas.ti to extract and code data in one process. Where Erasmus focused on key concepts and themes within the articles and sought to maintain conceptual relationships among them, Gilson et al. followed the usual thematic analysis process (Thomas and Harden 2008) of, first, systematically and inductively coding the extracted data for meaning and content, and second, a generating and organizing descriptive themes derived from the data. Berlan et al. and Erasmus et al. also identified themes summarizing the data they extracted, allowing them to describe the terrain of work reported.

It is common practice in qualitative synthesis to display the extracted data, concepts or descriptive themes in matrices, to support the comparative analysis entailed in synthesis. In Walt and Gilson’s synthesis, the use of the pre-existing conceptual framework in data extraction enabled its subsequent comparison with data drawn from each article. Erasmus presents the details of the four articles used in this meta-ethnography in tables, whereas Gilson et al. use tables to compare and contrast experience across the two selected policy areas (decentralization and reproductive health) and Erasmus et al. to compare and contrast differences between disciplinary traditions, amongst other issues. As the articles also demonstrate, other presentation approaches (such as the use of different fonts or counts of articles in different categories) can also help to illuminate patterns in the data presented.

**Interpretation and synthesis**

Synthesis adds an additional step to the more conventional process of systematic review—a step in which judgements are made to generate new perspectives or new ideas. ‘At their core, interpretive methods for synthesis entail a process of qualitative re-interpretation or re-analysis of text-based forms of evidence’ (Pope et al. 2007, p. 75).

The process of judgement making and the types of judgements made differ between synthesis approaches. In meta-study, differences in disciplinary and theoretical perspectives are considered, as explored in Erasmus et al.’s mapping of implementation literature, whereas in thematic analysis, differences in contexts or types of participants are examined (e.g. Gilson et al. compare differences between policy areas and between provider groups).

Framework and thematic synthesis, meanwhile, go beyond describing and summarizing data across studies, to, in a second step, produce fresh interpretations of the phenomena under review. Within the more descriptive approach entailed in mapping the body of implementation literature, Erasmus et al. also used thematic synthesis to develop an interpretive analysis of the articles’ underlying ideas about how to improve implementation. Gilson et al., meanwhile, present a thematic synthesis around the forms and sources of power used by front line workers in implementation and the factors influencing their use, and Walt and Gilson use framework synthesis to generate fresh ideas to embellish and strengthen the Shiffman and Smith...
(2007) framework of agenda setting. The bit in the middle article (Berlan et al.), finally, uses narrative synthesis to develop some analytic clarity about the processes that fall between agenda setting and implementation, identifying seven distinct categories. Where framework or thematic synthesis is based on comparison across empirical data (Gilson et al. and Walt and Gilson), meta-ethnography develops theoretical insights through the translation of concepts across studies. Erasmus presents a worked example of this approach, which entails taking concepts from one study and recognizing them in another study (Thomas and Harden 2008).

The judgements entailed in these syntheses raise, for some, concerns about bias or inappropriate comparison across context. Such judgements must clearly be developed through systematic and reflexive processes, although being transparent is not straightforward. Involving multiple analysts can be important in developing and testing judgements, as well as iteratively testing whether abstractions are translatable across contexts (Pope et al. 2007; Thomas and Harden 2008). For example, Gilson et al. compare data across articles, across policy areas and across provider groupings, as well as comparing judgements among three analysts, in generating a synthesis about the practices of power used in implementation. Critical engagement between data and theory—as specifically demonstrated in Erasmus and Walt and Gilson—is also important for some forms of synthesis (Pearson 2007), whereas being methodologically reflexive is always necessary (Dixon-Woods et al. 2006).

Ultimately, moreover, it is the judgements entailed in interpretive synthesis that provide new ideas and insights—moving beyond the findings of any individual study included in the synthesis.

The contributions of this set of articles
The first contribution of these articles is methodological. They demonstrate the overall approach of qualitative synthesis and its application in HPA. They outline the key synthesis steps, identify challenges and present some novel elements.

Second, the articles make substantive contributions to HPA work in LMICs—by mapping current bodies of work, identifying research opportunities, developing new insight about policy processes and generating new concepts and frameworks to use in future empirical research (Table 1).

Together the syntheses provide a platform from which to conduct further and richer HPA empirical work and from which to pursue additional secondary research that will deepen our knowledge base. One HPA methodological recommendation is to conduct research that acknowledges and investigates the influence and power of ideas, understandings and meanings as part of every policy stage (Yanow 2003).

Finally, although the articles are framed as knowledge-support synthesis, they can also inform action to strengthen change in health systems (decision-support synthesis). The conceptual insights and thinking frameworks that the articles provide—whether on the forces shaping agenda setting, key activities in policy formulation, the influences over implementation and front line workers (street level bureaucrats) or their exercise of power—all offer ideas about how to influence and manage policy change more effectively. For reflective practitioners, whether policy makers or managers or advocates, these synthetic insights have practical value.

Table 1 Summary of the articles’ contributions

<table>
<thead>
<tr>
<th>Policy stage</th>
<th>Article contributions</th>
<th>Article</th>
</tr>
</thead>
</table>
| Agenda setting        | • Maps literature, identifying limits of current work and opportunities for future research  
                        • Confirms and develops the Shiffman and Smith (2007) framework, supporting its future use in prospective, cross-national and/or cross-policy studies of agenda setting | Walt and Gilson |
| Policy formulation and adoption | • Maps literature, identifying limits of current work and opportunities for future research  
                                 • Provides analytic clarity about specific processes encompassed within this policy stage  
                                 • Identifies specific questions to guide future research for this policy stage | Berlan et al. |
| Implementation        | • Maps literature, identifying limits of current work and opportunities for future research  
                        • Highlights key influences over implementation and possible implications for implementation management  
                        • Provides directions for future implementation research  
                        • Provides and deepens insight into the factors shaping street level bureaucrats behaviour  
                        • Illuminates the possibility SLBs may support (rather than only oppose) new policies  
                        • Demonstrates how discretionary power is exercised in daily decision making by front line providers  
                        • Points to strategies for managing this discretionary power  
                        • Generates a framework that can be applied and tested in future empirical work to deepen insights into discretionary power | Erasmus et al.  
                        Gilson et al.  

Acknowledgments
Thanks to all members of the synthesis team, to the reviewers of the early draft papers presented at the Oxford 2010 workshop and all other participants.

Funding
This special edition is the product of the project ‘Strengthening health policy analysis research and training through a focus on approaches to comparative and synthesis
analyses’, funded by the Alliance for Health Policy and Systems Research.

Conflict of interest statement: None declared.

References


Thomas J, Harden A. 2008. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Medical Research Methodology 8: 45.

