Evaluating and synthesizing qualitative research: the need to develop a distinctive approach

Rosaline S. Barbour MA PhD and Michael Barbour BSc Dip Lib ALA
1Department of General Practice, University of Glasgow, Glasgow, UK
2University Library, University of Glasgow, Glasgow, UK

Correspondence
Dr Rosaline S. Barbour
Department of General Practice
University of Glasgow
4 Lancaster Crescent
Glasgow G12 0RR
UK

Keywords: critical appraisal, qualitative research, synthesis, systematic review

Abstract
The growing popularity of qualitative research has led to calls for it to be incorporated into the evidence base. It is argued that, in seeking to respond to this challenge, it is important that we recognize the important differences between qualitative and quantitative research and that we take this into account in developing a distinctive approach. This paper outlines the distinctive contribution made by qualitative research with regard to the nature of the curiosity involved, the iterative research process and its treatment of data, analysis and findings. We caution against simply importing templates developed for systematic review of quantitative work, and make suggestions with regard to developing a new model for evaluating and synthesizing qualitative work. The proposed new model takes a critical look at some of the assumptions underpinning systematic review, such as the process of literature searching and selection of relevant material. Although there is potential for checklist items – such as purposive sampling, respondent validation, multiple coding, triangulation and grounded theory – to be used over-prescriptively in evaluating qualitative papers, it is argued that a more creative engagement with these concepts could yield a distinctive approach more appropriate for this type of work. Moreover, we speculate that some of the questions thereby raised might be usefully applied to consideration of established procedures for reviewing quantitative work.

Introduction
In recent years, we have witnessed a welcome move away from the previous polarization between quantitative and qualitative approaches to research. We are now in a phrase that is characterized by collaborative ventures – frequently multidisciplinary in nature – that aim to fuse quantitative and qualitative methods to provide more comprehensive understandings of the phenomena and research questions that we seek to study. Whilst there is undoubtedly scope for mixing methods and enhancing the rigour of each by engaging critically with the other (Barbour 1999), there are also dangers in under-emphasizing the important differences between the two paradigms; that which renders them distinctive. To acknowledge incompatibilities and their source is not to deny the enormous potential of mixed method designs. However, it is important that we avoid producing a homogenized and bland synthesis, which overlooks incompatibilities rather than developing solutions, compromises or simply deciding that it is most appropriate to pursue parallel strategies. There may, then, be limits to the rapprochement between quantitative and qualitative models.

The issue of systematic review of qualitative research throws many such incompatibilities into sharp focus. Whilst the drive to incorporate qualita-
tive research into the ‘evidence base’ is testimony to its growing respectability, this may prove to be a poisoned chalice. Most attempts at systematic review of qualitative research have simply borrowed and sought to impose a template designed to evaluate quantitative work, which has been modified only slightly in order to apply to qualitative papers, as Booth (2001) has also pointed out. We wish to argue here that not only is such an approach likely to miss important aspects of the contribution that qualitative research can make, but also that engaging critically with the problem of evaluating qualitative research raises some important questions for literature reviews and synthesizing of research findings more generally.

Qualitative research makes a distinctive contribution that raises particular challenges for evaluating and synthesizing such work. It is distinguished from quantitative approaches with respect to the nature of curiosity involved, the research process itself and the way in which data and findings are produced: these headings are used below to outline the challenges for evaluating and synthesizing qualitative research.

Challenges for evaluating and synthesizing qualitative research

Curiosity

Research paradigms are often viewed as a given, without recognizing the importance of the match between the individual researcher and the research approach embraced. Qualitative research asks very different questions and taps into what can usefully be viewed as a different sort of curiosity. Writing more than 40 years ago, Trow (1957, p. 33) argued: ‘cobbler think leather is the only thing. We ask research questions which are amenable to being answered using the methods with which we are familiar and in which we have developed skills’. More recently, Tricia Greenhaugh (1997), who has been largely instrumental in persuading the medical establishment to adopt qualitative methods, makes the point that individual are probably predisposed to asking what are either essentially quantitative or essentially qualitative questions. Whilst quantitative work seeks to measure and specify, qualitative researchers are likely to find that their questions may broaden out and that these are developed and built upon as the research progresses. Rather than setting out to test a clearly defined hypothesis, qualitative research excels at generating hypotheses (Mason 1996). Hence, qualitative researchers are likely to delight in having their own ideas challenged and a core activity of qualitative research involves ‘problematizing’ concepts and rendering strange the mundane or taken-for-granted (Barbour 2000).

The research process

Qualitative research has frequently been described as involving an iterative process (Marshall & Rossman 1995; Mason 1996). Thus projects are inherently flexible, with the potential to change or shift emphasis and even direction as they unfold. As the focus of most qualitative research is on issues that are salient for respondents, it is impossible, at the outset, for the researcher to determine the precise content and scope of the data to be generated. Interviewees or focus group participants may therefore play a significant role in alerting the researcher to further aspects that merit investigation. This is reflected in the sort of tools developed (such as interview schedules or focus group topic guides), which tend to be open-ended and to evolve throughout the course of studies.

Likewise, qualitative sampling strategies are not concerned with achieving representativeness (Mays & Pope 1995), but rather with reflecting the diversity (Kuzel 1992) within the groups or phenomena being studied. Purposive sampling, as described by Kuzel (1992), involves the researcher in speculating, at the outset of the study, as to the dimensions likely to influence respondents’ perspectives and experiences. Sometimes this type of sampling is referred to as ‘theoretical’ sampling, as it may draw on what is previously known, or what is ‘theorized’ with regard to what constitutes important characteristics. Such ‘theorizing’ may involve the researcher in reviewing both the existing qualitative and quantitative literature base of published studies. However, further theorizing is likely to be part of the qualitative research endeavour, even as data are generated and subjected to ongoing analysis, which may uncover unanticipated patterns related to the expression of particular
perspectives or production of accounts/explanations by respondents.

As power calculations are rendered unnecessary and because the overall numbers involved are likely to be relatively small it is possible – indeed often desirable – to augment samples by including other individuals or groups, in order to explore further the usefulness or relevance of such emerging explanatory frameworks. Glaser & Strauss (1967), in the original formulation of ‘grounded theory’, advocated returning to the field systematically to test such tentative hypotheses. The tighter time-scales and budgets of the current research climate do not generally afford us the opportunity to indulge in this luxury. However, it is possible to carry out some further interrogating of tentative hypotheses in a small way by utilizing such opportunistic theoretical sampling, thus affording enhanced analytical – in this case, comparative – purchase.

Data, analysis and findings

Recent commentators (e.g. Blaikie 1993; Seale 1999) have questioned the usefulness and accuracy of the oft-cited distinction between deductive reasoning (seen as characterizing the quantitative paradigm) and the equivalent inductive reasoning (hailed as the hallmark of the qualitative approach). Nevertheless, we would argue that arguments or explanations are built up in markedly different ways in the two broad traditions. Whereas quantitative research seeks to explicate the exact nature and strength of associations or relationships, qualitative research tends to develop an ever-widening explanation, drawing on diverse literatures and sometimes even crossing disciplinary boundaries.

Mason (1996) has made the important distinction between ‘collecting’ and ‘generating’ data, underlining the role of the qualitative researcher in engaging with respondents to produce data, using her/himself as a research tool. A similar distinction relates to the production of ‘findings’ rather than ‘results’ (Barbour 2000). The marriage between researcher and project, discussed above, is underlined by the tendency of qualitative researchers to write themselves into their accounts and presentation of findings. Rather than attempting to let data speak for themselves, the researcher is required to interpret data, building up an argument, considering alternatives and exceptions, and to provide a description of how findings were reached. Outcomes cannot thus be detached from the process of data analysis and the reasoning engaged in by the researcher. The need to render transparent this process presents one of the most important challenges for the qualitative researcher, particularly where journal word limits are strictly enforced.

Checklists

Checklists (e.g. Popay et al. 1998) have been developed to respond to the need to critically appraise qualitative papers and critical appraisal is a cornerstone of systematic reviews. Whilst the acknowledgement that there exists such a thing as bad qualitative research was long overdue, Barbour (2001) has argued elsewhere that an over-rigorous application of checklists can be counter-productive. Step-wise templates cannot hope to encompass the broad-ranging arguments and descriptions of what is essentially an iterative rather than linear (Marshall & Rossman 1995) research process. They also offer tantalizing possibilities to the researcher faced by word limits and appear to offer ‘off the peg’ descriptions. All too often, however, researchers succumb to the temptation to use such ‘technical fixes’ as ‘purposive sampling’, ‘multiple coding’, ‘respondent validation’, ‘grounded theory’ and ‘triangulation’, not just as shorthand descriptions but as appeals to credibility and rigour (Barbour 2001). Thus, such procedures may be invoked after the event – at writing-up stage – rather than having been applied in the course of the research.

Had such checklists been a feature of the history of qualitative research to date, it is likely that many of the most influential papers in this genre would have been de-selected from review exercises on the grounds of failing to fulfil all the stated criteria and some might never have been published. Although good research design and systematic analysis make for rigorous qualitative research and generally accompany the most valuable theoretical or conceptual insights, the two are not entirely interdependent. Moreover, what counts as good research design and analytical rigour is also subject to change over time, as is evidenced by the debate as to which items merit
inclusion in checklists and the degree of emphasis which should be placed upon each. A good example is provided by an examination of the brief history of ‘grounded theory’ (Glaser & Strauss 1967), which has recently enjoyed a resurgence. Although subsequent commentators (e.g. Melia 1997) have highlighted the need to develop a pragmatic version of ‘grounded theory’, rather than relying solely on the content of data to provide theoretical insights, tentative theoretical constructs can, as Glaser & Strauss (1967) argued, be derived from analysis of data generated by empirical research. Many qualitative papers, accordingly, occupy the grey area between presentation of the outcomes of empirical research and abstract theoretical discussions, and either make scant reference to the content of data sets or are highly selective with regard to the parts or concepts explored. Insofar as the researcher provides a description of the process of analysis and developing of analytical categories, it is possible to determine how rigorous or systematic their work has been. In many cases, however, papers are also likely to be judged, ultimately, on more subjective criteria, such as the persuasiveness of the writer (Seale 1999) or the perceived quality of the theorists cited, interpreted and critiqued.

Aggregating or comparing?

It has been argued that qualitative research projects involve an iterative process. This means that they are likely to start out with general questions, which are progressively refined in the course of the study. Rather than attempting to answer identical questions to those posed by other studies in the same topic area, there is a sense in which each qualitative study remains essentially self-referring, in that the research question is progressively honed and tailored specifically to address the data that have been generated. Although they may refer in papers to similarities and differences in findings between their own and others’ studies, before moving on to ask further questions of their own data, qualitative researchers do not generally set out specifically to refute or corroborate earlier work. Given the stress on context that characterizes the qualitative research endeavour, not only do concepts such as ‘refutation’ and ‘corroboration’ not make a great deal of sense; neither are they the stuff of which qualitative research curiosity is made. This has important consequences for the way in which the qualitative literature base is built up. Rather than being strictly cumulative, with each new study engaging with and adding to the evidence base, it is likely that much more casual and partial reference will be made to other studies. This can, of course, be taken too far, and can lead to researchers advancing similar explanations using different labels, rather than seeking to explore the reasons for similarities and differences in findings (Morse 2000). Such unacknowledged and unexplored duplication of effort also makes for hard work for the potential reviewer. However, the iterative – or inclusive – nature of qualitative research, which tends to bring different sources to bear as findings are described and as tentative explanatory frameworks are presented and interrogated, poses significant challenges for existing approaches to systematic review, which rely on the development of exclusion criteria.

Literature searching

There are also significant challenges with regard to literature searching, which forms an integral part of the systematic review process. The new-found acceptability of qualitative research means that it is being published in an ever-widening range of journals, including mainstream medical publications. However, given the transferability of qualitative research skills and their capacity both to address a wide variety of substantive topic areas and even to cross disciplinary boundaries, qualitative research has always found its way into a bewilderingly wide array of journals. Whilst it is relatively easy to search for qualitative papers in mainstream journals, the available electronic databases do not even cover all of the more popular qualitative journals. Although there have been attempts to develop search strategies to improve the scope of searches (e.g. Grant 2001) these have failed to address the issue of the adequacy of the coverage of databases such as Medline. All too frequently potential reviewers rely exclusively on a Medline search, ignoring the limitations of its essentially American – and clinical – focus. If a review is to aspire to being systematic or comprehensive it is important to carry out searches on
other databases, such as Excerpta Medica (EmBase) (which provides greater coverage of European journals; also of pharmacology and psychiatry). Because much qualitative research is published in nursing and social science journals, it is also essential to make use of databases such as the Cumulative Index to Nursing and Allied Health (CINAHL), the psychological literature database (PsycINFO, previously PsychLit) and Applied Social Sciences Index and Abstracts (ASSIA). The National Research Register (NRR) provides information on ongoing research projects and the Health Management Information Consortium (HMIC) is an important source for grey literature (such as reports). The input of professional librarians/information specialists is crucial, but often overlooked.

Rather than relying on freetext or researchers’ own terms, recourse should be made to the Thesaurus, which explains how terms (i.e. keywords) are used in the context of each database (Lowe & Barnett 1994). In addition to giving definitions for terms, database Thesauri are also designed to take account of grammatical variants, and both anglicized and American spellings. However, they are less sensitive to the nuances of terminology as applied to a wide range of disciplines. This is an important shortcoming, given the multidisciplinary pedigree of the qualitative research literature. Databases operate with a restricted number of keywords, which usually describe the more established general or broad-based topic areas and methodologies. Comprehensiveness therefore cannot be achieved. Although Thesauri provide definitions of the terms used to index material, these are unlikely to correspond with researchers’ specific interests and focus. Furthermore, there can be no definitive ‘search strategy protocol’ for any given research topic. As each bibliographical database has a different focus, keywords need to be modified accordingly, depending on whether the database is subject-specific (e.g. EmBase or CINAHL) or whether it aspires to covering a broad range of related disciplines (e.g. Social Sciences Citations Index). Although there is a facility for searching between databases within OVID (which hosts several ‘files’, including EmBase and Medline) search strategies do not translate in practice – apart from those situations where a keyword is a precise clinical term.

**Developing a new model**

Often perceived as the first step in putting together a research proposal, literature searching is not a discrete activity to be performed at the outset of a research project. Given the iterative nature of qualitative research and the potential for the focus of the research to shift, it may be necessary to expand the search strategy beyond the more general requirement of repeating and updating searches. Although most discussion of identifying the relevant qualitative research literature has, to date, revolved around the conduct of systematic reviews and contributing to the evidence base, there are, in effect, several reasons for seeking to identify and use a body of qualitative research papers. The researcher, as reviewer, may be seeking findings relevant to the substantive topic area; may be looking for discussions about using complementary or divergent methods; or may even be searching for theoretical literature that may illuminate the processes or phenomena under study. Thus, systematic review is not the only reason for undertaking a comprehensive literature search, and an approach that attempts to be inclusive rather than to develop strict exclusion criteria, might yield more useful material.

Rather than seeking to aggregate findings from several studies or attempting to establish the extent to which other studies corroborate our own, qualitative researchers may be particularly interested in identifying studies with divergent findings. These may be used to further interrogate their own tentative explanations, with these discrepant findings used to provide further information about the possible limitations of emergent explanations. This can be achieved through paying attention to the contextual details of other pieces of research. It has frequently been pointed out that the hallmark of qualitative research is its capacity to provide context, but we do not always provide sufficient detail to enable our work to be evaluated. Accordingly, several commentators (e.g. Hoddinott & Pill 1997) have called for an ever-expanding list of background information to accompany papers submitted to journals, in order to facilitate the review process. However, it is debatable to what extent it is possible to anticipate all information likely to be of relevance to the reader or reviewer. As it is common practice for researchers
carrying out meta-analysis of quantitative studies to contact the original investigators (in order to glean additional information about assumptions and procedures employed) it is curious that this should not automatically be the first port of call for those engaged in a similar endeavour with regard to qualitative research. Indeed, rather than the somewhat adversarial model, which tends to be adopted in quantitative meta-analysis, a collaborative model may be the most appropriate with respect to synthesizing qualitative work. This would mirror approaches to team-working and it is likely that a prospective rather than a retrospective model would yield the most fruitful results.

In addition, qualitative researchers may wish to identify publications of theoretical significance, which may derive from study of completely different substantive topics and which may involve interrogating discipline-specific databases. Many of the available examples of cross-fertilization in terms of theoretical insights that transcend substantive topic areas are the result of serendipitous face-to-face collaboration between researchers involved in parallel research projects (e.g. Exley & Letherby 2001, who used two separate studies, concerned, respectively, with the experiences of terminal cancer patients and infertile couples, to theorize about disruption to the lifecourse and issues of identity). However, thoughtful literature searching could potentially identify many more such possibilities. In this respect, one of the greatest challenges lies in developing creative search strategies in order to facilitate the identification of relevant theoretical leads. A useful starting point is provided by databases, such as CINAHL and Social Science Citations Index, which cover citations in published articles and which also encompass monographs and edited collections, where theoretical contributions – as opposed to reports of research findings – are frequently to be found.

Although an over-proscriptive approach to checklists can be counter-productive, we wish to argue that commonly listed checklist items can, nevertheless, provide a useful starting point for developing an alternative – iterative/comparative rather than aggregative – model for synthesizing qualitative research. This would rely on the core concept of the constant comparative method (Green 1998) and would seek to build up more comprehensive explanations by taking account of and seeking to incorporate divergent findings. Thus, literature searching could embrace a variant of purposive sampling, with searches based on the key concepts of relevance to researchers reviewers rather than those pertaining to the substantive topic of study and with particular attention being paid to discrepant findings or parallel study populations of theoretical rather than substantive relevance. Gaining satisfactory analytical purchase on qualitative data has always involved an element of serendipity, via personal contacts, chance reading or browsing in the library. It is important that in our rush to comply with the imperatives of systematic review and synthesis of research, we do not lose sight of the fertile ground traditionally visited and capitalized upon in wide-ranging scholarship.

There has recently been considerable debate regarding the role of respondent validation. Whilst it is questionable to what extent this is a necessary or worthwhile endeavour with regard to corroborating interpretation of our data (Bloor 1997; Barbour 2001), a broader – and ultimately bolder – definition of relevant literature holds considerably more potential in terms of informing and enriching our understanding of qualitative data. Shrinking time-scales for research and the scramble to climb aboard the evidence bandwagon have narrowed the scope of the literature that is brought to bear in analyzing and evaluating qualitative research. We could reap enormous benefits though paying attention (as did scholars such as Goffman – see Strong’s 1983 appreciation) to a much more diverse literature, including first-person accounts, novels, newspaper coverage, etc. Although such material might not make a significant contribution to the evidence base, it may be invaluable with regard to encounters with individual patients and may be appropriate for producing patient information. A good example is provided by the DIPEx project (Herxheimer et al. 2000; Ziebland 2000), hosted by the University of Oxford, which links a video and audio database of patients’ experiences with evidence-based information about the illness and treatment options and a range of other resources.

Provided that we can successfully develop search strategies to allow us to identify parallel data sets, another checklist item – multiple coding – offers the opportunity for carrying out secondary analysis of
others’ data sets with the object of engaging in dialogues – actual or virtual – in order to develop/refine our explanations, rather than for the more mundane purpose of validation (as specified in many checklists). Although multidisciplinary research teams undoubtedly raise some challenges, they allow data to be subjected to a range of disciplinary gazes (Barry et al. 1999), drawing on a wider store of theoretical frameworks and insights. The ensuing exchanges and the resulting refinement of interpretations and explanations provide for a much more comprehensive and conceptually productive review than do traditional approaches based on triangulation, with its restrictive focus on internal validation. Grounded theory is a checklist item that is invoked more often than it is practised (Bryman & Burgess 1994) and current time-scales mitigate against one of its central tenets, that of returning to the field to further test emergent theories or hypotheses (Glaser & Strauss 1967). However, there is considerable potential in identifying and interrogating complementary literature in this way/to this end, allowing it to stand in for further fieldwork. Such approaches, however, require new and more collaborative models of working with information specialists to ensure that qualitative research harnesses such expertise rather than merely being subsumed to the discipline and strictures of systematic review, secondary analysis and meta-analysis.

Conclusion

The ultimate fallacy may lie in viewing critical appraisal, secondary analysis and meta-analysis as separate activities. Given the iterative nature of the qualitative research enterprise, does it not make more sense to view these as forming part of a broader seamless process, built on the concept of the constant comparative method? Qualitative research is frequently associated with representing the voices of those otherwise disenfranchised, and generally employs a bottom-up approach. Rather than seeking to import and impose templates and methods devised for another purpose, qualitative researchers and reviewers should look, instead, for inspiration from their own modes of working and collaborating and seek to incorporate these, forging new and creative solutions to perennial problems, rather than hoping that these will simply disappear in the face of application of pre-existing sets of procedures. A critical approach to synthesis and evaluation of qualitative research may also yield dividends with regard to interrogating and further developing established approaches to evaluating and synthesizing quantitative work.

References


Booth A. (2001) Cochrane or cock-Eyed? How should we conduct systematic reviews of qualitative research? Paper presented at the Qualitative Evidence Based Practice Conference, held at the University of Coventry, 14–16 May.


