What have sampling and data collection got to do with good qualitative research?

Abstract

Objective: To highlight the importance of sampling and data collection processes in qualitative interview studies, and to discuss the contribution of these processes to determining the strength of the evidence generated and thereby to decisions for public health practice and policy.

Approach: This discussion is informed by a hierarchy-of-evidence-for-practice model. The paper provides succinct guidelines for key sampling and data collection considerations in qualitative research involving interview studies. The importance of allowing time for immersion in a given community to become familiar with the context and population is discussed, as well as the practical constraints that sometimes operate against this stage. The role of theory in guiding sample selection is discussed both in terms of identifying likely sources of rich data and in understanding the issues emerging from the data. It is noted that sampling further assists in confirming the developing evidence and also illuminates data that does not seem to fit. The importance of reporting sampling and data collection processes is highlighted clearly to enable others to assess both the strength of the evidence and the broader applications of the findings.

Conclusion: Sampling and data collection processes are critical to determining the quality of a study and the generalisability of the findings. We argue that these processes should operate within the parameters of the research goal, be guided by emerging theoretical considerations, cover a range of relevant participant perspectives, and be clearly outlined in research reports with an explanation of any research limitations.

Keywords: Qualitative research; methods; evidence; interviews.

Lisa Gibbs

The McCaughey Centre, VicHealth Centre for Promotion of Mental Health and Community Wellbeing, School of Population Health, University of Melbourne, Victoria

Michelle Kealy

Mother & Child Health Research, School of Public Health, La Trobe University, Victoria

Karen Willis

School of Sociology and Social Work, University of Tasmania

Julie Green

Centre for Community Child Health, Murdoch Childrens Research Institute, and Department of Paediatrics and Faculty of Education, University of Melbourne, Victoria

Nicky Welch

Centre for Physical Activity and Nutrition Research, Deakin University, Victoria

Jeanne Daly

Mother & Child Health Research, School of Public Health, La Trobe University, Victoria

P ublic health research using qualitative methods produces studies that can range from an exploratory study with modest implications for practice to well-developed, generalisable studies. The contribution that a study can make to public health practice and policy rests on several core features of sound qualitative research. In common with other empirical studies, qualitative research starts by justifying the research problem by reference to the literature. Qualitative research then defines a theoretical framework for the study, identifying the theoretical concepts that are relevant and will be employed in the study.[1] The next step is to collect data according to a sampling plan, following which there is the analysis of data and reporting of research findings.[2] In this paper, our focus is on sampling and data collection.

There are inconsistencies and gaps in the literature regarding appropriate appraisal of qualitative research.[3] We propose that sampling and data collection are critical to determining the quality of a study. We use the underlying model of a hierarchy of evidence-for-practice[3] to discuss the role of sampling and data collection in determining the strength of the evidence for decisions made in practice or policy settings. This has particular relevance as a guide for researchers seeking publication and reviewers of submitted articles, given recent concerns about the quality of qualitative papers being submitted for publication.[4] One of the biggest problems noted was the lack of information provided about sampling, providing little opportunity

to assess the generalisability of the findings. In our discussion of sampling and data collection processes, we start with the studies that generate the least reliable evidence-for-practice. For the purposes of this paper, we confine ourselves to interview studies, recorded verbatim on audiotape and then transcribed, as the most commonly used method of data collection.

The hierarchy of evidence model provides four types of study that all produce differing levels of evidence for health policy and practice: the single case study, descriptive study, conceptual study and generalisable study. As in quantitative research, the single case study is limited by a very small sample but it can provide interesting and important information about a setting. Descriptive studies typically provide an overview of a setting, describing a range of experiences or activities without exploring these differences further. Both case studies and descriptive studies may provide important new information about a problem, often indicating the need for further research, and may therefore be worth publishing if the limitations of the studies are clearly acknowledged. They provide a weak basis for practice or policy decisions. In contrast, high-quality conceptual studies and generalisable studies both draw on a theoretical framework for sampling and analysis, with generalisable studies providing a more comprehensive analysis of differences in experience.[3]

Starting the project

In most public health research projects, funding is provided for the analysis of a specific problem, justified as important and unresolved in the literature. Qualitative health researchers also review the theoretical literature for relevant concepts and theories that form a theoretical framework for the study. At this stage, quantitative and qualitative research processes diverge. Quantitative public health researchers enter the field with a set of variables that are measured using a validated instrument, usually with a sample statistically representative of the community or population of interest. Engagement with the community may be seen as a source of bias, undermining the objectivity of the study. Qualitative researchers identify theoretical concepts likely to be important to the study and identify a setting where it is likely that data relevant to the problem will be readily available. In many high-quality qualitative research studies, researchers then immerse themselves in the field before starting data collection. The quality of a qualitative research project may well rest on the extent of the understanding of the problem gained at this early stage.

Entering the field involves becoming acquainted with the research setting and may commonly involve the ethnographic processes used by early anthropologists to be accepted within a community so that the researcher becomes a trusted and all-but-invisible observer. A good long period spent in entering the field allows researchers to test preconceptions about the research problem and to identify relevant sources of information. It is common to amplify interview data with researcher notes about the setting and with other relevant material collected on site.

Entering a research field is always demanding. In some research settings, it is fraught with difficulties. In a much-quoted example, Booth and Booth[5] studied parents who had learning difficulties, a vulnerable group whose often inarticulate views are easily overlooked in research findings. The researchers were particularly careful to let potential research participants take the initiative in participating by introducing the study through a trusted professional worker supporting the family. This was followed by a long process of establishing trust and building rapport, collecting data through a series of interviews. Leaving the field and withdrawing from personal contact was another slow process to ensure the field relationship was not exploitative. In recognition that people with learning difficulties often have limited social networks, the withdrawal occurred at a pace that was comfortable for each family. The emphasis of the researchers was on sensitive relationships: "the validity of the data is the stuff of the relationship between the interviewer and the informant" (p. 421).

Many researchers reporting well-developed studies record that they spent long periods of time entering the field and even more time on retaining a precarious presence. Pyett[6] recognised that sex workers were working outside the margins of society and had every reason to distrust researchers. Her solution was to embark on a long and often challenging task of engaging with sex workers in a collaborative manner so that they participated actively in the research, including data collection, with one participant included as a co-author on some of the papers.

An intensive engagement in the field of research provides an opportunity to identify the people who can provide the best information about the research problem. It also informs researchers about the people who are most likely to give an opposing view, providing researchers with an indication of how to diversify their sample to analyse differences in experience. It is critical to the generalisability of the research findings that the views held by a particular group are fully understood. Data collection can be extended to establish whether a variety of views is present, perhaps in the same group, perhaps in another group, and fully exploring these, too.

The problem of immersing oneself in a research field is more difficult if more than one setting needs to be studied to allow different perspectives to emerge. Galloway and colleagues[7] report a study of iron supplementation, conducted under the auspices of the MotherCare Project, a United States Agency for International Development (USAID)-funded maternal and child health project. The study was conducted as a cross-country comparison of eight less-industrialised countries in which the aid project operated. The authorship represented researchers immersed in each of the different settings. Any research team trying to establish a research presence *de novo* in this range of research settings would face a formidable task.

Because of time, resource and other practical constraints, immersion in the field occurs to varying degrees, but it should be recognised that compromising this step may limit the quality of the evidence generated. A testing circumstance occurs when access to the research setting is restricted. The research goal of Tuckett's study of nursing homes[8] was to "explore the meaning

of truth telling within the care provider-aged resident dyad" (p. 48). Tuckett based the study in a range of nursing homes but he was familiar with only one setting as a result of his professional clinical work. He encountered the problem of 'gatekeeping' in another setting where he had to rely on nursing staff to identify eligible resident participants. One institution restricted access to people involved in the care relationship and one nursing home failed to complete its involvement in the study. These problems of entry to the research field seriously compromise the design of the study and set limits to the validity of its conclusions.

Sample selection and data collection

After time spent in the field, researchers should have a good understanding of the most appropriate way of drawing a sample, and any other sources of relevant information, that will yield rich data relevant to the research problem.

The researcher's conceptual framework for the study and existing literature on the topic guide the initial sample selection, but this sampling strategy is constantly rethought. Data analysis starts with the first data collected and the results of this initial analysis is fed back into the sampling process. This guided sample selection is a strength of qualitative sampling and is aimed at steadily intensifying understanding of the research problem. Researchers search for disconfirming cases to validate their analysis and they diversify the sample to address new theoretical concepts that emerge during data analysis. When the data being collected has become repetitive and no new issues are emerging, data saturation is achieved and it is considered appropriate to cease collecting data. Where a research problem is proving difficult to understand, there is the opportunity to diversify methods of data collection to substantiate interview findings. These methods could include observations of research settings, collecting visual records such as videos and photographs, and material from local media or archives.

It is in this process of obtaining a properly diversified sample, with no analytic cul de sacs, that careful entry into the field pays off. A well-developed study, with generalisable results, explores both the theoretical concepts from the original theoretical framework and new concepts and theories that emerge during data collection. An important purpose of diversification of the sample is to search for cases that do not fit the developing conceptual understanding of the data and to explore the nature and extent of these differences, thus contributing to the generalisability of the findings. Sometimes, it is possible to identify a clear reason for an exceptional case that does not undermine the plausibility of the theoretical analysis. Otherwise, failing to investigate deviant or outlier findings limits the quality of the study and the application of the findings.

In reality, many qualitative studies employ simplified sampling strategies and this has an important bearing on the level of generalisability of the results. Some studies use an opportunistic sample that is justified by showing that the sample is diversified according to demographic variables, but these demographic variables are not used as explanatory factors in the analysis. It is this link between sampling procedures and analysis that establishes the extent to which the sample is representative of other social groups. When this step is missing, studies can claim only a modest contribution to knowledge and it is a helpful guide to the application of the findings when these limitations are acknowledged.

Sampling and data collection and the strength of the evidence

Here we use the hierarchy of evidence model to focus on the way in which sampling and data collection determine the generalisability of the study and the generation of evidence for practice.

As in quantitative research, the **single case study** produces limited evidence but it may provide valuable insight about a setting. Some single case studies emerge during immersion in the field but identification may also come about by chance. These studies are hypothesis-generating with further research required to test the ideas generated, or they may represent one aspect of the early stages in an intensive community study.

An example. Fleming undertook a single case study to understand the experience of abdominal hysterectomy as experienced by one woman and her "significant others".[9] The strength of this study is in the generation of in-depth knowledge of the experience, unfettered by existing assumptions and theoretical frameworks. Although Fleming located the study in the existing literature on women's experience of hysterectomy, we are left unsure of the extent to which other women might have the same experience of hysterectomy.

Descriptive studies use methods closely resembling survey designs. Researchers enter the field with a predetermined set of variables for selecting a sample and there is thus little need for immersion in the field. Diversity is assumed by selecting from different population groups (often defined by demography) or different settings, but these factors may not feature in the analysis and thus possess little explanatory value. There is no attempt to return to the field to explore further any issues that arise during analysis, which is largely restricted to listing a range of experiences. Nevertheless, these studies may indicate the need for further research to explore interesting findings.

An example. Scott and colleagues examined the impact of the introduction of electronic medical records on four primary health care teams and one hospital in regional Hawaii.[10] They undertook 26 interviews with a range of staff across the various workplaces and identified a 'climate of conflict' associated with the implementation process. The authors describe their research as providing only a 'snapshot' view, while highlighting issues that are likely to arise in a workplace change situation

In **conceptual studies,** the initial sample is diversified according to concepts derived from the theoretical framework for the study. Analysis draws on these concepts to show how different groups differ in the way in which they experience a problem and to explain why this is so. These may be substantial studies, but sampling

strategies are limited. The sample may be limited to a small number of settings when perhaps other settings could produce contrasting data; alternatively, when differences emerge during analysis, there is no further sampling to explore the diversity of participant views.

An example. Gabbay and le May set out to study why clinicians often ignore the best scientific advice about interventions in their clinical decision making.[11] The theoretical framework drew on Michael Polanyi's[12] analysis of tacit knowledge to show that clinicians do take account of scientific research evidence but only in the context of other knowledge that comes from experience. Their data were drawn primarily from one clinical practice where they interviewed a range of clinicians and collected a wide range of other data including documentary material. They observed interactions during meetings and informal interactions in the practice. A second practice was used to check findings. They caution that their conclusions may not apply to practices with substantially different organisational structures.

Generalisable studies actively draw on a well-developed theoretical framework and sample for key theoretical concepts, diversify their sample to derive an explanatory model that is relevant to a broader range of settings, and can be implemented in practice with a high degree of confidence. An important purpose of diversification of the sample is to search for disconfirming cases, i.e. those participants who do not fit the developing conceptual understanding of the data, and to explore the nature and extent of these differences. This contributes to the generalisability of the findings by providing comprehensive explanation of the research issue. Given that these are intensive and extensive studies, we report two examples below which show many, if not all, of the features of high-quality qualitative studies.

Example 1. Over a number of years, Charmaz studied the impact of chronic illness on personal identity, the relation between body and self, meanings of loss and illness, and identity goals.[13] She conducted 115 intensive interviews exploring the body in illness; 16 of the study participants were followed longitudinally. An additional 25 highly focused interviews were then conducted with 12 of the original respondents from the longitudinal group and 13 new participants to pursue an understanding of issues about body and self that had emerged during the initial interviews. Further published personal accounts of chronic illness and disability were also collected and examined for statements about the body to explore other perspectives and to diversify the sample. Respondent characteristics such as gender, age, socio-economic status, marital status and diagnostic status were reported and each of these variables was then interwoven into the analysis of the results. For example, Charmaz noted the greater importance to women of appearance issues and yet also noted that compared with men, women showed greater resilience in dealing with the experience of illness. Her conclusions have implications for all chronically ill people: "By regaining control and coping with bodily changes, these people learn to live with their illnesses … through struggle and surrender, ill people paradoxically grow more resolute in self as they adapt to impairment" (p. 675).

Example 2. Warr was interested in the concept of social capital and in the way in which social networks operate in poor and marginalised settings, what she calls "discredited" neighbourhoods.[14] Her theoretical framework drew on the concept of stigma. She started spending time in a neighbourhood adjacent to an industrial zone, with a high proportion of public housing and a community of poor, largely unemployed residents. For seven months she attended a parents' group that met weekly and conducted community-based participant observation. At the same time, she organised consultative workshops, interviewed service providers and community workers and extended her study to include a slightly different neighbouring suburb because of mobility between the suburbs. The problem, however, was to gain access to the full range of resident experience in a setting where social stigma translates into shrinking social networks and the most vulnerable people do not attend community settings. She used a snowball sampling technique: initial contact through neighbourhood organisations provided her with her first interviews and she built on their friendship networks to extend and diversify her sample. Analysis involved mapping social networks and understanding the range of social practices encountered. The result is a comprehensive account of the experience of living in these neighbourhoods. It is the way in which she grounded her analysis in the concept of stigma that allows us to extend her findings to other discredited neighbourhoods.

Conclusion

The method used for data collection and the nature of the data collected depend on the research goal, the circumstances in which the study is conducted, and participant and community sensitivities. We accept that the limit to data collection is often practical and situational: research funds are exhausted, time is limited, ethics committees impose restrictions, or research access is compromised. For example, Tuckett's study of nursing home residents referred to above[7] had to exclude residents who had dementia or were acutely ill, too frail or too emotionally disturbed for participation. One resident died during the study and so was unavailable for follow-up interviews.

While acknowledging such practical difficulties, we suggest that there is an ideal process that allows the researcher to obtain the best possible sample and data. The complexities involved are not captured by the terms that are used in some research reports. Warr's study of stigmatised communities used snowball sampling, but there is nothing opportunistic or simple about her sampling strategy.[13] Gabbay and le May used a variety of data to substantiate interview findings but it seems inadequate to describe their extensive sets of data collection as "triangulation".[10] A claim to data saturation is only appropriate when it can be demonstrated that researchers have gained a full understanding of the variety of experiences relevant to their research problem. Constraints on achieving saturation are a legitimate reality of research and are best acknowledged by describing the limitations of the study.

In some academic journals there seems to be the expectation

that the methodological section of qualitative reports will include demographic information about the participants. Ethically, researchers should only be collecting information that is needed for the research. If demographic information is being collected, demographics should have a role in analysis conducted. Demographic diversity in itself does not indicate the diversity of sample that indicates a well-conducted study.

There are many excellent research papers, reports and books available to the qualitative researcher that address quality in the methodological processes. However, the complexity and detail contained in these documents can reduce their accessibility. This paper has highlighted some core considerations in the design, conduct and reporting of sampling and data collection in a condensed, readily accessible format. In deference to the proposed hierarchy of evidence-for-practice model for qualitative research, public health researchers are encouraged to employ a sampling and data collection approach that operates within the parameters of the research goal, is guided by emerging theoretical considerations, covers a range of relevant participant perspectives, and is clearly outlined in research reports with an explanation of any research limitations.

References

1. Willis K, Daly J, Kealy M, Small R, et al. The essential role of social theory in qualitative public health research. *Aust N Z J Public Health.* 2007;31(5):438-43.
2. Green J, Willis K, Hughes E, Small R, et al. Generating best evidence from qualitative research: the role of data analysis. *Aust N Z J Public Health.* 2007;31(6):545-50.
3. Daly J, Willis K, Small R, Green J, et al. Assessing qualitative research in health – a hierarchy of evidence-for-practice. *J Clin Epidemiol.* 2007;60(1):43-9.
4. Daly J, Lumley J. Dilemmas in publishing qualitative public health research. *Aust N Z J Public Health.* 2007;31(3):203-4.
5. Booth T, Booth W. The use of depth interviewing with vulnerable subjects: lessons from a research study of parents with learning difficulties. *Soc Sci Med.* 1994;39(3):415-24.
6. Pyett P. Researching with sex workers: a privilege and a challenge. In: Elias JE, Bullough VL, Elias V, Brewer G, editors. *Prostitution: On Whores, Hustlers, and Johns.* New York (NY): Prometheus Books; 1998. p. 368-75.
7. Galloway R, Dusch E, Elder L, Achadi E, et al. Women's perceptions of iron deficiency and anemia prevention and control in eight developing countries. *Soc Sci Med.* 2002;55(4):529-44.
8. Tuckett A. Qualitative research sampling: the very real complexities. *Nurse Res.* 2004;12(1):47-61.
9. Fleming V. Hysterectomy: a case study of one woman's experience. *J Adv Nurs.* 2003;44(6):575-82.
10. Scott JT, Rundall TG, Vogt TM, Hsu J. Kaiser Permanente's experience of implementing an electronic medical record: a qualitative study. *Br Med J.* 2005;331:1313-16.
11. Gabbay J, le May A. Evidence based guidelines or collectively construed 'mindlines'? Ethnographic study of knowledge management in primary care. *Br Med J.* 2004;329:1013-16.
12. Polanyi M. *The Tacit Dimension.* London (UK): Routledge; 1967.
13. Charmaz K. The body, identity and self: adapting to impairment. *The Sociological Quarterly.* 1995;36(4):657-80.
14. Warr D. Social networks in a 'discredited' neighbourhood. *Journal of Sociology.* 2005;41(3):285-308.