Part I: Qualitative research sampling - the very real complexities

A G Tuckett, RN, MA, PhD
School of Nursing, The University of Queensland, 11 Salisbury Road, Ipswich Queensland 4102, Australia

Abstract

This article discusses the complexities of qualitative research sampling. It analyses a research experience, the rationales for and limitations of qualitative research sampling. It examines the reality of establishing and maintaining a purposeful/theoretical sample and how data saturation symbiotically interacted with constant comparison to guide sampling. Additionally, sample limitations are countered. This methods paper is aimed at novice and experienced researchers in nursing interested in the practical reality of research, who are also mindful of the necessity for rigour.

Key words: Sampling, sample frame, purposeful/theoretical sample, constant comparison, bias

Introduction

This article discusses a research experience, in particular the rationales for and limitations of qualitative research sampling. This paper draws together a range of issues and complexities associated with sampling. Sampling is a core concern determining the ongoing success of a research project. Consequently, it is an issue requiring continual examination as practiced. A second paper will discuss maintaining rigour – the research strategies and operational techniques that were employed in concert with the sampling experience described here. In order for the reader to understand the researcher’s experience, choices, and theoretical position, the following section briefly details the research undertaken.
The study

The sample framing discussed in this article is based on recent qualitative research that aimed to explore the meaning of truth-telling within the care provider-aged resident dyad in high care (nursing home) aged care. The organisation in which the research was conducted represents the largest privately operated provider of aged care nursing in south-east Queensland, Australia. At the time of commencement of study, the organisation operated twelve high care (nursing home) aged care facilities representing a total of 1124 residents (beds). Research participants included 19 residents, 23 personal care assistants (PCA), 15 Registered Nurses (RN) and a single physiotherapist (PT) representing five of the twelve nursing homes. The research process comprised a number of phases (see Figure 1.0, below).

Grounded within the epistemology of social constructionism (Berger & Luckman 1966; Crotty 1998) and the theoretical stance of symbolic interactionism (Blumer 1969; Mead 1934), research data were collected through group discussion, personal journals, follow-up in-depth interviews, and researcher field notes (Tuckett & Stewart 2004; Tuckett & Stewart 2004a). Thematic analysis of data relied on practices within grounded theory in an attempt to comprehend participants’ truth-telling experiences and understanding (Boyatzis 1998; Strauss & Corbin 1998).

Discussion

Qualitative research sampling

There are ‘no hard and fast rules about numbers (however)…(q)ualitative research in old age and aging has used experiential cell sample sizes of from 10 to 100, with clustering around 50’ (Rubinstein 1994: 80). Others writing in this area suggest ‘12-20 (data sources)...when looking for disconfirming evidence or trying to achieve maximum variation’ (Baum 2002: 176).

Whilst there are no closely defined rules for sample size (Baum 2002; Patton 1990), sampling in qualitative research usually relies on small numbers with the aim of studying in depth and detail (Miles & Huberman 1994; Patton 1990). Seeking a richness of data
about a particular phenomenon, the sample is derived purposefully rather than randomly (Reed et al. 1996; Mays & Pope 1995; Ezzy 2002).

Additionally, for qualitative sampling, criteria typically define the process as

- embodied within a reasonably flexible research design, in which sampling criteria may change as the study unfolds,
- participants are sought serially (Higginbotham et al. 2001) – that is, depending on who and what has come before so that ongoing sampling supports the emerging theorising (ideas about ideas),
- sampling continues until the researcher recognises no new data were forthcoming – a point of data or information redundancy (Lincoln & Guba 1985), an ideal dependent upon some effort to seek out disconfirming or ‘negative’ cases (Baum 2002; Miles & Huberman 1994; Reed et al. 1996; Kuzel 1992).

This point of data or information redundancy is comparable to data saturation. That is, sampling continues until no new information is forthcoming or nothing new is heard in the case of interviewing (Patton 2002; Ezzy 2002; Higginbotham et al. 2001). The point of data saturation is contingent upon concurrent data analysis and data collection (and is discussed later).

The next section describes in detail the researcher’s sampling process guided by the above criteria.

Sample framing

![Sample framing diagram](image-url)
Sampling decisions were made and the sample framed (see Figure 2 above) according to the purpose of the study. Initially, the organisation and the research participants were selected according to the research aim and objectives (Ezzy 2002; Reed et al. 1996). Equally, the sample decisions were guided by reference to the research interpretive framework and realistically by practicalities and logistics (Miles & Huberman 1994).

**Selection criteria: getting started**

A total of five high care (nursing home) organisations were initially invited to participate in the research. These were investigated and according to the sample frame criteria a single, multi-site high care (nursing home) organisation was eventually selected (see Figure 2). As it is important to be mindful of the starting point for data collection, a decision had to be made about where to begin amongst the twelve nursing homes (Reed et al. 1996).

At the time, the decision was made based on ‘practicality and logistics’. Practically the researcher was familiar with the first nursing home (A) through his professional clinical work and the site was geographically accessible (Clavarino & Janda 2001). Logistically, this nursing home had been successfully audited and accredited by the Aged Care Standards and Accreditation Agency – allowing it to receive Commonwealth Government funding for the next three year period (Commonwealth Department of Health and Aging 1998). Successful auditing and accreditation meant that the potential research participants at the nursing home(s) were less distracted and more willing to engage with the researcher.

**Failed Negotiation of Entrée**

In the period from February 2000 until the commencement of June 2000, the negotiation of entrée at one nursing home laboured. The researcher made five nursing home visits, repeated the personal carer information sessions, and re-scheduled the resident information session. Over a ten day working period in May 2000, the researcher was informed that the interested research participant lists could not be located, was put ‘on hold’, asked to ring back but telephone calls were not returned. At this site, the researcher acknowledged that a significant ‘...ouch!factor’ or distracter (Alty & Rodham 1998: 275) was the Manager’s overriding responsibility for preparing the facility for Accreditation (CDHA 2002, August; CDHA 1998). The researcher gave up. (The nursing home in which there was the failed negotiation of entrée is not shown in Table 1).

**Selection criteria: participants**

Primarily, care providers (personal care assistants and Registered Nurses) were sought who had been working within a given nursing home for at least six months. In addition, care providers were sought who would be available for a follow-up interview within a four to six month period after the initial group discussion. No other criteria were enforced for initial care provider sampling. Furthermore, other care providers were not excluded so as not to stifle any chance capture of a multiplicity of perspectives.

Consideration was given to the clinical status of the elderly residents. Relying on the sample frame, the researcher knew that the selected organisation provided care for a substantial number of chronically ill, long term, aged residents. Consequently, it was
from this group of residents that research participants were sought, rather than residents that were acutely or terminally ill.

Acutely ill residents were excluded to eliminate any perceived added stress to them and to minimise the chance loss of participants due to death prior to the study’s completion. Also excluded were the grossly hearing or speech impaired, those not proficient in English, short stay (respite) residents, or those considered by the Manager physically incapable or too emotionally disturbed to participate. In brief, those residents included were those with adequate cognitive functioning (Clavarino & Janda 2001) or, as described by one Manager (C-RN1), residents having ‘slight cognitive impairment, but (who are) capable of making sound decisions for themselves and (not those) residents who are demented’ (researcher’s field journal notes).

**Group discussion and homogeneity**

The Administrator and/or the Manager in consultation with the Registered Nurses (team leaders) in each participating nursing home identified residents able to participate in open discussion (Yates et al. 1995) about the meaning and understanding of truth-telling. Their identification of potential research participants was framed by both the selection criteria and the research purpose. As a consequence of this involvement, some interested residents at one nursing home (D) were deemed by the Administrator as being ‘unable to participate in a group or write in a journal’ (researcher’s field journal notes).

Similarly, the Manager-Clinical Care Services at another nursing home (A) offered to ‘speak to a few key (personal) carers to try to get some interest (adding) the type of resident you need, the numbers here would be quite small’ (researcher’s field journal notes).

Furthermore, the nursing home care staff (PCA and RN) acted as “go-betweens” (Groger & Mayberry 1999), obtaining the residents’ agreement to participate in the information sessions. For care providers, as with the residents, the identification of potential research participants was framed by both the selection criteria and the research purpose. After the resident and staff research information session (part of the Negotiation of Entrée) at each nursing home, a sample for each group discussion was generated. However, not all those who attended the information session consented to participate, nor did they necessarily consent immediately.

Methods literature (particularly focus group literature) about the use of groups in qualitative research emphasises the requirement that the group be homogeneous (Morgan 1988). That is, members have homogeneous backgrounds and ought not know each other in order to foster ‘contribution by more participants’ (Khan & Manderson 1992; Thomas et al. 1992: 13). However, studies have successfully used groups comprised of members with different power, status, knowledge, and interest in the research outcome and comprising members known to each other (MacDougall & Baum 1997; Kitzinger 1994).
Heterogeneous group membership supports qualitative research that aims for a wide diversity of views (Kitzinger 1994). In this research, the group composition evolved as described above, whilst adhering to Rice and Ezzy’s (2000: 79) advice to use ‘careful consideration’ and ‘common sense’. Consequently, the researcher erred on the side of practicality and took the view that any group effect would be small (Nelson & Frontezak 1988) and compensated for through the use of the research participants’ personal journals (Tuckett & Stewart, 2004).

Separate discussion groups comprised the Registered Nurses, other care staff, and residents (Tuckett & Stewart, 2004a). That is, within each nursing home there was one discussion group for residents, one discussion group for personal care assistants, and one discussion group for Registered Nurses. Discussion groups for residents and care staff included members that may or may not have known each other. Groups were homogeneous only in the sense that members met the selection criteria, understood the purpose of the research, and were willing to share their views.

**Interview: purposeful/theoretical sampling**

For the sake of clarity, in the author’s research, purposeful sampling was recognised as having the same meaning as theoretical sampling (Higginbotham et al. 2001; Morse 1991; Brink 1991; Lincoln & Guba 1985). Purposeful/theoretical sampling attempts to select research participants according to criteria determined by the research purpose but also as guided by the unfolding theorising. Purposeful/theoretical sampling tends to be used in qualitative research (Miles & Huberman 1994).

Table 1.0 below indicates some of the sampling decisions made on the basis of emerging concepts (Strauss & Corbin 1998) and the practical contingencies that also impacted on these sample decisions. For example, on the one hand, participant A-RN1 represented an ‘atypical’ case and was interviewed on the grounds that she was both a care provider and a next-of-kin for one of the residents in her care (see the second paper for further discussion). On the other hand, participant C-R1 could not be included in a follow-up interview for practical reasons – she had died.

**Table 1: Theoretical implications and practical contingency governing sampling**

<table>
<thead>
<tr>
<th>Nursing Home</th>
<th>Participant</th>
<th>Rational: Theoretical implication (Rubinstein, 1994)</th>
<th>Practical contingency</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>R1, R2</td>
<td>R1: confirming:typical, M2r: ‘atypical’ re: control/role theme</td>
<td>PCA5: social movement</td>
</tr>
<tr>
<td></td>
<td>PCA2</td>
<td>PCA2: confirming &amp; verification: typical</td>
<td>PCA4: typical but also social movement</td>
</tr>
<tr>
<td></td>
<td>RN1</td>
<td>RN1(Team leader): verification &amp; ‘atypical’ re: role/family theme</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>R1, R2, R3: confirming &amp; verification; R2 variance &amp; R1, R3: control/role theme</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PCA1, PCA2, PT3</td>
<td>PCA1, PCA2: confirming &amp; verification: typical; PT3(Physiotherapist): perspective</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RN1, RN2</td>
<td>RN1, RN2: confirming &amp; verification, RN1(Nurse Manager)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>R2, R3, R4, R5: confirming &amp; verification R4, R3: awareness theory</td>
<td>R1: deceased R4: refused interview</td>
</tr>
<tr>
<td></td>
<td>PCA1, PCA2, PCA3</td>
<td>All confirming &amp; verification</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RN1, RN2, RN3</td>
<td>RN1(Nurse Manager): comparison RN2(Team Leader): comparison, RN3: confirming &amp; verification</td>
<td></td>
</tr>
</tbody>
</table>
Therefore, research participants sought for the follow-up in-depth interview were selected purposefully (theoretically), according to the following criteria (Ezzy 2002; Schwandt 1997). The participants' data

- either confirmed in some way ideas that were emerging (typical case) (Morse, 1991) or their data offered an insight deemed to be atypical (negative case) in the context of what was being theorised (Morse 1999) and
- whether typical or atypical, required confirmation and verification (contributed to member checking) (Keith 1994) and
- could be compared with participants in other settings (contributed to constant comparison of data).

At nursing home D, a unifying idea emerged – that of understanding truthful disclosure as 'easing and omitting'. That is, telling in a particular way and telling a particular 'what' in the resident's 'best interest' (D-RN1). This Registered Nurse was sought for a follow-up in-depth interview but she had left the nursing home and could not be located. The researcher was left to consider the congruence and complementarity (triangulation) of the participant's data from her personal journal (Greene & McClintock 1985; Tuckett & Stewart 2004).

This example represents one of the practical contingencies the researcher had to manage whilst making sampling decisions on theoretical grounds. Other practical contingencies included 'social movement' (geographical movement) of other care staff (Schatzman & Strauss 1973: 75), two residents’ deaths and the refusal by one resident, one personal carer, and one Registered Nurse to continue in the research. In these cases, the researcher was again left to consider the consistency of each participant’s data as it compared to the data from the participant’s personal journal and/or as it compared with other data. During the data collection-analysis process for nursing home D, the researcher became increasingly aware that he was reaching 'information redundancy' (Lincoln & Guba 1985: 202). However, nursing home E was purposefully selected on the basis that the Administrator’s representative had commented

'You'll have to come out to us...we have a different way of thinking in the country' (researcher’s field journal notes).

The researcher collected data at E, following this lead, concurring with Charmaz that what went on there may have been critical to understanding and unique ('different') (Charmaz 1990: 1162). However, the researcher identified ‘sufficient and quality data’
(Morse 1991: 135) at E in the group discussions – as evidenced by data saturation and decided no new information would be forthcoming in follow-up in-depth interviews.

**Sampling and data saturation: constant comparison**

A clarification needs to be made about saturation *per se*. In the author’s qualitative study, the underlying search was ‘not the amount of data but rather the richness of the data, not the total counts but the detailed descriptions’ (Carey 1995: 492). Unlike the quantitative approach, deliberate frequency counts were not conducted (Morse 1995). To some degree, the claim for saturation does include an element of faith (Cutcliffe & McKenna 2002). In describing sample size for this research, it must be stated that sampling was driven by the desire to learn in detail and in depth about the experience of individuals. Hence, the final decision about sample numbers was based on evidence of data saturation (‘redundancy’) which occurred when ‘no new information of significance (was) obtained’ for ongoing thematic development and theorising (Higginbotham et al. 2001: 236; Lincoln & Guba 1985: 202; Patton 1990).

Therefore, the decision that data saturation or data redundancy had been reached was facilitated through constant comparison of data (Glaser & Strauss 1967; Glaser 1999). That is, the researcher asserted that he had saturation ‘grounded in the empirical confidence attained from repeatedly comparing data to additional data’ (Cutcliffe & McKenna 2002: 614).

The researcher moved back and forth between the data and emerging tentative thematic identification and interpretation. In this process, he ‘witness(ed) reoccurring patterns and themes in the data’ (Cutcliffe & McKenna 2002: 614). Consequently, this constant comparison of data was contingent upon concurrent data analysis and collection (Rose & Webb 1998).

Consistent with the researcher’s theoretical position, the follow-up in-depth interview was considered a symbolic interaction (Blumer 1969). This meant that he introduced research participants’ ideas and perceptions that originated within a group or across differently constituted groups. The aim of this type of questioning was two fold

- on one hand, it facilitated group comparison and,
- on the other hand it stimulated a response that was either complementary or oppositional.

Other examples of research strategies for the constant comparison of data to facilitate decisions about data saturation and ongoing sampling (Keith 1994) included the keeping of a thematic log during the group discussion (Tuckett & Stewart 2004a) and in-depth interviews, writing marginal remarks during transcript reading, keeping detailed researcher field journal notes (Tuckett & Stewart, 2004) and the ongoing reading of literature. These strategies allowed for comparison of an experience or relationship and ‘juxtaposing data from each person against each other one’ (Charmaz 1990: 1168).

**Sample characteristics**

At the conclusion of each group discussion, demographic data were self reported on a demographic information sheet. The overall sample numbers of participants in each facility and their demographics are shown in the Table 2 and Table 3.
Table 2: Total sample size

<table>
<thead>
<tr>
<th>Participants →</th>
<th>Resident</th>
<th>Personal Carer</th>
<th>Registered Nurse</th>
<th>Total sample per site</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nursing Home ↓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>6</td>
<td>7</td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>B</td>
<td>3</td>
<td>5</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>C</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>D</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>E</td>
<td>2</td>
<td>5</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>Total sample per participant</td>
<td>19</td>
<td>23</td>
<td>15</td>
<td>57</td>
</tr>
</tbody>
</table>

Table 3: Participants demographics

<table>
<thead>
<tr>
<th>Demographics→</th>
<th>Age (years)</th>
<th>Sex</th>
<th>Length of Stay (months)</th>
<th>Nursing Experience (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants ↓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residents</td>
<td>61-97</td>
<td>15 female 4 male</td>
<td>1-48</td>
<td>-</td>
</tr>
<tr>
<td>Personal Care Assistant</td>
<td>21-56</td>
<td>All female</td>
<td>-</td>
<td>1-37</td>
</tr>
<tr>
<td>Registered Nurse</td>
<td>38-56</td>
<td>14 female 1 male</td>
<td>-</td>
<td>5-38</td>
</tr>
</tbody>
</table>

Limits to sample

As stated earlier, purposeful/theoretical sampling attempts to select research participants according to criteria determined by the research purpose but also as guided by the unfolding theorising. In reality, a number of issues arose that potentially undermined the essence of purposeful/theoretical sampling, namely, ‘gatekeeper bias’, ‘sample frame bias’, and practicality and logistics (Groger & Mayberry 1999).

Gatekeeper bias

Various levels of the nursing staff at each aged care facility took a role in the initial choice of residents and care staff to be sampled. That is, either the Administrator, Manager-Clinical Care Services, or the Registered Nurse (team leader), or combinations of these three consulted amongst themselves and potential participants, thereby having control of sampling. As such, they acted as gatekeepers to the facility (Groger & Mayberry 1999; Yates et al. 1995).

Sample frame bias

The sample was framed according to the purpose of the study. Consequently, those sampled were institutional care providers and the residents. In a sense, this sample frame bias, with its focus on the resident-care provider dyad, restricted sampling of other individuals who are involved in the care relationship, e.g. the resident’s family and doctor.
Practicality and logistics

A number of practical and logistical issues biased the sampling outcome. The researcher was part-time and therefore ‘competing demands sometimes interfered with the most expeditious recruiting (and analysis) efforts’ (Groger & Mayberry 1999). Negotiation of entrée delays due to the Accreditation process (previously discussed) in a climate of finite research time meant that sampling decisions were tempered by practicalities. In the case of the failed negotiation of entrée the researcher reflected on the claimed data saturation, having failed to secure access to a facility. That is, the researcher asked: “What if?” but equally reflected that this failed entrée may not have altered anything (Groger & Mayberry 1999).

Additionally, purposeful/theoretical sampling was influenced by:

• residents’ death – one C and one D nursing home resident participant died before the researcher’s return,
• the participants’ right to withdraw from the research – at D one Registered Nurse and one personal carer withdrew due to other commitments, and at nursing home C, a resident was ‘too upset’ to continue and
• ‘social movement’ (Schatzman & Strauss 1973: 75), the geographical movement of participants away from the original aged care facility. Two personal carers at nursing home B changed their work status and could not be contacted and failed to reply to the researcher’s written requests; at nursing home D, two Registered Nurses moved to another employer (one abroad) and one personal carer left. In these latter cases, no forwarding addresses were available to the researcher.

Counter to sample limitations

The issue of limitations to the research as a consequence of sample bias can be countered by the use of a number of research strategies. One of these strategies included the use of different methods of collecting data (Tuckett & Stewart 2004; Tuckett & Stewart 2004a). The researcher was then able to consider the congruence and complementarity (triangulation) of each participant’s group and personal journal data with the data from the participant’s in-depth interview (Greene & McClintock 1985). Additionally, the recruitment of replacement participants was deemed unnecessary because of the attendant data saturation and a research design reliant on a range of data collecting methods. Discussed in the second part of this article are these and other research strategies and operational techniques that were employed in an effort to maintain rigour and create quality qualitative research.

Conclusion

Both novice and experienced nurse researchers need to take seriously the issue of sampling in qualitative research if they are to be mindful of the necessity for rigour. This methods paper contributes to the dialogue within qualitative research literature that bridges theory and practice. It analyses a researcher’s experience of sampling Registered Nurses, personal care assistants, and residents in high level (nursing home) aged care according to a sampling frame. As such, it provides not only a sound theoretical discussion but also relies on practical examples and the consequences of purposeful/theoretical sampling and data saturation and the constant comparison of data that guided the researcher’s sampling. Finally, the very real complexities of qualitative
research sampling discussed here suggest a research design open to and ready for change.

Acknowledgements

I acknowledge the supervisory role of Associate Professor D.E Stewart, School of Public Health, Queensland University of Technology, Kelvin Grove Campus, Queensland, Australia. I would like to also thank colleagues of the Centaur Memorial Fund for Nurses, Queensland, for partially funding the research titled ‘Truth-telling in aged care: a qualitative study’. This paper is adapted from the author’s Doctoral thesis.

References


Cutcliffe J, McKenna H (2002) When do we know what we know? Considering the truth of research findings and the craft of qualitative research. *International Journal of Nursing Studies*. 39, 611-618.


Kitzinger J (1994) The methodology of focus groups: the importance of interaction between research participants. *Sociology of Health and Illness*. 16,1, 103-121.


Mead G (1934) Mind, self and society. Chicago, IL: Chicago University of Press.


Morse J (1999) Qualitative generalizability. Qualitative Health Research. 9,1, 5-6.


