Selection, collection and analysis as sources of evidence in case study research


Abstract

Background Case study research is a valuable way to explore and describe nursing phenomena in their natural contexts. Multiple sources of evidence are critical in this approach. It is imperative that the strategies for selection, collection and analysis of cases are considered and articulated in the early stages of planning, to avoid having large datasets which cannot be harmonised.

Aim To critically examine what is meant by 'multiple sources of evidence' and how they can be used in case study research. Two examples of case study research are used to illustrate the decisions the authors made during the selection, collection and analysis stages of the research.

Discussion These decisions included what sources would be used, rationales for their use, and how the data would be collected and analysed. In addition, multiple sources of evidence can result in large amounts of data so the use of NVivo to manage the data is described.

Conclusion Each source of evidence selected must have a clear purpose and relate to the study's objectives. Clarification of this during the early planning of any research is imperative.

Implications for practice The authors hope that the examples provided to illustrate how multiple sources of evidence are used will guide researchers conducting case study research.

Keywords

case study research, clinical skills, cross-case analysis, multiple sources of evidence, qualitative research, psychosocial interventions

Introduction

Case study research (CSR) is becoming increasingly popular in nursing research. One of its strengths is the opportunity to use multiple sources of evidence (Yin 2003, Burns and Grove 2005, McGloin 2008, Ryan-Nicholls and Will 2009). These can provide in-depth and multifaceted explorations of phenomena. However, selecting, collecting and analysing multiple sources of evidence can be challenging and complex. The purpose of this paper is to explore, using two examples of CSR, some of the practical issues in working with multiple sources of evidence and how to address them.

Multiple sources of evidence

CSR is useful for studying a phenomenon in its natural context (Stake 1995, Yin 2003). It is used to explore real-life experiences and situations, when the phenomenon and the context in which it occurs are of interest (Yin 2003, Luck et al 2006, Salminen et al 2006). It is an appropriate design for exploring events, situations, programmes and activities (Hancock and Algozine 2006, Berg 2009).

CSR is considered suitable for nursing research because it can explore and explain contemporary phenomena (Pegram 2000). However, increasing complexity in healthcare, coupled with increasing use of case study designs in nursing research, supports the need for an examination of this methodology (Anthony and Jack 2009).

CSR is classified in several ways (Casey and Houghton 2010): a case study can be exploratory, explanatory or descriptive (Yin 1994); and it can be holistic or embedded, single or multiple, instrumental or intrinsic (Yin 1994, Stake 1995). However, a critical characteristic common to all types of CSR is the use of multiple sources of evidence (Stake 1995, Bergin and While 2000, Yin 2003). This captures the complexities of phenomena, thus making the case more complete (Walsh et al 2004).

Another objective of using multiple sources of evidence is to enhance rigour (Houghton et al 2013a). Biases in the researcher’s collection and analysis of data can be counteracted by using multiple sources of...

The most common sources of evidence in CSR are interviews, observations and documents (Yin 1994, Stake 1995, Keyzer 2000); others include surveys, physical artefacts and archival records (Moriarty et al 2007, Baxter and Jack 2008). With such a broad range of sources available, efficient and effective collection of data is imperative (Darke et al 1998).

There are challenges associated with using multiple sources. There can be overwhelming amounts of data that require management and researchers can sometimes get ‘lost’ (Baxter and Jack 2008). In addition, it can be difficult to bring the data together during analysis for reporting (Baxter and Jack 2008).

Two examples of CSR will be presented to illustrate how the authors overcame these and other challenges, emphasising the need for appropriate strategies for selection, collection and analysis of data.

**Example 1**
The first example of CSR was a multiple case study exploring the role of the clinical skills laboratory (CSL) in preparing nursing students for clinical practice. Five cases were identified; each comprised an institution offering the bachelor of nursing degree and its affiliated hospital. The main objectives were twofold: to describe the strategies for teaching and assessing in the CSL; and to explore the factors that could help or hinder how nursing students implement their learned skills in clinical settings. The overall aim was then to explore how the CSL could prepare students for that experience through teaching and assessment strategies. Identification of the case at the outset ensured a well-defined and bounded context (Stake 1995, Casey and Houghton 2010). The study was qualitative and located in an interpretivist paradigm (Houghton et al 2013a).

**Selection**
To answer the questions posed, multiple sources of evidence were needed. Three were chosen: interviews, observations and documents. Each method of collecting data needed to have a specific purpose. Semi-structured interviews were critical for gathering views about teaching and assessing in the CSL, as well as staff and students’ perceptions of factors that influence students’ implementation of skills in clinical practice. Non-participant observation was conducted to identify what happened when students implemented clinical skills in practice. Documents provided the context and background to how the teaching and learning of clinical skills occurred in the CSL and in clinical practice, as well as what was expected of students in terms of competency throughout their education and on qualification. Each source had a specific function but with triangulation could provide a comprehensive account of each case and confirm the findings of the phenomenon of interest (Murphy and Casey 2009).

**Collection**
A total of 58 semi-structured interviews were conducted by the primary researcher (CH) across the five case study sites between 2008 and 2009. Interviews are one of the most important sources of case study information (Yin 2003). They can provide rich, detailed data and gather information that is not overtly observable (Patton 2002, Shah and Corley 2006).

Participants included educators, CSL managers and technicians, clinical nurse managers, staff nurse preceptors, clinical placement coordinators, nursing students across the four years of the nursing degree programme, and newly qualified staff nurses who had recently completed the programme.

Efforts were made to ensure the equal representation of all types of participant across the five case study sites. The researcher asked participants about their perceptions and experiences of the CSL, as well as of students implementing clinical skills in practice.

The most challenging aspect of this was obtaining access. The researcher needed to establish an open, transparent relationship with gatekeepers to educational and clinical areas across the five sites (Simons 2009). Regular meetings were held, and gatekeepers and participants needed to be reassured that there was no intention to audit practice in either the CSLs or the clinical setting. Ethical approval was obtained from the research ethics committees of each institution and their affiliated hospitals. Informed consent was obtained and confidentiality ensured.

Observational data provided the opportunity to confirm verbal reports (Thompson et al 2005, Frankfort-Nachmias and Nachmias 2008). Observation is commonly used in CSR as it offers insight into behaviours and practices as they occur in their natural settings. Non-participant observation was conducted in the clinical settings of the five sites (Hancock and Algozzine 2006, Watson and Whyte 2006, Simons 2009).
A pilot was used to make important decisions about optimal positioning, length of time and the method of documenting the observations (Casey 2004). Observations in the main study were conducted by the primary researcher (CH) at each site over a 12- were used to determine students’ implementation of clinical skills throughout the day. Mobile positioning and event sampling were used. Mobile positioning involved following a particular student to certain clinical areas while they performed certain skills, as advocated by Polit and Tatano Beck (2014).

Event sampling involved observing a particular event (Casey 2004), in this case meal time and medication rounds. This method gleaned rich data that would not have been gathered through interviews alone. Observing in a clinical setting where patients, other staff and members of the public are present posed ethical challenges which needed to be addressed (Houghton et al 2010).

The third source of evidence was documents. Simons (2009) defined a document as anything written or produced about the research context or site. Documents are important for augmenting evidence from other sources and are most often used in conjunction with interviews and observations to develop a better understanding of the phenomenon of interest (Merriam 1998, Yin 1998, Shah and Corley 2006).

In this example, documents included national and international guidelines for nurse education and how best to help students learn. Competency assessments from each site were examined to understand what students were expected to learn in clinical practice.

Finally, clinical placement timetables were accessed to ascertain when students learned skills in the CSL in relation to when they went out on placement. Table 1 summarises the sources of evidence used for Example 1.

**Analysis**

The three methods used created a large volume of data from five different sites and from many different perspectives and sources. In multiple CSR, ‘within-case’ and ‘cross-case’ analysis are needed (Miles and Huberman 1994, Stake 2006, Creswell 2013): within-case analysis provides a comprehensive description of each case (Creswell 2013, Houghton et al 2015); cross-case analysis examines themes across the cases to deepen understanding and explanations (Eisenhardt 1989, Miles and Huberman 1994).

Analysis was conducted, guided by Morse’s framework (1994) and implementing strategies proposed by Miles and Huberman (1994) (Houghton et al 2015). NVivo 8 software was used to manage the large dataset and helped with cross-case and within-case analysis (Bassett 2009).

NVivo enabled the analysis of multiple sources of evidence together. In addition, using ‘queries’ to check the frequency and source of propositions made from the data ensured the confirmability of findings (Houghton et al 2013b).

It is important to note that in the final presentation of findings, only the cross-case-analysed data were presented, to preserve the confidentiality of the research sites as they may have been identifiable from the within-case analysis.

Comparing data gathered from multiple sources can confirm the extent to which findings can be verified (Murphy and Casey 2009). NVivo provided an audit trail, which demonstrated how the data were integrated and analysed together within cases and across cases. Furthermore, query tools ensured that any issue described in the findings could be examined and compared from different sources, types of participant or research site (Houghton et al 2013b).

**Example 2**

The second example was a multiple case study which explored the experiences of mental health nurses (MHNs) trained in using psychosocial interventions (PSIs) in the care of patients with mental health problems. Consistent with the goal of understanding experience, a qualitative approach was adopted using a multiple case study design. Four public sector Health Service Executive (HSE) sites were used in the study.

| TABLE 1. Overview of sources of evidence in Example 1 |
|---------------------------------|---------------------------------|---------------------------------|
| **Semi-structured one-to-one interviews (n=58)** | **Non-participant observation (five days)** | **Documents** |
| **Junior nursing students (n=10)** | One day (12-hour shift) per site conducted in two-hour observation sessions | National and international organisational guidelines |
| **Senior nursing students (n=10)** | Competency assessment documents | |
| **Newly qualified staff nurses (n=8)** | Clinical placement timetables | |
| **Clinical staff (n=15)** | | |
| **Academic staff (n=15)** | | |
Selection
As in example 1, multiple sources of evidence provided answers to the overall aim and addressed the objectives of the research. In this study, the data sources comprised non-participant observations and one-to-one semi-structured interviews, supplemented by field notes. The observations captured potential PSIs that nurses use in practice.

The interviews allowed the participants to talk freely about the topic (Polit and Tatano Beck 2014), while the supplementary field notes helped to make sense of the observations and the contexts in which the nurses worked (Gerrish and Lacey 2010). The field notes also helped in obtaining enriched meanings for the interview data.

Collection
The primary researcher (SS) gathered the interview data between 2012 and 2013 from 40 PSI-trained nurses across inpatient and community settings: 33 from community settings, seven from inpatient settings. In total, 16 of the nurses were clinical nurse specialists, 12 were staff nurses, 10 were clinical nurse managers, and two were advanced nurse practitioners. A multidisciplinary team meeting was also observed to help keep the findings related to the context. Observational data were collected through continuous note-taking during and after the observations, followed by reflective commentaries. In total, 19 hours and 20 minutes of non-participant observations were completed across the four cases. These were undertaken at convenient times for the participants.

Ethical approval was initially sought from the research ethical committee of the institution where the researcher was studying and from the four different hospital research ethics committees. Informed consent was obtained and participants’ well-being was always given priority over the study. Table 2 provides an overview of the different types of participant and the numbers of interviews and observations in this study.

<table>
<thead>
<tr>
<th>Data collection methods</th>
<th>Participants</th>
<th>Community PSI-trained MHNs</th>
<th>Inpatient PSI-trained MHNs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Semi-structured one-to-one interviews</td>
<td>PSI-trained MHNs (40 interviews)</td>
<td>33</td>
<td>7</td>
</tr>
<tr>
<td>Non-participant observation</td>
<td>PSI-nurses (19 hours, 20 minutes) and multidisciplinary meetings</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Field notes (reflective notes: supplementary)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

PSI can refer to a wide range of interventions, so a semi-structured observation tool was adapted from the literature (Spradley 1980), based on the researcher’s clinical experience as a MHN. This provided a focus to achieve a more comprehensive understanding of the participants’ experiences of using PSIs in practice. An interview schedule was used to guide the interviews. The field notes were managed by adapting Spradley’s (1980) format, which considers three sorts of field note: condensed accounts, expanded account and a fieldwork diary. Writing field notes during the study helped to identify patterns in the data and served to increase the quality of the research’s findings (Stake 1995).

For example, the observational data enabled the descriptions to be checked against facts, as it was possible to clarify afterwards in the interviews any discrepancies regarding how nurses use PSI in practice. Therefore, by observing and interviewing the same participants, the dissimilarities between what participants said and did could be acknowledged (Mays and Pope 1995).

Houghton et al (2013b) stated that when data are gathered from multiple sources, comparisons can be made to establish the degree to which findings can be confirmed. The field notes helped to provide contextual data; they served to support the observations and interviews by enhancing meaning and generating a more holistic and complete account of the phenomenon. Therefore, each modality assisted in providing the indepth information required to address the research’s aim and objectives. However, one challenge was the use of the non-participant observations. It was clear from when MHNs were approached about the study that most of them were reluctant to be observed. Many MHNs had queries about consent, particularly from the client’s perspective. Mulhall (2003) said these issues are compounded by the unpredictability of observation work.

Participants were reassured that it was unnecessary to gain written consent from
the clients, because they were the focus of the research, not the clients. However, verbal consent was obtained from all clients interacting with the MHNs being observed.

**Analysis**

Ritchie and Spencer’s (1994) framework approach was adapted as it enabled flexibility throughout analysis, to detect important issues in the data and interpret the data’s meaning. This approach is consistent with multiple CSR. It uses five stages, each stage consisting of several cycles, with the researcher able to use two types of analysis: within-case and cross-case (Miles and Huberman 1994, Stake 2006, Creswell 2013). Within-case analysis involved outlining a thorough description of each case and its themes; cross-case analysis allowed themes to be analysed across the cases and the researcher could search for similarities and differences between the cases.

Arguably, an audit trail is the most important criterion for establishing the trustworthiness and credibility of a study. In this study, every step of the analysis was accounted for because of the diagrams and text used during the various cycles of the analysis.

**Conclusion**

Multiple sources of evidence are an important feature of CSR and can enhance the rigour of a study. This paper examined the practical challenges involved when collecting multiple sources of evidence. Each source of evidence must have a clear purpose and relate to the study’s aims and objectives. Articulation of these

**References**


purposes when planning the study can help the researcher to fully explore the phenomenon of interest, but in a bounded context. It may also prevent a novice researcher using CSR from getting lost in irreconcilable data.

It is the very purpose of multiple sources that can pose the greatest challenge: multiple sources provide a comprehensive picture of a case, but it is that thoroughness that can potentially make gatekeepers and participants wary. The researcher should make the nature and purpose of the research explicit to key individuals openly and transparently; the way in which the data will be presented must also be determined at an early stage from an ethical perspective, so that gatekeepers and participants can be reassured that they will not be recognisable in the final report.

An additional challenge illustrated in the examples pertains to the complexity of managing data from different sources. NVivo or similar qualitative analysis software is therefore a crucial way to store and manage the large datasets generated.

Ultimately, careful consideration of how each source of evidence will help to answer the research question ensures that the multiple sources used are appropriate. Planning for how the data will be managed will prevent researchers from becoming overwhelmed during analysis. For novice CSR researchers, using the term ‘selection, collection and analysis’ in articulating how the multiple sources of evidence are selected may help to structure and make more explicit the decisions required to answer the research question.


