Rigour in qualitative case-study research

Cite this article as: Houghton C, Casey D, Shaw D, Murphy K (2013) Rigour in qualitative case-study research. Nurse Researcher. 20, 4, 12-17.

Date of submission: May 2 2012. Date of acceptance: September 18 2012.

Abstract

**Aim** To provide examples of a qualitative multiple case study to illustrate the specific strategies that can be used to ensure the credibility, dependability, confirmability and transferability of a study.

**Background** There is increasing recognition of the valuable contribution qualitative research can make to nursing knowledge. However, it is important that the research is conducted in a rigorous manner and that this is demonstrated in the final research report.

**Data sources** A multiple case study that explored the role of the clinical skills laboratory in preparing students for the real world of practice. Multiple sources of evidence were collected: semi-structured interviews (n=58), non-participant observations at five sites and documentary sources.

**Discussion** Strategies to ensure the rigour of this research were prolonged engagement and persistent observation, triangulation, peer debriefing, member checking, audit trail, reflexivity, and thick descriptions. Practical examples of how these strategies can be implemented are provided to guide researchers interested in conducting rigorous case study research.

**Conclusion** While the flexible nature of qualitative research should be embraced, strategies to ensure rigour must be in place.

**Keywords** Multiple case study research, rigour
documentary analysis. Data were analysed using Morse's (1994) analytical framework, which outlines four important stages: comprehension, synthesis, theorising and re-contextualisation. These steps were achieved by using the strategies proposed by Miles and Huberman (1994) which include the following: open coding, broad coding, pattern coding, memoing, distilling and ordering, testing executive summaries, developing propositions.

All the data were managed using the software program NVivo 8. The findings of the study outlined students' experiences in clinical practice and how the teaching and assessment strategies in the CsL were perceived. Recommendations were made for maximising the strategies used in the CsL and potential strategies that should be considered in undergraduate nursing education.

**Approaches to rigour**

The four criteria proposed by Lincoln and Guba (1985) – credibility, dependability, confirmability and transferability – formed the framework for determining the rigour of the research.

- **Credibility** refers to the value and believability of the findings (Lincoln and Guba 1985, Leininger 1994, Polit and Tatano Beck 2006) and involves two processes: conducting the research in a believable manner and being able to demonstrate credibility.

- ‘Dependability’ is often compared to the concept of reliability in quantitative research and refers to how stable the data are (Graneheim and Lundman 2004, Tobin and Begley 2004, Shah and Corley 2006, Rolfe 2006).

- **Confirmability** refers to the neutrality and accuracy of the data (Tobin and Begley 2004), and is closely linked to dependability – the processes for establishing both are similar.

- **Transferability** refers to whether or not particular findings can be transferred to another similar context or situation, while still preserving the meanings and inferences from the completed study (Leininger 1994).

Specific strategies for ensuring rigour in this framework were needed. Table 1 illustrates four processes for determining trustworthiness in qualitative research and outlines the main strategies for addressing them. Each strategy is described in more detail and illustrated using the research example provided.

**Credibility**

**Prolonged engagement and persistent observation**

Lincoln and Guba (1985) highlighted how prolonged engagement and persistent observation can enhance the credibility of research. These skills require researchers to spend sufficient time in the field or in case-study sites to gain full understanding of the phenomena being investigated (Altheide and Johnson 1994). The lack of any new emerging data is evidence that saturation has been achieved.

In the CsL study, non-participant observations were conducted over a 12-hour shift in each of the affiliated hospitals in the five case-study sites. The lead author spent sufficient time in the field or case-study site to gain a full understanding of the phenomenon under investigation. During the final non-participant observations no new concepts were emerging, which was confirmed in the analysis, indicating that saturation had been achieved.

**Triangulation** Credibility can also be enhanced with triangulation, which uses several methods to study one phenomenon (Polit et al 2001, Holloway and Wheeler 2002, Burns and Grove 2005). The two main purposes of triangulation are to ‘confirm’ data and to ensure data are ‘complete’ (Begley 1996, Shih 1998, Casey and Murphy 2009). Confirmation is the process of comparing data gathered from multiple sources to explore the extent to which findings can be verified (Casey and Murphy 2009). If data gathered through different methods are found to be consistent, this can increase confidence in the credibility of findings (Knafl and Breitmayer 1991).

Completeness of data is concerned primarily with gathering multiple perspectives from a variety of sources so that as complete a picture as possible of phenomena can be portrayed (Shih 1998, Casey and Murphy 2009). In the context of case-study research, a major strength of the design is the opportunity to use different sources of evidence through triangulation (Yin 2003, Burns and Grove 2005, McGloin 2008, Ryan-Nicholls and Will 2009).

In the CsL study, triangulation was used to collect qualitative data about the role of CSL. For
confirmation, data gathered from multiple sources were compared to determine the extent to which findings could be confirmed. For example, observations were conducted in the clinical area to identify the factors that helped or hindered students to implement their skills in clinical settings. Interview participants were also asked to discuss their perceptions of the factors that helped or hindered students. The similarities in the results of these two methods confirmed findings.

Completeness of data was concerned primarily with gathering multiple perspectives from various sources so that as complete a picture as possible of phenomena was portrayed. Each of the three approaches to gathering data had specific advantages that, when used together, maximised the potential for in-depth insight and completeness into the cases and their context.

**Peer debriefing** Some authors advocate using an external colleague or ‘expert’ to support the credibility of findings (Appleton 1995, Burnard 2002, Casey 2007a). However, the usefulness of this approach is debated in the literature. Lincoln and Guba (1985) suggested that peer debriefing may make researchers’ interpretation of the data more credible if peers define the data in the same way. Analysis in qualitative research is an individual, unique process between the researcher and the data (Schutz 1994, Cutcliffe and McKenna 2004). Thus, no two researchers will interpret the data in the same way (Andrews et al 1996, Cutcliffe and McKenna 1999, McBrien 2008). Ryan-Nicholls and Will (2009) recommended using peer debriefing in qualitative research, but with caution. The aim should not be for independent analysts to arrive at the exact same coding and thematic structure as the researchers. Rather, the purpose should be to see if they agree with the data labels and the logical paths taken to arrive at those labels (Graneheim and Lundman 2004).

In the CSL study, the purpose of peer debriefing was to verify whether or not an expert would agree with the coding process. An expert in qualitative research was asked to code three interview transcripts. She was then asked to review the coding conducted in this research for comparison. The coding carried out by the expert corresponded to that of the researcher in that the same main issues for discussion arose. This enhanced the credibility of the study findings.

**Member-checking** involves allowing participants to read the transcription of their interviews to ensure that these have been accurately recorded and are therefore credible (Sandelowski 1993, Altheide and Johnson 1994, Koch 1994, Cormack 2000, McDonnell et al 2000, Stake 2006).

There are challenges to using member-checking for the purpose of establishing rigour (Koch and Harrington 1998). One issue relates to the stage in the research at which participants are shown the findings. If they are shown the verbatim transcripts, they will be able to acknowledge and respond to their own words (Glaser and Strauss 1967, Melia 1982). They will have no contribution to how their opinions and assertions are interpreted and portrayed. Koch (1994) said that respondents should read the constructions derived from the analysis. However, asking participants for feedback on the reported data following analysis brings different challenges. If the study results have been synthesised, decontextualised and abstracted from individual participants, there is no reason for individuals to be able to recognise themselves or their particular experiences (Sandelowski 1993, Morse et al 2002). This suggests that member-checking should occur following transcription rather than after analysis.

In the CSL study, member-checking was conducted following transcription, so that participants could acknowledge and respond to their own words. Each participant was sent a copy of the transcribed interview with an accompanying letter. Participants were given the option of contacting the researcher with any queries they had regarding the data. None of the participants had concerns or reservations about the content of the interviews. However, of the 58 participants, three were concerned about their use of language in the interview. Following reassurance, they agreed that the transcripts should remain unchanged.

**Dependability and confirmability**

**Audit trail** Rigour can be achieved by outlining the decisions made throughout the research process to provide a rationale for the methodological and interpretative judgements of the researcher. Koch (1994) believed that while readers may not share a researcher’s interpretation, they should nonetheless be able to discern the means by which it has been reached.

Discerning the means can be achieved by an audit trail, which is an essential component in a rigorous study (Bryar 1999, Ryan-Nicholls and Will 2009). To assess the trustworthiness of a study, it is necessary to examine the process by which the end-product has been achieved and present faithful descriptions recognisable to the readers (Rubin and Rubin 1995, Horsburgh 2003). The audit trail is maintained through comprehensive notes related to the
contextual background of the data and the impetus and rationale for all methodological decisions (Glaser and Strauss 1967, Rodgers and Cowles 1993, Ryan-Nicholls and Will 2009). NVivo can enhance the rigour of the research by providing a comprehensive ‘trail’ of decisions made during data collection and analysis (Richards 1999, Bringer et al 2004, Silverman 2010). Furthermore, the ‘query tools’ in NVivo allow the researcher to audit findings and guard against excessive emphasis on rare findings that happen to suit the researcher’s preferred argument (Bassett 2009, Silverman 2010, Bergin 2011).

NVivo provided a record of decisions during the analysis phase of the CSL research. These findings were illustrated in the final research report. In addition, the dependability and confirmability of the data were demonstrated by running queries within the qualitative data analysis software. NVivo was able to locate all the passages that matched the criteria set in a query. Locating ensured that any issue described in the findings was not the perception of just one person, but rather confirmed that a number of participants held the same opinion. In addition, this confirmation guaranteed dependability in different contexts.

Three types of query were used in this research: text search, coding and matrix queries. The text-search query involved searching the interview transcripts and observational field notes for specific words that were important terms or code labels for analysis. The text-search query was useful for identifying concepts throughout and following the analysis to check whether or not the concepts had been sufficiently represented in the original coding strategy and whether or not they were worthy of further exploration. An example of this text-search query related to the use of mannequins, which was a common theme across interviews with all participants. While the literature would refer to mannequins and Laerdal’s SimMan, participants sometimes used the term ‘dummy’ when referring to the simulation models. Therefore, when looking for repeated references and discussion surrounding the mannequins, ‘mannequin’, ‘dummy’ and ‘dummies’ were included in the text-search query. Once the words had been located, the surrounding paragraphs were read to ensure that they related to the topic. These passages were compared with the original coding to ensure that no interview data had been omitted that examined mannequins.

Coding queries were also used to search and retrieve all coding at a node that provides a storage capacity for references to coded text (Bazley 2007) or all data assigned to a specific attribute. An attribute stores factual information, such as demographic data about participants in order to seek patterns and ask questions about the data at a later stage (Richards 1999).

This approach was used to check or confirm propositions made from the analysis. For example, it was highlighted that an issue raised by the clinical staff was referred to as ‘teaching from scratch’. This term was used to describe the fact that some clinical staff repeated training already undertaken in the CSL. The proposition suggested that it was mainly clinical staff raising this issue. A coding query allowed for this to be confirmed. Table 2 (page 16) illustrates the results of this coding search. It is evident that it was primarily the clinical staff that referred to this issue.

Matrix-coding was the third query strategy used to check the rigour and repeated instances in the research. Matrix-coding queries provided the opportunity to compare multiple nodes and/or attributes as a numeric table (Bassett 2009). One example provided in Table 3 (page 16) related to the question of where participants thought students mainly learned their clinical skills. When a matrix-coding query strategy was used, participants’ views were clear. The number of participants who referred to this issue and their perceptions were also evident. The aim was to confirm the findings and ensure that the participants’ views were not misrepresented rather than to quantify their responses. In this example, it was apparent that most of the participants thought that students learned their clinical skills in the clinical area.

Reflexivity It is important that decision trails are not stripped of the personal contributions of researchers and that personal responses are also recorded (Jasper 2005). In most qualitative research, the researcher is considered part of the research instrument (Rodgers and Cowles 1993). Therefore, the credibility of a study rests on the procedures implemented and the self-awareness of the researcher throughout the research process (Stoecker 1991, Rodgers and Cowles 1993, Mantzoukas 2005).

Maintaining a reflective diary can be an important expression of reflexivity (Rodgers and Cowles 1993, Koch 1994, Koch and Harrington 1998, Johnson 1999, Jootun et al 2009). This reflexive account highlights how the researcher’s history and personal interests brought them to the research and demonstrates how the theoretical perspective affected data collection and research (Van Maanen 1991, Toffoli and Rudge 2006). A reflective diary should provide the rationale for decisions made, instincts and personal challenges that the researcher experienced during research (Primeau 2003, Rolfe 2006). In the CSL study, a reflective diary illustrated the transparency of
in providing detailed descriptions for the reader to make informed decisions about the transferability of the findings to their specific contexts (Lincoln and Guba 1985, Firestone 1993, Stake 1995, Bogdan and Biklen 2003, McKee 2004). The emphasis should be on creating ‘thick’ descriptions, including accounts of the context, the research methods and examples of raw data so that readers can consider their interpretations (Stake 1995, Popay et al 1998, Dawson 2009). Ultimately, the reader can decide whether or not the findings are transferable to another context (Graneheim and Lundman 2004). A rich and vigorous presentation of the findings, with appropriate quotations, also enhances transferability (Graneheim and Lundman 2004).

In the CSL study, thick descriptions for the purpose of enhancing the transferability of the study were provided. In the final report, similar to the work of Casey (2007b), detailed and appropriate descriptions were offered so that readers could make informed decisions about the applicability of the findings to specific contexts. The necessary details included accounts of the context and examples of raw data so that alternative interpretations could be considered. For this purpose, direct quotes from the participants were illustrated (Houghton et al 2012). In addition, excerpts from the field notes showed how the themes developed from the data.

Conclusion
This paper outlines the strategies that can be used to ensure the credibility, dependability, confirmability and transferability of qualitative research. While the flexible nature of qualitative case studies should be embraced, strategies to ensure the rigour of such studies need to be in place. The research example illustrated a practical approach to implementing these strategies. QSR NVivo is beneficial as a data management tool, and as demonstrated by this paper, it can provide a comprehensive audit trail to depict decisions made during the research process.

| Table 3 | Matrix coding query: responses to where students learn their clinical skills |
|---------|---------------------------------|---------------------------------|---------------------------------|
| Participant type | Skills learned mainly in the clinical area | Skills learned mainly in the clinical skills laboratory | Skills learned both in the clinical skills laboratory and in the clinical area |
| Clinical staff (n=15) | 11 | 0 | 0 |
| Academic staff (n=15) | 9 | 1 | 3 |
| Students and newly qualified nurses (n=28) | 20 | 1 | 1 |
References


Poppay J, Rogers A, Williams G (1998) Rationale and standards for the systematic review of qualitative literature in health services research. Qualitative Health Research. 8, 3, 341-351.


