



An interdisciplinary approach to secondary qualitative data analysis: what why and how

| | |
|------------------|--|
| Title | An interdisciplinary approach to secondary qualitative data analysis: what why and how |
| Author(s) | Rodriguez, Leonor |
| Publication Date | 2021-02-16 |
| Publisher | Edward Elgar |

An interdisciplinary approach to secondary qualitative data analysis¹: what, why and how

Leonor Rodriguez

Living in global data-rich societies implies the development of critical appraisal to evaluate the quality, authenticity and meaning of data that is increasingly more accessible. Data is readily available, which is an advantage; however, not all data is good data, ethically sourced and governed under the same laws and principles in every part of the world. At this time, vast amounts of data are being collected and archived worldwide, therefore the use of existing data for further analysis is increasingly more prevalent (Johnston, 2014). Data availability has increased due to the efforts of organisations to create and maintain datasets in open and accessible ways and the advances in statistical software which have facilitated greater ease of manipulation (Trinh, 2018).

Secondary data analysis has the capacity to effectively use and make sense of readily available data; however, this also comes with advantages and challenges that will be explored in depth in this chapter. A case is made on the multiple benefits of engaging in secondary data analysis whilst highlighting the potential difficulties that should be given careful consideration. Secondary data analysis is an under-used methodological technique and the awareness of its benefits and how to overcome its limitations may encourage its use further (Irwin, 2013; Smith, 2008).

This chapter provides general guidelines that can be useful in different fields, particularly targeted at practitioners, policy makers and researchers from different backgrounds. The chapter, however, targets all levels of expertise from the very novice to experienced users of qualitative data. As proposed and described in this chapter, secondary qualitative data analysis is defined as an innovative and creative yet rigorous and systematic research design that can respond to the fast-changing data availability across different sectors. Additionally, funders across different countries now encourage and expect researchers to consider data sharing as part of their funding proposals (Irwin, 2013). It is, therefore, a very pertinent time to explore, understand and engage in secondary data analysis.

WHAT IS DATA?

Primary and Secondary Data

Data has been defined as all sources of information that are produced by people in their contexts (Largan and Morris, 2019). Differentiating between primary and secondary data is important as researchers have used several definitions to describe both concepts. Boslaugh (2007) explained that the difference between primary and secondary data is the role of the person who collected the data and the person who is analysing it. Data collected by a researcher or research team for analysis under the topic of consideration and using procedures that fit the research problem is defined as primary data, whereas data collected by another person for another purpose is secondary data (Boslaugh, 2007; Hox and Boeije, 2005; Largan and Morris, 2019). The same data source can, therefore, be used as primary as well as secondary data. In secondary data, the person carrying out the analysis was not involved in the original research design and data collection (Boslaugh, 2007; Hox and Boeije, 2005; Largan and Morris, 2019). Primary data analysis is limited to analysis carried out to answer the research hypothesis proposed in the original study (Cheng and Phillips, 2014). All other analyses of data collected or analysed for other purposes are denominated as secondary analyses of existing data (Cheng and Phillips, 2014).

Datasets have a wide variety of types such as public government datasets and health care databases; these are characterised by having large sample sizes spread across a wider geographical area and are, therefore, generalisable. Private organisations (professional, community and voluntary) may have their own datasets (Dunn et al., 2015) and these sources of data can also be stored, archived and used for further analysis in secondary data analysis. Therefore, there are several types of data and data sources increasing the potential for further analysis of secondary data.

QUALITATIVE RESEARCH

This section provides a definition of qualitative research as a paradigm as well as providing a comprehensive approach to some of the different forms of qualitative data. Secondary qualitative data analysis will also be defined and introduced.

Qualitative research can have different approaches as they respond to different research question types and qualitative researchers come from a variety of research traditions; therefore, qualitative research is a complex combination of elements under the same name (Hancock et al., 2009). Qualitative research has been defined as a type of social research based on text data rather than numerical data and is aimed at understanding the meaning of human action in the context as it occurs

rather than testing a predetermined hypothesis (Carter and Little, 2007). Qualitative research examines how people make sense and give meaning to themselves, others and their daily lives (Hox and Boeijie, 2005). Essentially, it is the systematic collection, organisation and analysis of text originating from speech or conversation (Grossoehme, 2014).

Table 9.1 Characteristics of qualitative research

| Characteristics | Definitions |
|-----------------------------|--|
| Natural setting | The research study and data collection are carried out in a specific geographic location, time and context determined and controlled by the research participants. |
| Researcher as a participant | Researchers are perceived as participants by research subjects. |
| Subject-based communication | Research participants identify and select the topics of communication they wish to approach. |
| Subject intentionality | Researchers intend to capture and preserve participants' communication styles and symbols the way in which they intended. |
| Pragmatic | Research results produce insights into ongoing social processes and outcomes with the intention of solving an existing social issue. This may be intended to generate theory but not always. |

A unique characteristic of qualitative data is that it is usually generated through personal interactions between researchers and study participants (Irwin, 2013); qualitative researchers affect and are affected by the data (Grossoehme, 2014). This interaction between participants and researchers has to be based on a relationship of trust and care, which has important and significant ethical implications. The duty of care from researchers to participants is to guarantee their anonymity and safety (Irwin, 2013). According to Carter and Little (2007) this relationship is very significant as knowledge constructed by a researcher in a particular context and the interactions with participants will be different from the knowledge that would be constructed with different participants in a different place and time. Qualitative research involves collecting and analysing subjective data that is shaped by the social, political and cultural context at the time it was created (Ruggiano and Perry, 2019).

Another critical aspect of qualitative research is the researchers' thinking and interpretation which generally develops during the writing process (Carter and Little, 2007), therefore, having records of this analytical processes and decision-making processes are important. All of these elements, therefore, can have important implications. The use of qualitative data for secondary data analysis, context and researcher– participant interactions needs to be acknowledged and carefully considered as they are an inherent part of the data.

As mentioned above, qualitative research can include different methods and traditions under the same name; however, there are some common characteristics.

Chesebro and Borisoff (2007) described some of these and they are included in Table 9.1.

Another characteristic of qualitative research is the type of sampling, which is usually purposive, which means that samples are intended to serve a research purpose instead of being representative of a population (Carter and Little, 2007). In qualitative research, samples can be determined in advance or evolve gradually over the research process. Some qualitative approaches such as Interpretative Phenomenological Analysis and Narrative Analysis require small samples (Hancock et al., 2009). Different types of sampling are pertinent for qualitative research, for example theoretical sampling, convenience or purposive sampling,² quota sampling,³ snowball sampling⁴ and typical case sampling⁵ (Hancock et al., 2009; Mack et al., 2005). One of the aims of qualitative research is to achieve theoretical saturation, which means that additional cases do not change the coding frame; this is often used as an indicator of an adequate sample size (Hancock et al., 2009).

Validity and reliability in qualitative data are crucial to determine the quality of the data and, therefore, the findings of the primary research and potential for the data to carry out further secondary data analysis. Different strategies can be used to determine the quality and reliability of qualitative data. Validity refers to whether the final product portrays what it claimed to portray and reliability refers to whether the results can be replicated to obtain the same result (Grossoehme, 2014). Researchers should keep fieldwork notes to record the choices they make, which will facilitate replication (Hox and Boeije, 2005). Other methods to validate qualitative research include triangulation, peer debriefing and external audits (Hox and Boeije, 2005; Grossoehme, 2014). Member checking specifically refers to bringing back the findings to research participants and asking for their feedback to ensure the findings are valid and reliable and a true reflection of their views and opinions (Grossoehme, 2014).

Another aspect to consider regarding qualitative research validity and reliability is the influence that researchers have on the study; Grossoehme (2014) has described that the investigator is part of what is being studied and influences it, which does not devalue the study but instead enhances it. Qualitative researchers need to be aware of the impact they might have on the findings and keep a record of the decision-making processes; this was termed by Largan and Morris (2019) as 'reflexivity'. The availability of primary researchers' records provides the evidence for validity and reliability of the data generated. Secondary data analysts should ensure they have access to these records, understand the way data was generated, its strengths and limitations and how these may impact on the secondary data

analysis. Trinh (2018) highlighted that it is crucial for researchers to have a broad understanding of the limitation of a dataset as this will impact on study findings.

QUALITATIVE SECONDARY RESEARCH

Qualitative secondary research is a systematic method of using existing data to provide new ways of understanding which may be additional to or different from the original data and analysis (Largan and Morris, 2019). It can also be used to address other research questions, evaluate new hypotheses or use a different analytical approach which might change the conclusions from the original analysis (Boslaugh, 2007; Dunn et al., 2015). Researchers have highlighted qualitative secondary data analysis as a systematic research method with specific procedural and evaluative steps (Johnston, 2014). Qualitative secondary data analysis can be carried out on the original data or published data (Church, 2001). Some of the methodologies that exist which can be used for qualitative secondary data analysis are included in Table 9.2 below.

Benefits and Challenges of Qualitative Secondary Research

The following section provides a critical approach to secondary qualitative data analysis. The strengths of qualitative secondary data analysis include resource saving (time and money); it can be used on several different types of datasets; and has ethical benefits for research subjects. Qualitative secondary data analysis also has challenges, including the quality of the data; accessing suitable datasets; determining the suitability of the data for the proposed analysis; and ethical considerations with participants and data.

Benefits of qualitative secondary data analysis

Secondary data analysis is a flexible research methodology as it can be successfully used with small datasets as well as large datasets. Secondary data analysis can be very suitable for handling large datasets at an organisational, national or international level at a lower cost for smaller research teams who may not be able to afford the creation of large datasets with the staffing and resources available (Smith, 2008). The fact that these types of datasets have become more accessible for researchers increases the possibilities of using these large datasets for secondary data analysis (Smith, 2008). Secondary datasets can also be cross-sectional and longitudinal, which enables secondary data analysis of longitudinal analysis at a lower cost than if primary data was collected over time (Dunn et al., 2015). Datasets can also be obtained from different sources including international repositories, or internet-based data can also be analysed (Largan and Morris, 2019). Secondary data analysis can also be used to increase sample sizes and statistical power (Ruggiano

and Perry, 2019), as well as the validation of instruments and testing complex theories (Clarke and Cossette, 2000).

This means that a wider variety of research questions can be answered using this method as different sample sizes, types of datasets and data sources can be used. Secondary data analysis can also be used for educational purposes and skills building in students and early-career researchers, for example statistics, software skills, data analysis, report writing and dissemination (Smith, 2008). Flexibility, accessibility and reduced research costs have earned secondary data analysis the reputation of being a 'democratic' methodology as it is available at a lower cost for a wider variety of researchers including independent, early career researchers and students (Smith, 2008).

Regarding time and costs specifically, secondary data analysis can significantly reduce or eliminate data collection costs (Largan and Morris, 2019). In the absence of funding or research grants, developing longitudinal and extensive datasets can be challenging (Dunn et al., 2015). Collecting data requires long periods of time and travel costs for researchers if data is collected face to face. Additionally, some online platforms may charge for licences to carry out surveys and data collection online. Even though some databases have to be purchased, the expenses of salaries, transportation and so on are lower (Boslaugh, 2007; Dunn et al., 2015). Secondary data analysis also eliminates the need to provide incentives to participants (Dunn et al., 2015). Researchers can spend more time analysing data, as secondary data is usually already collected, cleaned and stored in electronic format (Boslaugh, 2007).

Supporting the argument that this methodology can reduce research costs, secondary data analysis can also allow projects to be developed and completed much faster and therefore the contribution of new knowledge can be achieved in a timelier manner (Johnston, 2014). Secondary data analysis can also be a way of increasing cost-effectiveness of publicly funded research as it adds value and extends the original investment (Brezna, 2016; Rodriguez, 2018).

Another aspect to consider as part of the benefits of secondary data analysis is ethics. Ethical considerations and the limits of informed consent are an object of discussion and consideration in qualitative secondary research due to the nature of qualitative data and the way it is generated (Irwin, 2013). Ethical considerations in secondary data analysis are important as ethically sound databases are the most suitable on which to carry out further analysis as they show an awareness of research integrity. This can be an indicator of high-quality research and data, which can be an advantage for secondary data analysts as it provides them with the confidence that their dataset is adequate for the study. Datasets that are ethically questionable should be discouraged as their findings could potentially be inadequate or harmful for participants.

An ethical strength of secondary data analysis is that this type of analysis usually implies a lower risk for participants. This is particularly so in research that explores sensitive topics as participants do not have to repeat their experiences, avoiding distress and potential re-victimisation (Largan and Morris, 2019). Secondary data analysis can reduce the risk and burden on participants particularly for vulnerable populations and hard-to-reach groups (Largan and Morris, 2019; Smith, 2008; Irwin, 2013).

Secondary data researchers have an obligation to evaluate the ethical circumstances under which the original data was collected. This can be a challenge as there are cultural differences and variations between committees and organisations in terms of what is considered ultimately ethical research behaviour. A general suggestion is that ethical decisions made in the primary and secondary data analysis are explicit in the documentation or reports attached to the datasets. As a basic ethical requirement, databases used for secondary analysis must be carefully anonymised before information is stored or shared so the identity of the original research participants is protected (Dunn et al., 2015; Largan and Morris, 2019). Research participants need to be informed about the use of their data in future research; when the specific use of the data is unknown, generic statements about the use of the data are required, for example, participants can be informed that their data will be available for future research and/or teaching purposes (Irwin, 2013). Johnston (2014) has suggested that secondary data analysis should abide by the conditions included and described in the original consent form, which may allow or impede the use of data for secondary analysis. Ethical considerations, therefore, can have both benefits and challenges for qualitative secondary data analysts. Other challenges of secondary data analysis are explored in the next section.

Challenges of qualitative secondary data analysis

One of the first and crucial challenges of secondary data analysis is determining the quality of the primary datasets to ensure these are of high quality, reliable as well as suitable for the proposed research questions (Largan and Morris, 2019). Determining data quality can be a challenge for researchers, as researchers need to know how to access existing datasets, evaluate their quality and have the necessary knowledge or funds to hire staff with the expertise to carry out this work and relevant evaluation (Dunn et al., 2015).

Another significant challenge to overcome is that secondary data has an inherent limitation: it was not collected for your specific research questions (Boslaugh, 2007; Smith, 2008) and therefore may not always be suitable to answer all research questions. Additionally, data is socially constructed, which means data may be impacted by underestimation due to administrative errors, conflation of figures, and official and unofficial exclusion of research subjects among others. Even though data

may have limitations, official statistics and data are crucial in social research (Smith, 2008). Therefore, to determine the quality of the data, secondary data researchers need to be aware of the original purpose of the data collection and the theoretical frameworks underpinning the data, the characteristics of the data, the target population and the wording of the questions used in the data collection (Boslaugh, 2007; Dunn et al., 2015) as this may have an impact on the potential findings and the capacity of the dataset to answer a specific secondary research question.

Even though qualitative secondary data analysis can have many benefits, some authors have argued that due to the nature of this kind of data, qualitative secondary data analysis is limited. As described earlier, qualitative data is generated by the interactions between researchers and participants and therefore the knowledge created is dependent on the context in which it occurs and the understanding of this is crucial to the understanding and analysis of the data (Irwin, 2013). Secondary data analysts may have limited access to the context in which data is generated and, therefore, could have a limited understanding of the data itself (Irwin, 2013). This issue, described as a limitation of qualitative research, can be approached with the use of detailed data collection diaries and data collection recording methods that can provide a detailed account of the data collection process and all the decisions made at different stages of a qualitative research study.

Secondary data researchers need to ensure the datasets can answer their proposed aims and research questions. To make an informed decision, aspects such as data transferability, the characteristics of the original sample and the population should be considered, as well as the effect of potential confounding variables such as nationality, gender and age (Largan and Morris, 2019). Additionally data may not have been collected or categorised in a manner that may suit the research question, for example, it may not have been collected in the geographical region or timeframe of interest, or categorised in the groups or variables required, for example, age may have been reported as a continuous variable or by pre-determined categories (Boslaugh, 2007).

Authors have argued that certain types of qualitative data may be less suitable for secondary data analysis due to the nature and purpose of the method and the subsequent data generated. For example, participatory-action approaches are aimed at understanding cultural and social arrangements, norms and values, with the purpose of empowering research participants; therefore Irwin (2013) has argued that it is unclear what secondary data analysis would add to this type of approach.

Credibility of the data is also a very important aspect to evaluate. Original datasets need to include an explicit audit trail describing the research and all the stages of data collection, selection and processes to anonymise and store the data. Information should be included regarding how data was collected, by whom, and the

extent of the training and supervision processes (Boslaugh, 2007). Additionally, information is needed about the sample, including sampling decisions, in relation to the research questions and in comparison to the wider population and underpinning theory (Irwin, 2013).

Johnston (2014) has also suggested that another aspect to determine data quality is ensuring that the primary researchers are well-respected in their field and have a good reputation regarding research integrity. Additionally, secondary data analysis is a unique opportunity for primary researchers to consider their own data for further analysis, as they have proximate knowledge of the data and the context in which it was created (Irwin, 2013), which gives them an advantage at the time of carrying out secondary data analysis. Another opportunity to capture and understand the context of the original data is to involve primary researchers through dialogue or as part of the secondary data analysis itself (Irwin, 2013; Johnston, 2014; Rodriguez et al., 2018). Even though determining the quality of the data is a challenge, there are ways to achieve this successfully and make informed decisions about the best datasets available for secondary data analysis.

Accessing datasets is also a challenge. Secondary data researchers may face difficulties accessing the datasets they require, as the most suitable datasets may be hard or expensive to access. Public data is readily available; however, accessing data usually requires documentation and completion of forms where researchers need to be clear and transparent about how they will use the data and how the data will be stored safely. Researchers have to be aware of the requirements and guidelines they need to follow according to the data repository or source from which they are accessing the data.

Legal aspects also need to be considered regarding copyright and intellectual property. Secondary data researchers need to ask for permission to use the original datasets and acknowledge the authors appropriately (Hox and Boeije, 2005). Special thought and consideration needs to be given to the use of public data online. Organisations and countries have specific guidelines regarding the use of data in the public domain and whether and how this data can be used for research purposes (Largan and Morris, 2019). Currently, defining public or private data is still controversial (Largan and Morris, 2019). Further consideration needs to be given to the process of anonymising data considered public and, at a minimum, researchers need to approach the owners of the source they are accessing to ensure they can carry out the research to the end as well as publish and disseminate the findings.

An important issue to consider in qualitative secondary data analysis is the careful storage of data, whether it is provided in hard or soft copies. Specific data repositories may provide specific guidelines of the minimum requirements for data

to be stored safely; this may include password encrypted computers and locked cabinets or rooms. Even if repositories or the original providers of data do not specify these, secondary data researchers have to ensure they comply with their local data protection policies and safeguard their data.

Closely linked with data handling and storage is data transcription. Data transcription has ethical implications also. Data used for secondary data analysis is usually already transcribed. However, if this is not the case, confidentiality protocols need to be in place for data transcribers, particularly if they are external to the organisation carrying out the analysis. Secondary researchers need to ensure transcribers have adequate mechanisms to protect and maintain the privacy of the data at all times and protect the anonymity of research participants (Grossoehme, 2014). As good practice, mechanisms and distress protocols need to be out in place and offered for data transcribers, particularly in cases of sensitive data.

Even though qualitative secondary data analysis has several challenges, these challenges can help researchers guarantee the success of their analysis and the suitability of datasets to answer their research questions fully. These challenges should not discourage researchers from using the method but can instead provide a solid basis from which to start the analysis and ensure that the investment and research will be a success. The next section of this chapter describes in detail all the stages necessary to carry out a qualitative secondary data analysis from research design to dissemination.

RESEARCH DESIGN, EXECUTION, REPORTING AND DISSEMINATION

Several stages of research design are required to complete a qualitative secondary data analysis. There are aspects that the secondary qualitative researcher needs to consider at every stage to ensure research quality and integrity⁶ are achieved. The role and impact of the researcher on the study will also be highlighted as a key component of the research process that requires self-reflection but also encourages the development of awareness, transparency and a further critical approach to the overall research process (Boslaugh, 2007; Dunn et al., 2015; Church, 2001). Even though this section is written in stages, research planning and design is not always a linear process. Planning usually involves simultaneous consideration of different aspects including a literature review, measures available, subject pools, analytical methods, researcher expertise and budget available (Clarke and Cossette, 2000).

Research methods are defined as the practical activities of research (Carter and Little, 2007). These methods are crucial as they determine the final research product; it is, therefore, important to avoid selecting methodologies based on how familiar, fast or easy to implement they are for the researcher, but instead focus on the

research question and how this can be approached in the best possible way (Carter and Little, 2007). Qualitative data collection methods can include observation, interviews, focus groups, texts, official documents and images. Data can be collected face to face and also online (social media, blogs, interviews). Different analytical methods can also generate different accounts of the same data; therefore, selecting a method should be a purposeful and careful decision, based on the research question and objectives that will be answered (Carter and Little, 2007). All components of a research design should be synchronised with each other, starting with the theoretical background that will underpin the research topic, data collection methods, data analysis and ultimately how data is disseminated. The first stage that enables this congruency is the literature review as it provides the theoretical background to support the research and its different components

1. Literature Review

The first stage of any research study is identifying a topic of interest and carrying out the literature review relevant to the topic to determine the current body of knowledge available on this topic at the time, as well as identifying the knowledge gaps that would support the need for further research in the field. The literature review stage can also help identify other researchers in the field, organisations or research centres that have previously conducted research in the field of interest (Johnston, 2014); this can support researchers to make informed decisions about methodologies and types of data that can be used in their own research. As qualitative secondary data implies the use of previously existing data, researchers need to ensure that they are providing an innovative approach to the data either through using innovative methods or generating new knowledge by analysis aspects or perspectives of the data that have really not been known or carried out before. Defining the research question is the crucial next step.

2. Defining the Research Question

The research question is defined from the topic of interest you want to study. Research questions are crucial as they guide the research process (Hancock et al., 2009). Identifying a suitable research question for secondary data analysis can happen in two ways: identify a research question and seek a dataset that would suit, or, alternatively, select a dataset and then formulate a suitable research question to fit the data (Johnston, 2014). These methodologies to identify the research question were described by Cheng and Phillips (2014) as the 'research question driven' approach and the 'data driven' approach. After the topic has been chosen and a suitable question is determined the population of interest will be specified.

3. Specify the Population

Researchers may be interested in a specific group such as children or adolescents, or people of all ages; this will depend on the proposed research question. Additionally, the research may require access to the full population or a sample which may or not be representative in the case of qualitative data.⁷ Datasets may also include a national sample or a group of people confined to a smaller area. For this reason, qualitative secondary data researchers may use complete datasets or parts of them only, if it is appropriate for the research question and research objectives. Researchers need to consider, however, that drawing on subsets of data can have consequences for sample size and representativeness (Clarke and Cossette, 2000). Determining the population of interest is important to identify suitable datasets that have worked with those populations specifically or as part of a wider population. Data for the population of interest must be available and easily identifiable from other populations to ensure the data is pertinent, valid and useful for further analysis with that specific cohort. For example, if a dataset was carried out originally with children in primary and secondary school but the secondary data analysis is only focused in primary school children, the dataset should have presented and stored the data in a way that enables the separation required to identify the age of the child.

4. Topics and Themes Included in the Analysis

Informed by the literature review and their own expertise, secondary researchers need to decide which topics are relevant to answer the research question and the potential characteristics of participants that are relevant to explain the findings. Some of these characteristics may include age, gender, ethnicity, income, education level and geographical location, among others. Datasets need to be revised to ensure this information is included and in a format that is useful and relevant for the secondary data analysis desired so that these topics can be analysed and their potential influence on the findings explored and understood in the analysis. Researchers also need to identify how the data was collected and stored and which theoretical background underpinned the decision making around the storage of data. Qualitative data usually consists of words or texts in the form of interviews, focus groups or any other, or speech in oral or written form, sometimes including creative methods such as art or photography. It would be desirable for anonymised original recordings or transcripts to be available for secondary researchers to make their own decisions about data analysis. Different methods of data analysis have their own style of organising data, for example into themes (e.g. thematic analysis) or perhaps into frequencies and percentages (e.g. content analysis). Based on these theoretical models, researchers should revise how variables are coded and if these need to be and can be recorded for further analysis. Differences in definitions,

classification and coding can impact on the reusability of the data. This last aspect is, therefore, linked to selecting relevant datasets.

5. Selecting Relevant Datasets

Secondary researchers need to decide which datasets are relevant and appropriate to answer their research question. This may include national surveys, hospital records, census data or qualitative datasets such as interview transcripts. Data is also available from official data archives whose purpose is to acquire, archive and disseminate data for secondary analysis.

The selection of datasets can be informed by the literature review to determine which types of methodologies and datasets are used to address the research topics and questions. Data collection methodologies may also have an impact on the quality and type of data collected, for example participants may be more open and honest about sensitive topics over telephone interviews or surveys, rather than face-to-face interviews. All the challenges of determining the appropriate dataset described previously in the chapter are relevant at this stage of the research design.

6. Evaluate the Quality of the Datasets

The quality of the data in the sets, as described in more detail in the previous section, needs to be evaluated in terms of missing data and the themes and topics included. Datasets used for secondary data analysis need to include information about the data collection process, data cleaning, missing or incomplete data, and overall data quality. At this stage, decisions need to be made about whether to continue to use this dataset or whether to find an alternative one that meets the quality necessary to carry out a successful study where the research question can be fully and adequately answered. Original questionnaires, interview scripts and any other text or material used for data collection should be available. Published reports can be useful to understand the original methods and procedures carried out in the primary study (Johnston, 2014).

Another issue to consider is time. Databases may be outdated due to the lapse of time between primary data collection and the proposed secondary analysis. Some datasets may be subject to a bargain period which will delay the time in which data can be released. In specific fields such as information and technology, the time frame of data collection is paramount and should be less than three years (Johnston, 2014) for findings to be valid and relevant. Secondary researchers need to determine whether the timescales and questionnaires used for data collection reflect current thinking in the field or if these are outdated (Clarke and Cossette, 2000). Knowledge of the field or contacting experts in the field will help guide decision making around suitable timeframes to ensure the data is still relevant and worth using to carry out

the secondary analysis (Clarke and Cossette, 2000). The next step is focused on selecting the methodology to carry out the secondary data analysis.

7. Determine the Secondary Data Analysis Methodology

Qualitative research encompasses a variety of methods. This chapter is not intended to provide an in-depth description of qualitative research methodologies; however, Table 9.2 compiles some of the most common qualitative methods that could be used to carry out secondary data analysis (Braun and Clarke, 2006; Elo and Kyngas, 2008; Embuldeniya et al., 2013; Grosseohme, 2014; Hancock et al., 2009; Noblit and Hare, 1988). Details of the methodologies, the usual aims of these methods, recommended sample sizes, and design and analysis considerations are included to provide a general overview of these methods; however, researchers need to become more familiar with the research method of their choice before carrying out the analysis, whether it is primary or secondary analysis.

The analysis of secondary data can be as challenging as primary data. Researchers, depending on their skills, knowledge and resources available to them, must decide the appropriate methodology that will suit their analysis of qualitative data as well as determining if the use of software is appropriate. Software packages available may require a licence and/or permission for its use. These costs need to be included in the research budget at the research design stage.

8. Findings, Conclusions and Dissemination

Selecting a suitable method of data analysis will usually have an impact on the way research findings are structured and presented. As emphasised earlier, all stages of a research design are relevant and should be carefully considered to provide a congruent research study. Researchers need also to be aware of the impact of the research findings on their participants, how they are portrayed in the research findings and conclusions and whether or not the findings will have a positive impact on their lives, such as changing policy and practice towards improving services and participants' lives directly or indirectly.

The research question is crucial again to guide the presentation of the data (Hancock et al., 2009). A good way to present the findings is using the heading and subheadings, categories or themes that emerged during the analytical phase (Hancock et al., 2009). These findings need to be supported and validated by participants or document and text quotations. Some qualitative data, for example content analysis, is instead presented as quantitative data including frequencies and percentages in the form of text, diagrams and or graphs. Authors must be careful to clarify that these figures are not statistically significant (Hancock et al., 2009).

Researchers must be aware of their target audience and how reports and other dissemination materials need to be designed and structured to reach the audiences of interest effectively. This will contribute to increasing the potential impact of the research, whether it is to generate theory, inform policy makers or inform practitioners and service provision. Careful consideration should be given to the levels of literacy and ages of potential readers and populations targeted to ensure the message

Table 9.2 Qualitative research methodologies

| Method | Aim | Sample | Design | Analysis |
|------------------------------|--|--|---|--|
| Ethnography ^a | Understand people and culture | Purposive. People under investigation such as observation, share some diaries, journals. The parameters such as region, religion, periods of time in the field. social and or familiar shared experiences. | Combination of methods researcher spends long as region, religion, periods of time in the field. | The researcher attempts to interpret the data from the perspective of participants. Start the analysis after the first few interviews. Name important words and phrases and determine how they relate to each other to begin developing theory. Participant's voice is broken into fragments and put together. |
| Grounded Theory ^a | No existing hypothesis to test. | Determined by the research question. Theoretical and sampling is used, is participants are selected based on theoretically achieved. Participant's own words relevant constructs. are used as the codes. | Various methods of data collection are such as interviews observations. Data collected until data saturation is achieved. | The researcher should start from a position of knowing nothing about the topic to allow the theory to emerge. Data collection and analysis are simultaneous, emergent theory informs data collection. Categories generated are synthesised into theory. |
| Phenomenology ^a | This is an approach to explore people's everyday life. Search for meaning is a social construct of a particular group of people. | Relatively small sample sizes, detail is more important than generating findings that are applicable across people or situations. | Any design is suitable, depends on the type of data needed to understand a phenomenon and its meaning. | Researcher immerses in the data as a whole, texts are coded, meaning units are categorised using participant's own words and used to create thematic statements to describe the experience. |

| | | | | |
|--------------------------------|---|--|--|---|
| Thematic Analysis ^b | Identify, analyse and report patterns in data. Not bound to a pre-existing theoretical framework. | Smaller samples as analysis is time-consuming. | The essentialist–realist method explores the experiences meaning and the reality of research participants. Constructionist methods explore how social discourses shape events, realities, meaning and experiences. | Familiarisation with the data, generations of initial codes, theme search, theme review, naming themes and producing reports. |
|--------------------------------|---|--|--|---|

| Method | Aim | Sample | Design | Analysis |
|---------------------------------|---|--|--|---|
| Content Analysis ^c | Quantify phenomena by turning words into categories. | Purposive. be used to speech texts and such and used for theory-testing. | Characteristics of language are understood in the context and meaning of the text. Deductive content analysis is as newspapers interviews. | Words are turned into categories to build a conceptual model. |
| | Synthesis of qualitative research findings. | Depends on the inclusion criteria selected; the amount of studies or documents included will respond to that inclusion criteria for example date of issue, topic, population, etc. | Conventional content analysis involves counting the frequency of words, phrases or themes to quantify them. Make a judgement about the current state of knowledge, the quality, scope and comprehensiveness of qualitative reviews. | Reciprocal translation of studies to determine relationships between concepts by keeping the meaning of the original data but generating new conceptual models. |
| Interpretative | Consists of two components, phenomenological to understand how participant make sense of experiences; however, this requires an interpretation from the researcher. | Purposive | The researcher has an insider’s perspective. | Data is coded for emergent themes and to find connections to identify higher order themes |
| Discourse Analysis ^f | Focused on text and talk as social practices. | Study of any language in use; this may include speech, newspapers, policy documents etc. | Researchers are focused on discourses organised to persuade or present a particular view of the world. | Linguistic repertoire: search for patterns in the words that are used. Rhetorical strategies: explore the way words are used. |

Sources: (a) Grossoehme (2014); (b) Braun and Clarke (2006); (c) Elo and Kyngas (2008); (d) Noblit and Hare (1988); (e) Smith et al. (2009); (f) Largan and Morris (2019).

delivered by the research is understood and acted upon if that was the intended outcome.

Researchers can also think of creative methodologies to disseminate their findings (Ben, 2017; Lapum et al., 2014). Research reports and academic papers are the most traditional methods to disseminate research; however videos, podcasts, radio interviews, newspaper articles and other creative methodologies such as art and theatre have also been used for research dissemination particularly targeted at non-academic populations and the findings presented back to research participants.

As described in this section, qualitative secondary analysis usually follows the same stages as a primary analysis. Probably the only obvious difference is that primary research usually includes data collection. Secondary data analysis works with pre-existing data collected, but it still involves a thorough process of identifying, judging and selecting the most suitable dataset. This will still require knowledge and a critical approach from the researchers, although it may take less time and cost to complete the task, but it is the most crucial task of qualitative secondary data analysis. A suitable dataset will ensure a good secondary data analysis that is ethical, reliable and capable of fully answering the proposed research question.

QUALITATIVE SECONDARY ANALYSIS: CASE STUDY

As an example of a qualitative secondary data analysis, this section describes a real research study that was carried out in 2018 to evaluate the impact of mentoring relationships on adolescent empathy (Rodríguez et al., 2018). The purpose of the case study is to exemplify some of the methodological considerations described in the previous sections. The case study is intended as an example only as all decisions in qualitative secondary data analysis will depend on each individual study and informed by specific research questions.

This qualitative secondary data analysis was funded by the Irish Children's Research Network as part of their Prevention and Early Research Initiative (PEI). The aim of this initiative was to archive datasets from research studies focused on prevention and early intervention with children and young people. Data archives would be subsequently available for further analysis and service development in Ireland.

As this secondary data analysis was part of the PEI Initiative, datasets were identified by data curators that were working for the Irish Children's Research Network. The researcher, therefore, had access to a list of potential datasets that could be used for secondary data analysis. Based on her own research interest, the current database was selected first and subsequently the researcher designed a research question that would fit both the topic of interest and the database.

This secondary data analysis was justified as empathy and had not been explored in the original study. The secondary analysis built on the first analysis and provided additional evidence of how to improve mentoring programmes, maximise the benefits for young people and expand on the body of knowledge to inform policy and improve practice in the field of mentoring and empathy (Rodriguez, 2018).

1. Literature Review

A thorough literature review was carried out to explore the current knowledge base and theories relevant to the study. The main concepts explored were mentoring and empathy as well as the existing knowledge regarding the relationships between these two concepts. Other concepts included as relevant were the Big Brothers Big Sisters mentoring programme and research dissemination with young people.

Due to the vast amount of definitions identified in the literature review for the term 'empathy', one of the first challenges was to identify a suitable theoretical framework and definition to underpin this study. Combining and critically analysing the various definitions of empathy, the authors operationalised empathy in the way that it would be defined and included in this secondary data analysis. For the purposes of this study empathy was defined as passive and active empathy. Passive empathy was defined as the ability to sympathise with the emotional states of others, but not taking any action to relieve the person in distress. Active empathy refers to affective reactions and actions towards the person in distress or need (Rodriguez et al., 2018). Clearly defining a study variable was crucial to carry out the secondary analysis, as this was the underpinning anchor that guided the subsequent methodology and analysis.

The literature review also provided evidence of the need for further research in the field, as research focused on mentoring in youth empathy was limited. This was a crucial element as secondary data analysis need to be carefully justified to ensure these are adding to the body of knowledge and it is not a repetition of previous studies and a waste of time and resources.

2. Defining the Research Question

The qualitative secondary data analysis was focused on achieving one of the research objectives, specifically to explore the role of empathy in the experiences of youth involved in mentoring relationships. The research question was defined as a result of the research interest of the authors but also the suitability of the dataset to explore and successfully approach this research question. In secondary data analysis the topic and underlying objectives can emerge from a variety of contexts, real life issues or even personal interests. The only crucial aspect is to ensure that the dataset

suitable to answer the research question thoroughly exists and is available to carry out a secondary data analysis, otherwise primary data collection may be needed.

3. Specify the Population

The original study included young people as the original sample; for this reason the secondary data analysis was adapted to include the age range in the dataset, and therefore there was guaranteed data availability. An innovative aspect of this secondary data analysis was the inclusion of young people in the validation of results and research dissemination, therefore the secondary data analysis at all stages was carried out and adapted to fit the age of the population available. This contributed to building the consistency and congruency at all stages of the study as adolescents themselves were actively engaged in results and dissemination in a way that was age-appropriate and relevant to them. The possible limitation of this is that it can only be claimed that the findings are pertinent to this age group and further research will be needed to target other age groups and populations to understand how mentoring can impact on empathy development.

4. Topics Included in the Analysis

The secondary data analysis described in this case study was focused on exploring the role of mentoring relationships in adolescent empathy, therefore the variables of interest explored in the analysis were: mentoring relationship, passive empathy and active empathy. The research team had to read thoroughly and carefully determine if the interviews had approached these topics in a level of detail that could provide useful insights into answering the research question. Judging the original data, however, can be subjective as there are no guidelines to determine how much information about a topic is 'enough'. A possible criterion used for this can be data saturation; however, this may not become obvious at an early stage of data analysis. Saturation may not be achieved until full or advanced stages of analysis are complete and several transcripts have been analysed. Previous experience in primary research can facilitate this decision as researchers may have the knowledge of past experiences of when 'enough' information was achieved; however, this remains a limitation of secondary data analysis particularly for novice researchers.

5. Selecting Relevant Datasets

Data for this secondary data analysis originated from the Big Brothers Big Sisters (BBBS) of Ireland: Evaluation Study undertaken by the UNESCO Child and Family Research Centre, NUI Galway (Dolan et al., 2010). This was a mixed methods evaluation consisting of a quantitative randomised control trial and nine longitudinal qualitative case studies of mentoring relationships including the views of young people, mentors, mothers and case workers.

The sample for the original study was purposeful; however, it included a balance between young people's gender and between urban and rural areas. Both variables could act as confounding variables and, therefore, these variables were included and considered in the analysis. Half of the sample came from single-parent households and this variable was also considered in the analysis. The issues and needs of young people involved in mentoring were mostly economic disadvantage, poor social skills, shyness, and withdrawal. Mentors (n = 73) ranged between 18 and 55 years, and 55 per cent of them were female. The majority had third-level education (80 per cent) and were in full-time employment (70 per cent). This shows a methodological strength and a careful awareness of the potential confounding variables that could impact on the findings. These variables were explored in the analysis to identify the potential impact or influence of these characteristics in empathy and mentoring relationships.

A critical analysis of the database was carried out to identify the strengths but also the limitations of the original data and determine whether these limitations could be resolved to the extent that the dataset would be adequate for further analysis.

The strengths of these databases were:

- Data was complete, and all nine pairs of transcribed interview were available for further analysis. Therefore, all case studies were available and it was suitable for further analysis.
- The original purpose of these case studies was to explore the mentoring relationship; therefore, this was a very similar subject matter as this secondary data analysis intended.
- The original data collection was carried out with semi-structured interviews. As Irwin (2013) suggested, semi-structured interviews produce data that is more independent of the context and the primary researchers, suggesting that the data generation process was transparent and, therefore, more suitable for secondary data analysis.

Limitations of this database were:

- The topic of empathy was not specifically approached in the original study, therefore the topic was not going to be explicitly found in the data or in the level of depth in which it could have been found if the topic had been included in the original interview scripts.

6. Evaluate the Quality of the Datasets

The original dataset was generated by the UNESCO Child and Family Research Centre, NUI Galway, which has a long-standing history of research and evaluation and a strong reputation of high-quality research and research integrity. These

considerations provided an initial level of security to further explore the quality of the data.

Another consideration which provided support to determine the quality of the dataset was the availability of published reports describing thoroughly the research process, methods, analysis and findings. Having this information provided the main researcher with enough detail about the data and the potential and limitations of this data to make an informed decision about its suitability for a secondary data analysis. Another aspect considered in determining the quality of the data was the involvement of two of the original authors, which was a great support in understanding the context of the original data, its collection and analysis.

Regarding ethical considerations, one significant and important advantage of this secondary data analysis was the inclusion of two of the original researchers. Original consent forms did not suggest the use of data for archiving in the future and the academic institution where this study took place required these consents to carry out secondary qualitative and quantitative research. However, research participants had given their original consent to the original authors to use their data for research purposes. Their involvement was crucial to comply with basic consent guidelines and procedures.

A strength of this database specifically supporting the validation of the dataset was data triangulation.⁸ Case studies included the views of different informants: young people, mentors, mothers and case workers. All of this contributed to compare the quality of mentoring relationships from different points of view and added a layer of validation to the findings. For all these reasons, it was deemed that the dataset was of high quality and suitable to be used in a secondary data analysis. Secondary data analysts need to be explicit in their reports about the criteria used to determine data quality and the reasons that justified and informed the selection of a specific dataset for their analysis.

7. Determine the Secondary Data Analysis Methodology

Based on the definition of empathy selected for this secondary data analysis, references to passive and active empathy were going to be identified and quantified in the interviews. This required a methodology capable of quantifying examples of passive and active empathy in the transcripts. Based on this systematic quantification of passive and active empathy, content analysis (Elo and Kyngas, 2008) was the method selected to explore the role of mentoring in adolescent empathy due to the capacity of this method to systematically evaluate patterns and quantify the frequency of words and concepts. A description of this method is included in Table 9.2 above. Secondary data analysts also need to be explicit about the methodological decisions made and the reasons why a specific method is

chosen. In this case, the underpinning theory was the main influence in the selection of the method, in line as well with the research question.

8. Findings and Conclusions and Dissemination

Findings and conclusions of this secondary data analysis were disseminated in a final report that was published as a hard copy (Rodriguez et al., 2018). This report also had a digital version that was shared in the NUI Galway, UNESCO Child and Family Research Centre and the Children's Research Network websites. As a result of this secondary data analysis a research paper focused on the ethical implications of secondary data analysis with young people was also published (Rodriguez, 2018). These dissemination methods are standard in the institution where the research team was based, therefore, these methodologies of dissemination were expected. This does not mean, however, that these methods were the only ones possible; other options were also allowed. This enabled the research team to be creative, although more time and costs had to be invested as several methodologies of dissemination were used. These costs were foreseen from the budget proposal; however, available budget and time can also restrict the type and amount of dissemination activities possible.

One of the strengths of the presentation of the findings of this qualitative secondary analysis was the inclusion of a research summary and key findings that are effective formats to share research findings quickly and effectively with targeted groups. These may be printed, posted or made available through newsletters or relevant sources online. The main objective of these types of data dissemination methods is to generate impact quickly, even though the details of the study are compromised.

The report was also designed to target relevant populations such as practitioners and policy makers. Implications of the findings for these groups were explicit and easy to find. To achieve this, a separate section explicitly targeted at practitioners and policy makers was included in the report, clearly marked and included in the index. Including this section contributed to the clarity of the message that was delivered and targeted the message to increase the possibility of creating an impact. The only risk of including these summaries would be that people do not read the full report and the context of the findings can be compromised.

This qualitative secondary data analysis also had a significant dissemination activity which consisted of the design and recording of a video. The creation of this video was included as one of the research objectives in the design stage and the cost of this video was calculated and included in the research proposal. Videos are effective ways of sharing research findings with researchers and policy makers but also with the lay public who may have very limited access and interest in academic publications. The video was also a suitable and interesting way to target young

people, not only to share the findings but because they were interested in the creative process, and dissemination was youth-led. The limitation of this is evidently that the scientific underpinning and research design are not included; therefore, the origin of the findings is not shared and can be decontextualised. This is relevant for non-academic circles.

The youth-led dissemination process was carried out in two phases: (a) consultation with the youth advisory group; and (b) youth-led production and filming. The youth advisory group was created with the support of young people currently involved in the Big Brothers Big Sisters Programme in the same age range as the sample included in the secondary data analysis. New information sheets and consent forms were provided for young people and their parents before they took part in the Advisory group. Young people met the video producer and lead researcher to discuss the research findings and ensure that they were relevant and pertinent to their own mentoring experiences. The purpose of this was to adapt the findings to current service users to ensure that findings were relevant and used for the 'public good' (Bishop, 2009; Rodriguez, 2018). Following this consultation process, young people provided ideas on how to turn the findings into a script and story-line for the video. A second session was used to agree the final version of the script and film the video. Young people were given a choice to be actors or to carry out other production roles such as props, using the cameras and recording if they wished to do so. Post-production and edits were carried out by the professional video producer hired for this purpose. The video was uploaded to websites and relevant social media to maximise dissemination of the findings.

CONCLUSION

This chapter has provided a detailed introduction to qualitative secondary data analysis. Qualitative secondary data analysis is a recent research methodology which has huge potential to support and advance the creation of knowledge in different fields, topics and populations. As with any other research method, it has benefits and limitations that need to be carefully considered before deciding if this method is appropriate. Even though it is a relatively flexible method, it may not be suitable to respond to specific research questions, or datasets required to carry out this type of secondary analysis may not be available or may have very low quality. Determining the quality and suitability of datasets remains the main challenge for any researcher thinking of carrying out secondary data analysis and this is the crucial stage to determine whether or not the research is worth carrying out.

Qualitative secondary data analysis is a structured method that will only be successful within a congruent research design from the research topic and question to all stages of research analysis and dissemination. This chapter has emphasised the

need for all research parts to form a synergy that is supported by and within the qualitative research paradigm. From that basis, researchers can make their choices based on their knowledge, skills, creativity and personal interests. Secondary data analysts need to be explicit about the decisions they made at every stage of the analysis, the same as is expected for primary researchers to ensure research validity and reliability.

Although this chapter provided an example from practice on how to determine the quality and suitability of a dataset, some aspects such as determining how much data is 'enough' to answer a research question are still subjective decisions that should be informed by evidence and practice in the future, and the literature still lacks clear guidelines around this crucial topic. Although qualitative secondary analysis has great potential, it continues to be a developing field that would benefit from further experiences from practice, space for further reflection, understanding and improvement for the future.

NOTES

1. Qualitative secondary data analysis should not be confused with qualitative meta-analysis. The latter has been described as 'the aggregating of a group of studies for the purposes of discovering the essential elements and translating the results into an end product that transforms the original results into a new conceptualization' (Schreiber et al., 1997, p. 314). Qualitative secondary data analysis can be carried out on a single dataset, excluding comparisons from other datasets or research studies.
2. Participants have the relevant characteristic to answer the research question. Time, resources and research objectives usually inform the sample size as well.
3. Type of purposive sampling which specifies how many people and with which characteristics are included in the sample. These characteristics can include age, gender, marital status, location, etc.
4. Research participants already included in the study use their social networks to help identify potential research participants.
5. Research participants have a particular profile that responds to the typical characteristics of a particular issue or phenomenon of interest for the study.
6. Research integrity refers to the ethical conduct of scientists around the world. The Singapore Statement on Research includes four principles of ethical conduct in research which are honesty, accountability, professionalism and stewardship (Resnik and Shamoo, 2011). Ensuring research integrity entails setting standards for robust and ethical conduct and reporting of research (Bhui et al., 2019).
7. A representative sample is a subset of a population that accurately represents the wider group being examined. This is usually calculated by using statistical methodologies.
8. Triangulation consists of the use of multiple methods and data sources on informants to develop a comprehensive understanding of specific phenomena. This method is also used to validate qualitative research findings through the comparison of different data sources.

REFERENCES

- Ben, B. (2017) 'Enclothed knowledge: The fashion show as a method of dissemination in arts-informed research', *Qualitative Social Research*, **18**(3), 2–24.
- Bhui, K., Lee, W., Kaufman, K. and Lawrie, S. (2019) 'Ensuring research integrity: Setting standards for robust and ethical conduct and reporting of research', *The British Journal of Psychiatry*, **215**, 381–2.
- Bishop, L. (2009) 'Ethical sharing and reuse of qualitative data', *Australian Journal of Social Issues*, **44**(3), 255–75.
- Boslaugh, Sarah (2007) *Secondary Data Sources for Public Health*. Cambridge: Cambridge University Press.
- Braun, V. and Clarke, V. (2006) 'Using thematic analysis in psychology', *Qualitative Research and Psychology*, **3**, 77–101.
- Breznau, N. (2016) 'Secondary observer effects: Idiosyncratic errors in small-N secondary data analysis', *International Journal of Social Research Methodology*, **19**(3), 301–318.
- Carter, S. and Little, M. (2007) 'Justifying knowledge, justifying method, taking action: Epistemologies, methodologies and methods in qualitative research', *Qualitative Health Research*, **17**(10), 1316–28.
- Cheng, H. and Philips, M. (2014) 'Secondary analysis of existing data: Opportunities and implementation', *Shanghai Archives of Psychiatry*, **26**(6), 371–5.
- Chesebro, J. and Borisoff, D. (2007) 'What makes qualitative research qualitative?', *Qualitative Research Reports in Communication*, **8**(1), 3–14.
- Church, R. (2001) 'The effective use of secondary data', *Learning and Motivation*, **33**, 32–45.
- Clarke, S. and Cossette, S. (2000) 'Secondary analysis: Theoretical, methodological and practical considerations', *Canadian Journal of Nursing Research*, **32**(3), 109–29.
- Dolan, P., Brady, B., O'Regan, C., Canavan, J. and Brumovska, T. (2010) *Big Brothers Big Sisters of Ireland Evaluation Study, Report 2*. Galway: Child and Family Research Centre. National University of Ireland Galway.
- Dunn, S., Arslanian-Engoren, C., DeKoekkoek, T., Jadack, R. and Scott, L. (2015) 'Secondary data analysis as an efficient and effective approach to nursing research', *Western Journal of Nursing Research*, **37**(10), 1295–307.
- Elo, S. and Kyngas, H. (2008) 'The qualitative content analysis process', *Journal of Advanced Nursing*, **62**(1), 107–15.
- Embuldeniya, G., Veinot, P., Bell, E., Bell, M., Nyhof-Young, J., Sale, J. and Britten, N. (2013) 'The experience and impact of chronic disease peer support interventions: A qualitative synthesis', *Patient Education and Counselling*, **92**, 3–12.
- Grossoehme, D. (2014) 'Overview of qualitative research', *Journal of Health Care Chaplaincy*, **20**(3), 109–22.
- Hancock, Beverley, Ockleford, Elizabeth and Windridge, Kate (2009) *An Introduction to Qualitative Research*. National Institute for Health Research for the East Midlands/ Yorkshire and the Humber.
- Hox, J. and Boeijs, H. (2005) 'Data collection, primary vs. secondary', *Encyclopaedia of Social Measurement*, **1**, 593–9.
- Irwin, S. (2013) 'Qualitative secondary data analysis: Ethics, epistemology and context', *Progress in Development Studies*, **13**(4), 295–306.
- Johnston, M. (2014) Secondary data analysis: A method of which the time has come. *Qualitative and Quantitative Methods in Libraries*, **QQML**(3), 619–26.
- Lapum, J., Liu, L., Church, K., Yau, T., Ruttonsha, P., Matthews, A. and Retta, B. (2014) 'Arts-informed research dissemination in the health sciences: An evaluation of people's responses to "The 7.024th Patient" art installation', *SAGE Open*, 1–14.
- Largan, Claire and Morris, Theresa (2019) *Qualitative Secondary Research: A step-by-step guide*. London: Sage Publications.

- Mack, Natasha, Woodsong, Cynthia, MacQueen, Kathleen, Guest, Greg and Namey, Emily (2005) *Qualitative Research Methods: A Collector's Field Guide*. United States: Family Health International.
- Noblit, George and Hare, Dwight (1988) *Meta-Ethnography: Synthesizing Qualitative Studies*. London: Sage Publications.
- Resnik, D. and Shamo, A. (2011) 'The Singapore statement on research integrity', *Accountability in Research*, **18**(2), 71–5.
- Rodriguez, L. (2018) 'Secondary data analysis with young people: Some ethical and methodological considerations from practice', *Children's Research Digest*, **4**(3).
- Rodriguez, L., Dolan, P. and Brady, B. (2018) 'Exploring the impact of mentoring relationships in adolescent empathy: A mixed methods approach'. Galway: UNESCO Child and Family Research Centre, National University of Ireland Galway.
- Ruggiano, N. and Perry, T. (2019) 'Conducting secondary analysis of qualitative data: Should we, can we, and how?', *Qualitative Social Work*, **18**(1), 81–97.
- Schreiber, R., Crooks, D. and Stern, P.N. (1997) 'Qualitative meta-analysis'. In J.M. Morse (ed.), *Completing a Qualitative Project: Details and Dialogue* (pp. 311–26). Thousand Oaks, CA: Sage Publications.
- Smith, E. (2008) 'Pitfalls and promises: The use of secondary data analysis in educational research', *British Journal of Educational Studies*, **56**(3), 323–9.
- Smith, J., Flowers, P. and Larkin, M. (2009) *Interpretative Phenomenological Analysis. Theory, Method and Research*. London: Sage Publications.
- Trinh, Q. (2018) 'Understanding the impact and challenges of secondary data analysis', *Urologic Oncology*, **36**, 163–4.