Sampling Strategies in Qualitative Research

In: The SAGE Handbook of Qualitative Data Analysis

By: Tim Rapley
Edited by: Uwe Flick
Pub. Date: 2013
Access Date: August 18, 2021
Publishing Company: SAGE Publications Ltd
City: London
Print ISBN: 9781446208984
Online ISBN: 9781446282243
DOI: https://dx.doi.org/10.4135/9781446282243
Print pages: 49-63

Sampling Strategies in Qualitative Research

TimRapley

Put simply, sampling really matters. It matters in relation to an array of issues, for the whole trajectory of the analytic process, from initial questions asked about a phenomenon to the presentation of your work. Given that the claims that qualitative researchers want to make are routinely based on working closely with relatively small numbers of people, interactions, situations or spaces, it is central that these are chosen for good analytic reasons. Above all, sampling should never be the product of ad hoc decisions or left solely to chance. It needs to be thoughtful and rigorous.

There are some good discussions of the range of key conceptual issues about sampling (see e.g. Guba, 1981; Mitchell, 1983; Ward Schofield, 1993; Sandelowski, 1995; Williams, 2002; Gobo, 2004) alongside some useful, user-friendly, introductions to more practical considerations (see e.g. Patton, 2002; Charmaz, 2006). In this chapter, I want to explore sampling through a different narrative, one which uses a single case to demonstrate a range of issues researchers face in relation to sampling. I am going to offer a reasonably detailed account – although sadly far too brief – of a research project I undertook, in order to explore some of the pragmatic and theoretical issues you can face. Initially, I will explore issues of sampling prior to entering the field, both in relation to proposal writing alongside the forms of knowledge that can inform your ideas. I will then explore the evolution of the sampling practices over the life of the project – from exploratory rounds of sampling to those more focused on conceptual development – always outlining the iterative relationship between sampling and analysis. Finally, I will turn to sampling in relation to the presentation of data. Interwoven throughout this account will be some traditional overviews of the key debates and procedural issues that you need to consider. However, first, I will introduce the research project.

The Context of the Case

The case I want to explore focuses on delay in diagnosis for children with juvenile idiopathic arthritis (JIA). JIA is a form of arthritis that affects both children and adolescents. As soon as the diagnosis is suspected, these patients need to be referred to a paediatric rheumatology team, to get confirmation of the diagnosis and to get access to the effective treatments now available. The research team included a consultant in paediatric rheumatology, with whom I had worked closely on other projects. I need to stress that this was an extremely practically orientated project, funded by a research charity, Arthritis Research UK[1], that was seeking practically orientated findings.

Note: Genres of Sampling and the Generalizability Question

Sampling can be divided in a number of different ways. At a basic level, with the exception of total population sampling you will often see the divide between random sampling of a representative population and non-random sampling. Clearly, for many more quantitative-minded researchers, non-random sampling is the second-choice approach as it creates potential issues of 'bias'. However, in qualitative research the central resource through which sampling decisions are made is a focus on specific people, situations or sites because they offer a specific – 'biased' or 'information-rich' – perspective (Patton, 2002). Irrespective of the approach, sampling requires prior knowledge of the phenomenon. Knowledge is essential in order to establish how 'typical' your sample is of the phenomenon alongside understanding the potential diversity, or variance, within the phenomenon. The higher the variance, the larger the sample required.

Within more quantitative work, when working with a random sample you need to be able to classify the population in order to generate a representative random or quota sample. That assumes various things. In survey work, you need to have enough a priori information to inform the design of the sample. Routinely, you would work with some kind of proxy for the issues that you are interested in, often based on socio-demographic data. In more experimental work, like randomized controlled trials, you conduct some prior research in order to establish the variance in the phenomenon, as documented by some outcome measure, in order to undertake sample size calculations to detect significant differences. However, routinely within research, especially social science, the focus is on issues – like actions, interactions, identities, events – where we do not have sufficient knowledge of the distribution of phenomena in order adequately to inform sampling issues.

When not sampling the total population, random sampling relies on large samples and attempts to minimize sample errors. You can then begin to claim statistical representatives. As Gobo notes:

> There is no evidence that the sampling assumptions underlying the natural sciences (i.e. that cases are interchangeable because they are *equal* and *distributed at random* in the population) work well in the social sciences. On the contrary, in society almost nothing is random, there are social inequalities affecting people's position in the population. (2004: 441; italics in original).

So, notwithstanding the problems of adequately understanding the distribution of a phenomenon to inform sample design and size, you cannot assume a random distribution.

Now, the logic behind this is that a representative, and ideally random, sample will mean that the findings are generalizable (in a statistical sense). As Gobo (2004) highlights, too

often these two terms are used interchangeably, without reflection on what separates them. Representativeness connects to the questions about the sample whereas generalizability connects to questions related to the findings. Working with a representative sample does not automatically lead to generalizable findings; between these two issues are potential 'measurement errors', connected to a wide array of practical problems. Relatedly, working with a non-representative sample does not mean you can automatically assume that generalizability is not possible. For example, work within the tradition of conversation analysis has repeatedly demonstrated interactional practices, such as preference organization, that are routinely used across a wide variety of domains of everyday and institutional talk. In this way, the findings have theoretical generalizability (see Maxwell and Chmiel, Chapter 37, this volume).

Public Accounts of Sampling Strategies

Prior to undertaking the research, we had to offer the funding agency and the medical ethics committee an outline of our sampling strategies. Given that these audiences expect a reasonable degree of certainty and structure, the approach to sampling was simply introduced as 'purposive sampling' and later described as 'maximum variation sampling', with the variation defined by:

- duration of delay from onset of symptoms to first assessment by paediatric rheumatology;
- observed complexity of referral pathways (assessed by the known number of contacts with health and social care professionals).

Note that, even at this stage, the sampling strategy is not defined in terms of socio-demographic solutions; rather the strategy is led by the phenomena of delay and complexity. Given the focus of the research, this makes practical sense. Also, terms like 'maximum variation strategy' can have a currency with these types of mixed audiences. Certain key phrases have become part of the grammar of applications, in much the same way that terms like 'grounded theory' are found in data analysis sections.

We also outlined that we would focus on two patient groups. We proposed to sample new referrals to the paediatric rheumatology team, recruiting between 10 and 15 patients diagnosed within the prior nine months to minimize problems of recall. We also aimed to recruit between 10 and 15 more established patients, those who had been with the team for over nine months, in order test the emerging ideas. As we were working with children and adolescents, each 'case' would be understood through talking to families – so parents, guardians and, if they wanted to take part, the patients themselves. We would also interview the health and social care professionals either responsible for or involved in the referral to the paediatric rheumatology team – so our understanding of each case could be expanded to include these actors. We offered some sense of certainty around the numbers we would recruit, albeit offering the numbers within a range. This had practical value, in terms of offering the readers a sense of the work that the project would involve – so they could establish how

plausible the project appeared given the time frame and resources requested.

Making Sense of the Phenomena

Centrally, I already had access to four forms of data, which could assist in understanding the potential variation in the phenomenon:

- A review of 152 patients' case notes had already been conducted and published (Foster et al., 2007). This outlined that over 75% of patients exceeded 10 weeks from onset of symptoms to first paediatric rheumatology assessment. The median interval was 20 weeks and ranged from less than 1 week to 416 weeks (eight years!).
- I also had access to a Masters student's data set that updated the previous published review, and covered over 200 patients. Again, this outlined a similar range of delay.
- I had read some of the (limited number of) papers on delay in diagnosis, which outline some of the factors tied to delay. So, for example, in rheumatoid arthritis in adults, the central issue was family doctors not recognizing the patient's problems as disease related.
- Finally, as I had already worked with members of the paediatric rheumatology team, I had access to them. I was able to discuss their impressions of the range of issues faced. I was often confronted by one type of narrative, an atrocity story, where a child had been subjected to extensive delay through the incompetence of a range of medical practitioners who were, for whatever reason, unable to see the child's problem as arthritis related.

In this way, I could begin to gain a sense of some of the issues I *might* want to focus on over the life of the project. None of this offered a firm direction as to where to go next. And, over the life of the project, during rounds of sampling, I would return to these sources of information to inform my analysis.

Note: Qualitative Approaches to Generalizability

Qualitative research has recently grown in popularity and shifted in focus beyond documenting the unique and particular, in part due to funding from evaluation and policy-orientated sources. In this context, considerations about sampling, alongside considerable debate and discussion, have become more central (Ward Schofield, 1993). As Dingwall notes:

> The one-off case study, conceived and executed in magnificent isolation, has no place in modern social science and little more than anecdotal value to a policy maker trying to understand how an organisation works. (1992: 171)

In this context, in part as a reaction against the positioning of qualitative research as less vital and relevant given its refusal to undertake random sampling with large numbers – due to a fundamental asymmetry in goals (e.g. Lincoln and Guba, 1985) and inability in practical terms, given time, resources and funding (e.g. Hammersley, 1992) – alternative understandings have emerged. Various authors have argued, to various degrees of success, that qualitative research is bounded by different epistemological and ontological orders. As such, alternatives have emerged, for example:

> For the naturalist, then, the concept analogous to generalizability (or external validity) is transferability, which is itself dependent upon the degree of similarity (fittingness) between two contexts. The naturalist does not attempt to form generalizations that will hold in all times and in all places, but to form working hypotheses that may be transferred from one context to another depending upon the degree of "fit" between the contexts. (Guba, 1981: 81)

And in this situation, given adequate information about the context, it is for the reader to make the connections to other similar contexts, to judge the 'degree of "fit"'.

Alongside Guba's 'transferability', we have such concepts as 'analytical generalization' (Yin, 1994) 'moderate generalization' (Williams, 2002) and 'empirical generalizations' (Hammersley, 1992), among others. Hammersley (1992) argues that you need to establish that the people or settings are in some way 'typical' of the population to which you want to generalize. He suggests establishing this through reference to published statistics, embedding qualitative research within or alongside survey research, or working with multiple cases, in terms of either people or sites, and exploring the variance. In this way, empirical generalization is possible when the case, or cases, are in some way demonstrated as representative of the population. The case can only be generalized to defined settings over a defined period of time and, for Hammersley, it is for the author to define these other similar contexts. He contrasts this with what he refers to as 'theoretical inference', inference to a class of people, situations or sites in any setting or time. In this way, a case's adequacy is its ability to generate formal theories – with hypotheses, theoretical propositions, logical inferences or casual connections – that can be tested and verified in further empirical work in the same class of people, situations or sites (see Mitchell, 1983, and Silverman's, 1985, discussion of these ideas). As such, atypical or particularly interesting single cases would be ideal places to sample, as they would offer a rich space to generate and test theoretical principles (see Maxwell and Chmiel, Chapter 37, this volume).

An Initial Round of Sampling (*n* = 3)

I then engaged in a very exploratory round of sampling. I asked the team to suggest three different families of patients. I wanted to speak to people from three areas of referral: one that was fairly rapid, so under 10 weeks; one that was typical, so about 20 weeks; and one that was over a year. Following Patton, I saw this sample as 'illustrative not definitive' (2002: 236) – as a way to begin to explore the phenomenon. He notes that, 'It is important, when using this strategy, to attempt to get broad consensus about which cases are typical – and what criteria are being used to define typicality' (2002: 236). As I discovered, this was not a simple process. Below are extracts from an edited field note I wrote after meeting with some members of the team:

> *The role of the team secretary, as part of distributed knowledge/memory of the team, is key. The secretary and one of the nurses looked through a list to offer a selection of about 10 patients.*
>
> …
>
> *They had a key question – what is quick, routine and long? 24 weeks of history is 'long' for them, but 'routine' as far as prior research shows. Is 'quick' a fast diagnosis, via Accident & Emergency, or within the official target of ten weeks? Also, the nurse's caseload was tied to her vision. She deals with more complicated cases. They both thought about route to diagnosis and where interested in finding referral from an unusual source (like Plastic surgery or ophthalmology) – so they were thinking in terms of 'untypical' cases?*
>
> …
>
> *This was NOT an easy or smooth process. Original list was questioned and modified by consultant and then we returned to some of those on the original list!*

They finally agreed on six names of patients and three families agreed to be interviewed. Although the process was illustrative of a range of issues, I just want to focus on a few. Generating consensus on something that is a 'typical' case involved extensive discussion. The discussion itself was illuminating, highlighting taken-for-granted aspects of individual and team reasoning about how they categorize cases. Through this process, my understanding of 'typicality' was questioned and extended. So, rather than just focusing on issues of typical delay in relation to time, typicality should also include the route the patient took.

In conducting and then analysing these interviews, I discovered something interesting. In talking to these parents I got slightly different accounts from that presented in the patient's notes. For example, what the team categorized as a typical case of a 'quick' referral emerged as a more complicated process that lasted about 11 months. This might stem from parents' re-evaluation of prior symptoms. At the time of diagnosis they may not have told staff about the onset of some symptoms as they felt they had little to do with the illness, but, with growing knowledge about the disease, they now understood them as first signs of onset.

Also, the case that was typical of ‘long delay’, over one year, was, at this point in the project, an ‘atypical’ or ‘deviant case’. The child had received a diagnosis of JIA at about 3 years old and the parents were told he was too young for further tests and he was given a short course of physiotherapy. After this, they where told he was fine and discharged. He did not complain of any problem for another five years and then was referred straight to the team. In this way, he received a rapid diagnosis but inappropriate care.

Note: Purposive Sampling Strategies

If you look at the literature on sampling, you can soon be overwhelmed by the diversity of approaches people write about. So, for example, Sandelowski (1995) refers to three approaches – maximum variation, phenomenal variation and theoretical variation – all described as purposeful.[2] Gobo (2004) refers to four: purposive, quota, emblematic and snowball. Patton (2002) refers to 16 different types – including critical case, stratified purposeful, snowball and convenience – all again described under the label purposeful.

Personally, I find Patton's list very useful to think with. He presents you with 16 different labels to work with, to think about, and this is incredibly useful as a way to sensitize your sampling strategy. It enables you to realize that you have choices, that you should be making choices and that those choices can have an impact. However, the issue is not that you have been able initially to sample five ‘typical cases’ of rapid referral, but rather that you have got five cases and you have thought through issues of how typical are they, what connects them, what divides them. As Sandelowski notes:

> These determinations are never absolute; depending on the purpose, analytic frame, and phase of an analysis, any one case can be a case of and about more than one thing and can, therefore, be analytically (re)located among other cases. (1996: 527)

So being able to call a case ‘typical’ is useful. Initially, you might know from some other source, say statistical data, the funder, colleagues or even other respondents, that a specific site is ‘typical’. However, you need to question such a position – it might be ‘typical’ in the way that others have understood the issue, but your research might render the phenomenon in a different way

Thinking about and categorizing your sampling strategies does not always occur prospectively or over different rounds of sampling. For example, Draucker et al. (2007), after an initial recruitment flyer, discovered they had 110 calls from people interested in taking part in their study. Given the nature of the focus of the study, people's experiences

of sexual violence, they felt they had to interview those 43 who met the criteria sooner rather than later. They undertook an initial round of coding of 43 interviews, and developed initial codes and concepts. Rather than conduct more interviews, they re-explored their own data set, searching within this, initially for 'intense' cases, so undertaking a form of intensity sampling. Intensity sampling refers to 'excellent or rich examples of the phenomenon of interest, but not highly unusual cases' (Patton, 2002: 234). They looked again at their data set through various sampling approaches, and in one area, when conducting 'extreme or deviant case sampling', re-interviewed one of the participants.

In some senses, the reality is a lot simpler than thinking about which of Patton's 16 labels fit. It is enough to make good, analytically driven, thoughtful, decisions. Poor sampling decisions, those driven by lack of access, response, knowledge, time or resources can lead to sampling driven by opportunism or convenience. Pragmatic considerations, especially in relation to access to institutional sites, situations or hard-to-reach people, do have their place (see Hammersley and Atkinson, 1995). However, as Murphy et al. note, 'opportunistic sampling will be seen as the method of last resort in anything other than the most exploratory research' (1998: 93). Centrally, being able to describe your sampling as in some way strategic offers increased confidence in your work. There is a rhetoric of expertise that is embedded in such work. But this is beyond sheer rhetoric. It is about doing good analysis.

Exploring the Phenomenon in No Particular Order: (*n* = 14)

After conducting the three interviews with three types of 'typical case', I decided to interview the families of recently diagnosed patients. I had no particular logic about whom I approached, the only criterion they had to fit was that they had had a diagnosis in the last six months. Despite wanting to interview families with fresh memories of the experience, families were contacted at least one month post-diagnosis so as to avoid burdening the parents.

In sampling strategy terms, I undertook the least analytically strong option. I undertook something similar to what Patton refers to as convenience sampling:

> doing what is fast and convenient. This is possibly the most common sampling strategy and the least desirable. … Convenience sampling is neither purposeful or strategic. (2002: 241–2)

If my whole sample had been achieved by recruiting those who were easiest to hand, I would agree with Patton. For me, projects that *only* undertake such desperation sampling are generally problematic. However, as I was still in the initial stages, I wanted to explore the phenomenon, to get a generic sense of the potential issues and, with luck, to get a sense of the potential variance in the phenomenon.

I ended up conducting eight interviews with newly diagnosed families. As the unit of the analysis was paediatrics patients' route to diagnosis, I was interviewing parents, sometimes mothers or fathers on their own, sometimes both parents and, in one case, an adolescent child took part. The focus was on the very practical issues of what happened, in what order, alongside their emotional trajectory. For some of these patients, I also conducted parallel interviews with a health practitioner involved in their referral. I was still conducting very fine-grained coding, documenting the broad (and ever-growing) array of issues that were emerging in each new interview and constantly comparing the application of my codes with those that had gone before. However, at this stage, a potential key analytic issue was emerging, centred around the initial diagnosis the patients received from health professionals and how that impacted on delay. I kept returning to the same issue, within and across cases, and felt I might be getting somewhere.

Exploring the Phenomenon Through Somebody Else's Order: (*n* = 11)

So far I had conducted 11 interviews with families and 6 interviews with health professionals involved in the referral pathway. I felt I had begun to make some sense of the issues. Family resemblances were starting to emerge – especially around issues of initial decisions to seek lay and medical help. Fewer new patients were coming through the service. I discussed some of the recruitment issues with the team and we decided also to recruit more established patients. In clinics, the team were seeing new patients as well as those returning for regular three- or six-month check-ups and then thinking, ‘this would be an interesting case for Tim’. Their version of ‘interesting’ was often tied to specific issues of the case: for example, that the family had sought help from a private medical practice (a relatively uncommon thing in paediatric care in the UK) or that referral was ‘fast’ (i.e. under 10 weeks).

In this phase I interviewed another six families and a further five health professionals. The clinic staff felt they were offering interesting cases. For me, I was working with really quite ‘information-rich cases’, cases that, for whatever reason, could offer a new insight into the phenomenon. In technical terms I was using a mixture involving ‘deviant or atypical sampling’ (e.g. the parents who went private), ‘intensity sampling’ (e.g. the parents with relatively fast referrals) and ‘critical sampling’.

One case did turn out to be a ‘critical case’. Patton describes this as a case that makes a point dramatically: ‘[i]dentification of critical cases depends on the recognition of the key dimensions that make for a critical case’ (2002: 237). This was another account where, after an initial visit to a family practitioner, the parents were told the child had JIA, but there was nothing that could be done. After a period of time, the child no longer complained and the symptoms did not flare. Only five years later, when the mother noticed the child's restricted movement, did they return to seek advice. The child was then referred to the team and given the necessary medication and physiotherapy routines. This case was key, and critical for me in rethinking the analysis. For example, rather than understand the phenomenon under study as ‘delay in diagnosis’, we realized we needed to focus on delay in diagnosis *and* in receiving appropriate care. So a previous case from the initial round of sampling where the parents were told there was little medicine could do (discussed

above) was no longer to be understood as atypical or deviant. The phenomenon of ‘inappropriate care’ with a diagnosis of JIA was now central to our understanding. In Emmerson's (2004) terms, it was a ‘key incident’ in the trajectory of the analysis that enabled me to re-conceptualize the focus.

As noted above, our conceptualization of a ‘case’ included accounts from the patients' family members, patients, and health and social care professionals. So far, we had 10 cases, which included the health professionals' accounts and one teacher's account. Although they provided an additional layer of context, I felt that, analytically, they were often of relatively limited value. Also, it was proving very hard to contact practitioners, although once contacted they all agreed to take part, and then trying to arrange interviews was difficult, given their work schedules. I interviewed some of them on the phone, but found this less effective in generating a sufficient level of detail. Rather than spend more time on this, I decided to focus solely on families' accounts.

Note: Information-Rich Cases

Central to the success of purposive sampling is a focus on working with what Patton describes as ‘information-rich cases’ (2002: 230). These are the cases:

> from which one can learn a great deal about issues of central importance to the purpose of the inquiry, thus the term *purposeful* sampling. (Ibid.; italics in original)

Now, a case can range from an individual, a group, to an organization (and beyond). However, a case is not a naturally occurring object, it is a researcher's construct, a product of what Ragin (1992) refers to as ‘casing’. Centrally, through casing you are attempting to get information about some aspect of a particular phenomenon. As Miles and Huberman note, albeit it in relation to complex cases:

> you are sampling people to get at characteristics of settings, events, and processes. Conceptually, the people themselves are secondary. (1994: 33)

This might sound quite harsh, and beyond the limits of calls for qualitative research to do things like give others access to people's ‘voices’ or ‘lived experience’. But we are always only giving access to some aspect of that lived experience or organizational context. And if we take it that exploring a specific phenomenon is central to our research, we need to think about what makes up the focus of our casing.

What are the sampling units (or combination of units) that should guide your sampling? Rather than solely focus on the classic socio-demographic units, like age, ethnicity, etc., we need to think about more social, relational and conceptual units. For example, we could

consider structuring our sampling to focus on other issues:

- Actions – specific acts, processes, behaviours, intentions and motivations.
- Interactions – activities, formats, consequences and outcomes.
- Identities – roles, types, categories.
- Events – situations, rituals, ceremonies, temporal orders or trajectories.
- Settings and spaces – spatial (or conceptual) locations, organizations, milieu.
- Objects – devices, artefacts, electronic and paper texts.

Exploring the phenomenon is key, not being able to say ‘I observed X number of men and X number of women’. Relatedly, within-case sampling can also be important, especially in relation to more ethnographic studies. So, for example, in exploring a children's ward in a hospital, you may initially choose to focus on junior doctors. Over time, you may switch your focus to other actors in the setting, say parents. You may sample a related setting, those spaces where discussions about referring children to the ward first happen, such as the accident and emergency department or the children's day clinic. Or you might sample discussions about changing a child's medication and want to observe similar discussions across a range of contexts (e.g. with and without parents present, on the ward and in clinics) or a range of times (e.g. day, night, weekend). In this way, sampling is driven by emergent analytic issues.

Building Ideas, Challenging Assumptions (*n* = 17)

The fourth round of sampling shifted towards more conceptual development as well as focusing on some specific criteria. Analytically, I felt I had some sense of the key issues. I was becoming interested in exploring these further. So, for example, in trying to make sense of how people navigate through the system, I was interested in exploring the impact of people's prior knowledge and experience. I asked the staff to help me find cases containing one of the following three dimensions:

- Child diagnosed with JIA who had another significant illness that had been diagnosed prior to JIA emerging and for which the child received ongoing care.
- Family member who, through illness, had led the family to have significant contacts with the NHS (National Health Service).
- Family member who worked in the NHS in some capacity.

I conducted six interviews, covering six different families, two per dimension. Put very simply, I found that irrespective of how much prior knowledge the families had, the key facets of the trajectories echoed those of the other children without access to such potential ‘expertise’. This additional knowledge, in some cases, did impact on aspects of the trajectory, enabling the families to manage the problems they faced differently, but

the same problems existed. The addition of a previously diagnosed illness added to delay, in that practitioners' diagnostic focus was confined to explaining the new emerging symptoms through that particular lens. Rather than seeing these cases as somehow ‘atypical’, they echoed a broader issue throughout the data set: that the practitioners' search for a new diagnosis was routinely constrained by prior diagnostic expectations. All these accounts not only enabled me to develop these hunches, but confirmed my understanding of other issues alongside raising new issues.

One of the tensions I had to contend with was that between the conceptual development of my ideas and the possibility that the staff knew enough about a patient's trajectory to tell me who to interview. So, I was interested in exploring what can be best described as the role of ‘chance’, ‘luck’ or ‘timeliness’ in the patient's trajectory, as I had seen how central this had been. For example, in one case, the mother met a nurse, whom knew the family, in a hospital corridor. The nurse could see that the mother was distressed. This led the nurse to get her colleague, an adult practitioner, to ‘glance’ at the patient's x-ray results over their lunch. Her colleague was the first to suspect a diagnosis of JIA. I did not sample for this explicitly, due to the rather ephemeral nature of the issue, but rather expected (or maybe hoped) it would emerge in interviews. It did. It helped me conceptualize the work parents sometimes undertook to engineer these ‘chance’ encounters, to increase the possibility of encountering someone who might offer a diagnosis or treatment that made sense.

In this phase, my coding and analysis were becoming a lot more focused. I was selecting within the material, selecting specific stretches of talk for more in-depth analysis. Such coding work has strong family resemblances to the practices of sampling, especially theoretical and within-case sampling, in that strategic choices are made about what issues to focus coding on in order further to explore, challenge and confirm emergent ideas.

Note: Theory and Sampling

Despite the range of things written about the relationship between theory and sampling, there appear to be two main approaches to using theory to inform sampling.

First, following the tradition of theoretical sampling in grounded theory (in whatever version, see e.g. Glaser and Strauss, 1967; Glaser, 1978; Strauss and Corbin, 1990), after an initial round of sampling (driven by a priori ideas) to generate ideas, your next choice of person, site or situation is driven by the need to develop and elaborate on your emerging conceptual ideas. In grounded theory terms, you undertake theoretical sampling to help develop codes and categories, to understand variation in a process, to saturate properties of categories and to integrate them. In this way, your sampling decisions are emergent, progressive and inductive. Your task is artfully to choose a next case in order to progress the development of your emergent conceptual ideas. The focus here is not to demonstrate

empirical generalizability, in terms of choosing cases that might show others that you have sought variation to represent the population in some way. The focus is on developing the shape – the robustness – of your emergent categories and substantive theory. In this way, the demonstration of adequacy is understood in the transportability of the theoretical ideas.

Second, another tradition exists – one that receives less attention, but is potentially equally useful to consider. This is where the initial and often subsequent sampling decisions are driven by a priori theoretical ideas. This can take multiple forms. In such circumstances you may be exploring, testing and refining the ideas of an existing theory. Silverman notes that:

> in a case-study, the analyst selects cases only because he *[sic]* believes they exhibit some general theoretical principle. His account's claim to validity depends entirely on demonstrating that the features he portrays in the case are representative not of the population but of this general principle. (1985: 113)

So, for example, Silverman (1984) undertook some observations in a private medical clinic. He wanted to test Strong's (1983) theoretical ideas about the rituals of interaction between doctors and patients. Strong's work was based on extensive observations in public medical clinics, mainly in the NHS in the UK, so a private clinic offered an excellent space to test and refine Strong's theory. You will note that, in this example, as with others (see Murphy et al., 1998), this is often focused on ethnographic research where the choice of site is key. Given a lack of resources, use of more than a few sites is rare. So very good theoretical reasons for sampling a particular case can be central in claims-making. Of course, you need to choose a theory that is reasonably well recognized. Such a priori theory can also help support the selection of specific people, situations, times or places within a case.

Finding a Good Practical Solution to the Puzzle

It was during the fourth phase of sampling, when I sampled for specific theoretically driven issues alongside some typical cases, that I had my ‘eureka moment’. I met one of the team in the corridor and she asked me how it was going. I explained to her what I felt were the key issues emerging from my analysis. As I walked away, something in my account of the work kept coming back to me. I had used the analogy of the game ‘snakes and ladders’. Then, over a three-hour period, I wrote a conceptual memo. The title, albeit rather elaborate, offered up the main issues I wanted to explore through writing it:

> *MEMO – Snakes and Ladders – Persistence, Luck and Knowledge – Or how persistence (in symptoms and seeking a solution) combined with good and bad luck connect you with people with*

> *the relevant knowledge*

In the memo, I conceptualized the main issues that were central to the phenomenon of delay in diagnosis. In writing it, I moved between a large number of different documents, including coded and uncoded interview transcripts, summaries of interviews and codes, field notes from interviews and a range of types of memos. In and through this process I checked my ideas, sought out disconfirming moments from the cases, and brought together the hunches and leads I had tried to explore.

I knew my conceptualization worked – it made sense of the data, it offered up a coherent account of the phenomenon. Given that I had now developed a conceptual model that I felt made good sense of all the 'variance' of the phenomenon, I went on to carry out a more structured review of my data set, to explore whether I had any deviant or atypical moments or features in my cases that did not fit. I did not find anything that meant I needed to re-evaluate my ideas. With prior conceptualizations of the data, I had found exceptions that made me reconfigure my ideas.

Also, prior to writing the memo, I felt that, at least in terms of the accounts I was listening to, I had reached something like repetition. I was seeing the accounts as having very clear family resemblances. Some of the ideas were key; they were emerging again and again. More data would not help me understand or expand on my ideas any further. What was missing was some kind of conceptual model that linked the various ideas I was working with. All my attempts to offer something like a coherent account of the phenomenon either were far too unwieldy, or had too many exceptions, with many aspects of cases as atypical. However, repetition is not sufficient justification for stopping sampling. As Glaser notes:

> Saturation is not seeing the same pattern over and over again. It is the conceptualization of comparisons of these incidents which yield different properties of the pattern, until no new properties of the pattern emerge. (2001: 191, cited in Charmaz, 2006)

My eureka moment had enabled me to reach something like what grounded theorists call *theoretical* saturation. And in re-reviewing the already-collected data set, field notes and memos and engaging with the new data I was gathering, 'no new properties of the pattern' were emerging. However, I did not stop sampling at this point, as I wanted to test and refine this conceptual model. It still held over the next five interviews during this phase of sampling, albeit with some minor tinkering.

A Few More for Luck (*n* = 2)

By this stage, I now had a very good sense of my data; the issues and concepts were well developed; I had had my 'eureka moment'; and I had what I felt was a good analytic narrative. I had undertaken 34 interviews with families and 11 interviews with health and social care practitioners. The final stage emerged in part through pure opportunity and in part due to the 'doubts' many researchers feel. Two new patients had just been seen in the clinic and were now on the ward. I was told that these were both 'really interesting' cases. I did both interviews. Both cases were very interesting, and could be classified as 'intensity' cases.

Conceptually, I did not need to do them. I felt I had a coherent account – that I had reached saturation. However, I decided to do them for two reasons: first, the clinical team felt they would be useful; second, irrespective of how beautiful your conceptual ideas are, you always have a nagging doubt that you might have missed something. I do not feel we should ever be overly confident. We should always be open to having our ideas challenged. At this point, after working with these two cases, we decided that we should close recruitment. Clearly, I could have gone on, but in terms of time, money and resources, as well as imposing on people's lives, this would have been overly intrusive.

The Public Faces of Research

Over the life of the project, I conducted 36 interviews with families. I had spoken to mothers (*n* = 34), fathers (*n* = 9), teenage patients (*n* = 5), grandmothers (*n* = 2) and an aunt (*n* = 1). I had also undertaken 11 interviews with professionals involved in the care pathway of these JIA patients: orthopaedic surgeons (*n* = 4), paediatricians (*n* = 3), a paediatric immunologist (*n* = 1), a GP (*n* = 1), a nurse (*n* = 1) and one non-health professional (*n* = 1), a primary school teacher.

The sampling strategy was purposive; it was designed to explore, test and refine our emerging ideas. It was not designed to replicate the pattern of delay shown in cohort studies but rather to explore and map the diversity of factors that impact on that. However – and this was not planned – the sample closely matched prior quantitative research. You may remember that I noted that a prior study of 152 patients' case notes found that over 75% of patients exceeded 10 weeks from onset of symptoms to first paediatric rheumatology assessment (median 20, range 1–416 weeks). In our purposive sample (*n* = 36), over 73% of patients exceeded 10 weeks from onset of symptoms to first paediatric rheumatology assessment (median 22, range 1–362 weeks).

Now, this happened purely by chance. I had never sat down and tried to work out which patients we would need to recruit to get something that mirrored the prior quantitative work. This was excellent news as it added another layer of confidence that the sample was not somehow atypical of the population as conceptualized through statistical means. This was also excellent news when presenting the data to audiences that demanded a specific version of representativeness. However, throughout the life of the project I had attempted to recruit a range of ‘typical’ cases. My sampling was driven by both theoretical and what Sandelowski (1995) refers to as phenomenal variation. I had tried to explore the range of cases the clinic sees, to explore the variance in the phenomenon in terms of both the substantive issues (like length of delay) and the emergent theoretical issues (like knowledge of navigating health systems).

I want to focus on one last issue: the sampling we undertake when we present our data to others (see also Barbour, 1998). At the time of writing this chapter, I have only presented the data at three time points, namely two posters and one conference presentation. The posters were presented prior to my ‘eureka moment’. With these, in part, given the limited space on a poster, the posters focused on demonstrating the key ideas I was working with. They simply reported on specific concepts, with very few quotes of what I was thinking about.

For example, in relation to how parents conceptualize initial physical signs, the poster text reports:

> *As parents notice 'low grade' and often subtle physical and behavioural changes they rationalise observations as normal (e.g. 'drama queen'), adopt a 'wait and see' approach, and/or provide modifications to compensate (e.g. getting child new shoes). When symptoms are severe, escalate or continue, they then seek a medical opinion.*

Sampling specific accounts to demonstrate the phenomenon was offset by generic descriptions of process across the data set. The only direct quote in this section is 'drama queen', which I could have got from any number of transcripts.

Later in the analytic process, I offered a different presentational style. In drafting my conference presentation, I was unsure about how best to present the data. I tried various options, searching through my transcripts and memos to discover good exemplars. By 'good' exemplars I mean those that demonstrate specific aspects of an idea through concise and clear language. However, this became rather messy, as too much contextual detail was needed to place each issue in context and shifting between accounts of a large number of cases meant that the message was diluted.

In the end, I went for a different solution, simply comparing two cases, one 'intense' and one 'extreme'. So, for example, in relation to how parents conceptualize initial physical signs, the text on the slide for the conference presentation read:

> ***Bella – Aged 3***
>
> *Started walking early, 8 months,*
>
> *'… we'd occasionally see certain days that she'd be a little bit stiff and we'd think, "well is that because she's young and a bit too young to walk?" so you kind of pass that off as something else'*
>
> *'… at first I was thinking, "wellshe's new to walking, her muscles are developing, is she stiff because them muscles are developing, how long do we let this go on?"'*

And this discussion of the initial signs continued to another slide. After presenting the two cases, I directly compared their key features. As Sandelowski notes, in relation to more case-based research approaches:

> cases [are] conceived as singular combinations of elements constituting each case that are compared to singular combinations of elements constituting other cases. (2011: 157)

The intense case was typical, in terms of delay in diagnosis, and the extreme case was atypical in terms of a very rapid diagnosis and appropriate care. Irrespective of time from onset of symptoms to diagnosis and care, both cases represented all the key aspects of the conceptual model. I was interested in demonstrating how a specific configuration of the same elements led to a different outcome. Now, I could have presented any of the 36 cases, as each demonstrates a particular configuration. However, these two were chosen, in part for

their rhetorical impact, as they both presented well, given the relatively short time I had.

Closing Comments

The aim of this chapter was to offer you, the reader, 'technical access' to the lived practices of sampling. I hope you can begin to make sense of some of the issues that can weave through sampling. Ideally, your sampling strategy should be something that evolves over time, that emerges through a mutual relationship with desk, field and analytic work.

The obvious question, given the subject of the chapter, is to ask: why choose this specific research project as my single case? It was chosen for a range of reasons. In part, for opportunistic and pragmatic reasons – I am still working trying to write up the findings, and so it is still fresh in my mind. I also chose it because, given the widespread use of interviews, it was potentially a 'typical' case in terms of the methods. However, in terms of the phenomenon of applying sampling strategies to qualitative research in a strategic way, I am not sure how I would classify it. Hopefully it is a typical or intense case. Hopefully, it reflects elements of the lived practices of researchers using a wide array of methods and methodologies. Maybe trying to undertake sampling, in what I hope is a reasoned and thoughtful way, might mean it is a relatively 'atypical' case. However, exploring deviant cases can be an extremely useful thing in itself.

We should also ask: how generalizable (see Maxwell and Chmiel, Chapter 37, this volume) or transferable (see Murray, Chapter 40, this volume) is this case? Sadly, I have not got enough data on other people's approaches to sampling-in-action, in order to understand how empirically generalizable it is. This is a product of the relative paucity of accounts that describe the lived practices of sampling. However, it does relate to the more theoretical accounts about sampling. It also shares similarities with some of the practical issues that fellow researchers have discussed with me. So, despite an underlying argument running through the chapter about the utility of exploring and documenting the variance in the phenomenon, I can do little to demonstrate this case's empirical generalizability. It is for you, the reader, to decide. Hopefully, you can see this case as an ensemble of the very practical, contingent, analytic and theoretical issues that researchers are faced with. Perhaps you can think with some of these ideas in order to inform your thinking about your own sampling practices.

Notes

1. The research project was 'Exploring the pathways of referral for children with incident juvenile arthritis' Arthritis Research UK (Grant No: 17738).
2. For me, 'purposive sampling' and 'purposeful sampling' are synonyms.

Further Reading

Charmaz, Kathy (**2006**) 'Theoretical sampling, saturation, and sorting', Chapter 5 in *Constructing Grounded Theory: A Practical Guide through Qualitative Analysis*. London: Sage. pp. 96–122.

Gobo, Giampietro (**2004**) 'Sampling, representativeness and generalizability', in Edited by: **CliveSeale** et al. (eds), *Qualitative Research Practice*. London: Sage. pp. 435–56.

Patton, Michael Q. (**2002**) *Qualitative Research & Evaluation Methods*, 3rd edition. Thousand Oaks, CA: Sage (1st edition 1990). See the section entitled 'Purposeful sampling', pp. 230–47.

References

Barbour, Rose, S. (**1998**) 'Engagement, presentation and representation in research practice', in Edited by: **Rose S.Barbour** and **GuroHuby** (eds), *Meddling with Mythology: AIDS and the Social Construction of Knowledge*. London: Routledge. pp. 183–200.

Charmaz, Kathy (**2006**) *Constructing Grounded Theory: A Practical Guide through Qualitative Analysis*. London: Sage.

Dingwall, Robert (**1992**) '"Don't mind him – he's from Barcelona": Qualitative methods in health studies', in Edited by: **JeanneDaly** et al. (eds), *Researching Health Care*. London: Tavistock/Routledge. pp. 161–75.

Draucker, Claire B., **Martsolf, Donna S.**, **Ross, Ratchneewan** and **Rusk, Thomas B.** (**2007**) 'Theoretical sampling and category development in Grounded Theory'. *Qualitative Health Research*, 17: 1137–48. http://dx.doi.org/10.1177/1049732307308450

Emmerson, Robert M. (**2004**) 'Working with "key incidents"', in Edited by: **CliveSeale** et al. (eds), *Qualitative Research Practice*. London: Sage. pp. 457–72.

Foster, Helen E., **Eltringham, Michal S.**, **Kay, Lesley J.**, **Friswell, Mark**, **Abinun, Mario** and **Myers, Andrea**. (**2007**) 'Delay in access to appropriate care for children presenting with musculoskeletal symptoms and ultimately diagnosed with juvenile idiopathic arthritis', *Arthritis Care and Research*, 57: 921–7. http://dx.doi.org/10.1002/art.22882

Glaser, Barney G. (**1978**) *Theoretical Sensitivity*. Mill Valley, CA: Sociology Press.

Glaser, Barney G. (**2001**) *The Grounded Theory Perspective I: Conceptualization Contrasted with Description*. Mill Valley, CA: Sociology Press.

Glaser, Barney G. and **Strauss, Anselm** (**1967**) *The Discovery of Grounded Theory*. New York: Aldine de Gruyter.

Gobo, Giampietro (**2004**) 'Sampling, representativeness and generalizability', in Edited by: **CliveSeale** et al.

(eds), *Qualitative Research Practice*. London: Sage. pp. 435–56.

Guba, Egon G. (**1981**) 'Criteria for assessing the trustworthiness of naturalistic inquiries', *Education, Communication and Technology Journal*, 29: 75–91.

Hammersley, Martyn (**1992**) *What's Wrong With Ethnography? Methodological Explorations*. London: Routledge.

Hammersley, Martyn and **Atkinson, Paul** (**1995**) *Ethnography: Principles in Practice*, 2nd edition. London: Routledge (1st edition, 1983).

Lincoln, Yvonne S. and **Guba, Egon G.** (**1985**) *Naturalistic Inquiry*. Newbury Park, CA: Sage.

Miles, Matthew B. and **Huberman, A. Michael** (**1994**) *Qualitative Data Analysis: An Expanded Sourcebook*, 2nd edition. Thousand Oaks, CA: Sage (1st edition 1984).

Mitchell, J. Clyde (**1983**) 'Case and situation analysis', *Sociological Review*, 31: 187–211. http://dx.doi.org/10.1111/j.1467-954X.1983.tb00387.x

Murphy, Elizabeth, **Dingwall, Robert**, **Greatbatch, David**, **Parker, Susan** and **Watson, Pamela**. (**1998**) 'Qualitative research methods in health technology assessment: A review of the literature'. *Health Technology Assessment*, 2 (16).

Patton, Michael Q. (**2002**) *Qualitative Research & Evaluation Methods*, 3rd edition. Thousand Oaks, CA: Sage (1st edition 1990).

Ragin, Charles C. (**1992**) '"Casing" and the process of social inquiry', in Edited by: **Charles C.Ragin** and **Howard S.Becker** (eds), *What Is a case? Exploring the Foundations of Social Inquiry*. Cambridge: Cambridge University Press, pp. 217–26.

Sandelowski, Margarete (**1995**) 'Focus on qualitative methods: Sample size in qualitative research', *Research in Nursing and Health*, 18: 179–83. http://dx.doi.org/10.1002/nur.4770180211

Sandelowski, Margarete (**1996**) 'One is the Liveliest Number: The Case Orientation of Qualitative Research', *Research in Nursing & Health*, 19: 525–9. http://dx.doi.org/10.1002/%28SICI%291098-240X%28199612%2919:6%3C525::AID-NUR8%3E3.0.CO;2-Q

Sandelowski, Margarete (**2011**) '"Casing" the research case study', *Research in Nursing & Health*, 34: 153–9.

Silverman, David (**1984**) 'Going private: Ceremonial forms in a private oncology clinic', *Sociology*, 18: 191–204. http://dx.doi.org/10.1177/0038038584018002003

Silverman, David (**1985**) *Qualitative Methodology and Sociology*. Aldershot: Gower.

Strauss, Anselm and **Corbin, Juliet** (**1990**) *Basics of Qualitative Research: Grounded Theory, Procedures and Techniques*. Newbury Park, CA: Sage.

Strong, Phil (**1983**) *The Ceremonial Order of the Clinic*. London: Routledge & Kegan Paul.

Ward Schofield, Janet (**1993**) 'Increasing the generalis-ability of qualitative research', in Edited by: **MartynHammersley** (ed.), *Social Research: Philosophy, Politics & Practice*. London: Open University/Sage. pp. 200–25.

Williams, Malcolm (**2002**) 'Generalization in interpretive research', in Edited by: **TimMay** (ed.), *Qualitative Research in Action*. London: Sage.

Yin, Robert. K. (**1994**). *Case Study Research: Design and Methods*, 2nd edition. Thousand Oaks, CA: Sage (1st edition 1984).

http://dx.doi.org/10.4135/9781446282243.n4