FOCUS ON METHODOLOGY

Inductive content analysis: A guide for beginning qualitative researchers

D. F. Vears[1, 2] & L. Gillam[2, 3]

Abstract

Inductive content analysis (ICA), or qualitative content analysis, is a method of qualitative data analysis well-suited to use in health-related research, particularly in relatively small-scale, non-complex research done by health professionals undertaking research-focused degree courses. For those new to qualitative research, the methodological literature on ICA can be difficult to navigate, as it employs a wide variety of terminology and gives a number of different descriptions of when and how to carry it out.

In this article, we describe in plain language what ICA is, highlight how it differs from deductive content analysis and thematic analysis, and discuss the key aspects to consider when making decisions about employing ICA in qualitative research. Using a study investigating practices and views around genetic testing in children as an example, we provide a clear step-by-step account of analysing text using ICA.

Clear guidance on ICA will be useful for beginning researchers, especially those more familiar with quantitative biomedical and behavioural research, and for their academic and professional supervisors.

Keywords: qualitative content analysis; inductive content analysis; qualitative research methods

[1] Murdoch Children's Research Institute, The Royal Children's Hospital, Australia
[2] Department of Paediatrics, University of Melbourne, Australia
[3] Children's Bioethics Centre, The Royal Children's Hospital, University of Melbourne, Australia

Correspondence
Dr Danya Vears
Email: danya.vears@mcri.edu.au

Introduction

As health researchers who teach qualitative research methods, conduct research using qualitative methods and supervise the research projects of health professionals and healthcare students, we frequently need to consider the uses, limitations and boundaries of specific research methods. As such, we often search the literature trying to find useful resources with clear guidance to assist researchers and students in the health field who are new to qualitative methods.

Unfortunately, our attempts to identify such resources for one particular methodology, inductive content analysis (ICA), have led to considerable frustration for a number of reasons. We explain these reasons in detail below, with reference to the literature, but briefly they are as follows. First, the terminology and naming of the method varies, making searches for useful methodological articles and examples of how others have used it quite challenging. Second, descriptions of how to conduct ICA are often poorly articulated, leaving novice researchers confused. Third, there is a lack of good references that delineate the differences between ICA and other apparently similar analysis methods (in particular, thematic analysis and deductive content analysis).

To address these needs, this article aims to describe in easy-to-understand terms what ICA is and when it is appropriate to use it. We also outline in plain language a set of steps to use when undertaking ICA. These steps describe a process that we have found helpful but are not a series of rules that must be followed. Our aim is that this will be a resource to guide researchers new to qualitative research and to promote rigour in the conduct of qualitative research for those new to the field. This article is presented in a way that directly addresses the new researcher, not the expert. In order to make the general descriptions that we give easier to understand, we refer throughout to a research study conducted by one of the authors, Danya Vears. We describe this study in Box 1.

Overview of ICA

ICA is a method of analysing data. It is commonly used with text-based data, either written transcripts of verbal interactions or documents created in written form. These documents may be pre-existing documents written for non-research purposes (such as a doctor's referral letter or a submission to a government inquiry), or they may have been created for the specific purpose of a research study (such as a blog written by a research participant). ICA can also be used for images, but in this article, we will focus on textual data only. In very basic terms, ICA essentially involves producing an overall summary of the content of different individual texts in a data set (for example, a set of interview transcripts or referral letters). The critical feature of ICA, as we will describe in more detail below, is that the analysis is built up inductively, from a close reading of the texts rather than searching the text for a pre-determined list of content items.

ICA is one type of qualitative analysis. It is qualitative in that it aims to produce an understanding of the meaning/s of the content of the data set. It does not count instances

of content to calculate frequencies or find statistical correlations between different content items. In contrast, quantitative content analysis often involves counting (Krippendorf, 2004). As we discuss below, ICA has many similarities with thematic analysis, which is perhaps the best-known form of qualitative analysis, especially for interview transcripts (Braun & Clarke, 2006). One obvious initial similarity for beginning researchers is that both thematic analysis and ICA involve a process known as "coding", which involves labelling chunks of text in each interview transcript or document and using these labels to identify and bring together pieces of text that are similar, both within a document/transcript and across documents/transcripts. There are also differences between ICA and thematic analysis that make it important to be clear which method is being used. For reasons that we expand on below, we suggest that for researchers new to qualitative research who are undertaking research intended to be directly relevant to healthcare practice or policy, ICA is often a better choice.

In the literature, a myriad of terms is used to refer to ICA. While some authors refer to it as "qualitative content analysis" (Cavanagh, 1997; Graneheim & Lundman, 2004; Krippendorf, 2004; Polit & Beck, 2004; Vaismoradi et al., 2013), others use "content analysis" alone (Downe-Wamboldt, 1992; Weber, 1990) or "interpretive content analysis" (Ahuvia, 2001). While variation in terminology is a feature of qualitative research more broadly and is not necessarily troubling to experienced researchers, it is often very confusing for beginning researchers. We suggest using the term "inductive" rather than "qualitative" or "interpretive", because it is most informative about the analytic style of the method and helps differentiate it clearly from deductive content analysis, which is one of the forms of analysis that ICA may be confused with.

So, what defines ICA? In our view, its key characteristics are that it is an *inductive process* and involves *iterative coding*. By inductive process, we mean that the codes used to label the data are developed during the process of coding, based on the actual content of the data set. The codes are identified by the researcher within the data itself, or as is often said "arise" (Bennett et al., 2019; Sousa, 2014) or "emerge" from the data (Lichtman, 2014; Morse, 1994). This is in contrast to deductive content analysis, where the codes used to analyse the data are predetermined, generally based on previous research in the field and/or a conceptual framework or model. In deductive content analysis, the codes are developed before the data is collected, without reference to the data.

Involving iterative coding means that the process of coding is not done only once for each document/transcript but is refined on the basis of comparisons between documents/transcripts and then repeated. Each document/transcript is coded a number of times in more refined iterations each time. Because the coding is inductive and new aspects will continue to be identified from the data as more documents/transcripts are analysed, this iterative re-coding is a vital step to check that the newly identified codes were not "missed" in the earlier coding rounds. It is quite common for ideas that are identified in later transcripts to be present in earlier transcripts, perhaps in more subtle ways, but

not recognised when those transcripts were coded for the first time. So, the researcher is adding to and adapting the list of codes over the course of the analytic process. This is different from coding using deductive content analysis, where the list of codes, being pre-determined, is fixed and does not change as the analysis progresses.

The process of coding, followed by comparing, grouping and sub-dividing groups of codes, results in content categories and subcategories rather than "themes", for reasons explained below. A "content category" is a broad idea or concept within which a number of more specific content codes have been grouped. If the researcher is conducting interviews or focus groups, content categories may often be quite closely linked to the questions the researcher is asking participants, though this is not necessarily so.

Box 1. Project example: Carrier testing in children

Cystic fibrosis (CF) is a genetic condition that begins in childhood and causes a build up of thick mucus in the lungs leading to chronic lung disease. Thanks to her training as a genetic counsellor, Danya knew that when a child is found to have CF, some parents ask to have their unaffected children tested to see if they carry this condition. Carriers of CF are generally healthy, but if they have children with another carrier, they have a 25% chance of having an affected child. However, performing this carrier testing in children is not recommended by testing guidelines because there are concerns about potential psychological harms from children knowing their carrier status and desires to preserve their right not to know their carrier status until they reach adulthood and can make this decision for themselves.

At the time that Danya was starting her PhD, little was known about how genetic health professionals (i.e., genetic counsellors and clinical geneticists) respond to parents' requests for carrier testing in their unaffected children. There was also very little published research about why parents might want to know this information about their child, given it has no bearing on their health in childhood.

To explore these research questions, Danya undertook semi-structured interviews with both genetic health professionals (GHPs) and parents of children with one of three genetic conditions (CF, Duchenne muscular dystrophy and haemophilia). The interviews with the GHPs explored their experiences with parents requesting carrier testing in their unaffected children, how they responded to these requests (i.e., did they accede to or refuse testing) and their beliefs about why parents want to know this information. The interviews with parents explored their reasons for wanting to know their child's carrier status and their experiences with requesting carrier testing (i.e., did they or did they not receive testing for their unaffected child/children).

Danya used inductive content analysis to analyse the interviews from both GHPs and parents. This study was approved by the Human Research Ethics Committee at The University of Melbourne, Victoria (ID 1137204). Consent was obtained from all participants prior to participation.

For example, in Danya's study with parents of children with genetic conditions, her main research question focused on parents' reasons for wanting to know about the carrier status of their other children, who were not affected by the genetic condition. Her key interview questions were designed to elicit discussion of this, and "reasons for wanting carrier testing" was a key idea she wanted to explore in the interview transcripts to answer her research questions. The content categories she identified were different types of reasons for wanting testing. The term "theme" is sometimes used to refer to broad content categories identified by ICA, but we suggest "theme" is better reserved for use in thematic analysis. A theme that is developed using thematic analysis would typically be a more abstract or theoretical concept than a content category.

Our understanding of ICA is similar to that described as conventional content analysis by Hseih and Shannon (2005). It also shares similarities to the methods described by Elo and Kyngas (2008) and Graneheim and Lundman (2004), although the latter seem to describe further abstraction of the content categories into themes, which we see as a potentially confusing use of terminology for beginning qualitative researchers.

When is it appropriate to use ICA?

In our experience, there are several questions that one needs to consider when deciding whether ICA is the appropriate analysis method to use for a particular research project. Importantly, these are questions that should be asked when designing the project as a whole to ensure that the research question/s, conceptual framework, methods and intended output all align.

1. What is already known about the area of research?

ICA is a good type of analysis to use when there is little in the way of existing research in that area (Hsieh & Shannon, 2005; Vaismoradi et al., 2013) or where an approach that does not depend solely on the existing literature is desired. Because the development of content categories is inductive, starting from the data itself, it does not need an established theory or account as the starting point. If there is a standard body of literature about the area, with a widely accepted interpretation or model, a beginning researcher might decide simply to apply that existing model. In this case, deductive content analysis would be an appropriate choice. Deductive content analysis can also be used to test existing theories about a phenomenon (Elo & Kyngas, 2008; Hsieh & Shannon, 2005). Thematic analysis could also be appropriate when there is little known about the area of interest, so this first factor does not fully distinguish between ICA and thematic analysis. This leads us to the second factor.

2. What kinds of outcomes is the research intended to achieve?

ICA is very useful when the aim is to describe and understand the phenomenon under investigation (Bloor & Wood, 2006; Elo & Kyngas, 2008; Vaismoradi et al., 2013) in a

way that has immediate resonance for health professionals and has quite direct relevance for practice or policy (Thorne et al., 2004). In this sort of qualitative health research, the interview questions are likely to be quite specific and direct (rather than very broad) to elicit descriptions of the situation, event or experience. ICA is particularly appropriate when the researcher is aiming for a practical answer or application of the findings (such as to develop practice guidelines or policy). This is in contrast to a more theory-focused aim of explaining the worldviews behind the phenomenon or locating it within a more general theoretical understanding of the world. In the case where a more theoretical answer that goes beyond the particulars of the situation (Thorne et al., 2004) is sought, with the goal of making a contribution to the theoretical literature, we suggest that thematic analysis would typically be more appropriate.

How is data analysed in ICA?

Understanding the task at hand

Qualitative data analysis is vividly described by Janice Morse (1994), a leading qualitative methodologist, as follows: "You have to set the stage, organize your data, and organize your brain, so that you set the stage for having the insight" (p. 23). Although some articles list steps such as "formulate research questions" (Hsieh & Shannon, 2005) and "select data set/data type" (Downe-Wamboldt, 1992) as the first steps for data analysis, these steps are actually completed in earlier phases of the research process (Denzin & Lincoln, 2017) before data analysis begins. Once data has been collected to answer research questions, data analysis (that is, organising the data) can begin. Some authors put "select unit of analysis" as the first step in the analysis process (Downe-Wamboldt, 1992), which seems very sensible—identify what is going to be analysed before embarking on working out how to analyse it. Unfortunately for beginning researchers, the term "unit of analysis" is used in different ways in the literature. While some authors appear to be using it to refer to *a piece of data* that makes up part of the data set (such as a word, phrase, piece of text or even an entire interview) (Ahuvia, 2001; Downe-Wamboldt, 1992; Graneheim & Lundman, 2004; Zolnoori et al., 2019), others seem to be referring to a "unit of analysis" as a *concept* that the researcher might want to code for (a "code" would then be the label used to identify that concept) (Hansen, 2006; Schick-Makaroff et al., 2016; Zhang & Wildemuth, 2009). We suggest that a helpful way of understanding a "unit of analysis" is as a chunk of text (which may be short or long) that has a meaning in relation to the research question/s. This understanding incorporates key elements of these apparently competing definitions and directs the attention of the beginning researcher to the thing that matters most in ICA—namely, the meaning of the texts. The alternative term "meaning unit", as used by Erlingsson & Brysiewicz (2017), conveys this well. So, the first step is to identify meaning units, a process that we describe in more detail below.

Another source of confusion is the question of whether latent or manifest content is being analysed, which some accounts of ICA put much emphasis on. Manifest content analysis

has been described as "the analysis of the visible or apparent content of something" (Boyatzis, 1998, p. 16). In contrast, latent content analysis is "looking at the underlying aspects of the phenomenon under consideration" (Boyatzis, 1998, p. 16). For beginning researchers, it is important to be aware of this distinction, but not to become bogged down in it. We suggest thinking of the manifest/latent distinction as describing two ends of a spectrum. At one end (the "manifest" end), almost no inference from the words on the page to the meaning would be involved. The further along the spectrum you move, the more you are making an inference from the words on the page to the meaning. At the furthest end (the "latent" end), the inferential leaps are much larger. Boyatzis (1998) gives a good example of a particularly large inferential leap, the interpretation of dreams, where the interest is in the latent content believed to lie behind the manifest content of the words in the transcript. Other authors who provide detailed discussion of latent content analysis include Ahuvia (2001), Kleinheksel et al. (2020) and Thayer et al. (2007).

Some interpretation of the words in text is always required to articulate the meaning of the text. This is particularly so for transcripts of spoken words, where the words on the page may not readily convey the meaning that was quite apparent when the participant spoke to the interviewer. However, for a beginning researcher using ICA, we recommend not to take interpretation too far. Stay close to the manifest content end of the spectrum by using the common sense meaning of the words that you can infer from the immediate context.

Outlining the steps in the process of ICA

A first point to note is that analysis is ultimately a thinking process. The tools for doing data analysis can be a pen and paper or a software package. The thought process is the same no matter what tools you use. You will need to decide on what tools to use, but you can always start with pen and paper while investigating different software packages.

Step 1: Read and familiarise

Morse (1994) identifies the first cognitive process in data analysis as "comprehending". Before launching into the process of selecting and labelling chunks of text, read through all of your data so that you become very familiar with it. This is even more important if you have not done all of the data collection yourself (for example, if you are working with transcripts of interviews or focus groups that were conducted by another researcher). This step helps you think holistically—what are these texts or transcripts *about*?

Step 2: First-round coding—identify big-picture meaning units

We suggest that a helpful first step in coding is an extension of the cognitive process of comprehending. This starting step involves organising your data in "big picture" terms. A process for doing this is to identify and label sections of the texts according to their "big-picture meaning". We use the term "big-picture meaning" to refer to a "class" or "type" of content—a broad category of content that is relevant to the research question/s.

Depending on the research question/s, sections of text might be labelled as an "experience or (some sort of) event", a "fear for the future", a "reason for doing (some behaviour)", a "criticism of (some sort of) policy" and so on, without yet focusing on the details.

Looking for big-picture meaning involves reading the text or transcript carefully, asking yourself "what is this section *about*?" Remember that the aim of the analysis is to answer your research question/s, so these big-picture categories will be connected to the research question/s in some way. In identifying sections of text that have meaning in relation to the research question/s, you are beginning to identify units of meaning. For example, in coding the interviews with parents in the carrier testing project, some of the big-picture categories that Danya used in this initial coding step were "experiences with requesting testing for their unaffected child", "reasons for wanting to know child's carrier status" and "parents' reactions to receiving results" (Vears et al., 2016). We sometimes refer to this step in coding as "chunking", but more formally it is "big-picture coding". A sample of this coding is provided in Figure 1.

If you are analysing interview transcripts rather than other types of texts, the big-picture categories may track quite closely to some of the questions your participants were asked in the interview, depending on how the interview questions were designed for the purposes of eliciting responses about the phenomenon of interest. For other types of texts, the structure of the text may give some guide to the big-picture units of meaning to be coded. However, be careful not to simply follow the existing structure, as the content may be organised in a way that is not helpful in answering your research question/s. Usually, a section of text will be coded as only one big-picture category, but because of the way people speak and write, it sometimes turns out the same chunk of text needs two different big-picture codes (because it contains two meaning units). The size of these sections will depend on how much of the text relates to each big-picture category. There is no correct or pre-set size—it may be a paragraph, an entire page or a number of pages (see Figure 2).

Figure 1

Example of First-Round Coding in Two Transcripts

Transcript 1

Researcher: And what was it that made you decide to get carrier testing done in <sibling>.

Vanessa: Sure. We just as a family made a decision before the children were even born that we, going to be age appropriate but we were never going to hide anything from them or lie or you know just be an honest family um so we just think it's important from two perspectives. One, if <sibling> asks we want to be able to give him the answer and also when we start explaining to <child with CF> age appropriate about cystic fibrosis and how it works and all that sort of stuff, we'll be able to tell her how she got it and we think that at some point logically she's going to ask the question about <sibling> so we'd like to give her that information too. Um and that's in a nutshell is it. Just, we're just honest and if you can have the information we'd like to give it to them so and <sibling> can make his decision about what he chooses to do with that information when he's older.

Category: Reasons for wanting carrier testing in unaffected child

Researcher: Hmm. And so you went through the process of carrier testing with him. Um what did that involve?

Vanessa: Um they had a, they had a, I think it was the blood sample from the heel prick test? I think she said that she could put that through for the carrier testing? I don't think he got a blood test other than that. I do this they said they might need to do it again but they didn't and so she called I think about 2 weeks later. Um oh about 10 days later and ah she reconfirmed did I want to know (laughs)

Transcript 2

Researcher: And around that time were to thinking, well I guess once you had the diagnosis for <child 1 with haemophilia> finally, were you thinking about the implications for your daughter <female sibling>?

Violet: No um I don't think we started thinking about <sibling> until a lot later because we didn't know whether I was a carrier at that point. Um We didn't know what my levels were, not that that makes a difference to her but we didn't know how it would be affected and then we found out that I was a carrier and then we were like well I haven't had any issues but we were like well what happens if her levels are lower um you know what if she is a carrier. Um I watched my sister go through the testing um and that was 3 months of a lot of stress for her while she was doing uni and just went that's so hard like it really mucked her around. Um and I'm like well I wanted to be able to have answers when <sibling> asks me. So I didn't want to be like well you might be a carrier, you might not, we don't know rah rah rah. So yeah I was really keen to find out. I didn't want her to be in an accident and have really low levels and that sort of thing. I wanted to know what to tell the school and the paramedics if something happened.

Category: Reasons for wanting carrier testing in unaffected child

Researcher: Yeah

Violet: That sort of thing so it wasn't just a case of um just wanted to know, well I did want to know out of curiosity obviously but I didn't want her to go through the stress of testing and not worrying and that sort of thing. So we did it, I think she was three when we got her tested.

Category: Reasons for wanting carrier testing in unaffected child

Note: Transcript 1 shows a segment of a transcript of an interview with a parent of a child with cystic fibrosis and transcript 2 shows a segment of a transcript of an interview with a parent of children with haemophilia. In both, the "big picture", or broad category, of "reasons for wanting carrier testing in unaffected child" is coded (pseudonyms used).

Figure 2

First-Round Coding

Note: When coding "chunks" of the text into "big picture", or broad content, categories, the size of the chunk depends on how much of the transcript or text relates to that category. It may be a sentence (A), a paragraph (B) or an entire page (C).

Once you have completed the first round of coding, working through your texts or transcripts in small groups and comparing and adjusting codes as you go, you will have a full set of transcripts or texts that all have a large proportion of the words (but not necessarily every word) labelled with a big-picture code. You will also have a preliminary list of the big-picture categories that you have created, which is the first version of your coding tree or coding schema (see Figure 3). Remember that this is not a fixed or finished piece of analysis. It is a work in progress, which can be refined in the next stage of coding.

At the end of these initial steps, "line-by-line" coding (i.e., text where every line has its own fine-grained code), which so many authors describe as being a key aspect of qualitative data analysis (Charmaz, 2006; Elo & Kyngas, 2008; Thomas & Harden, 2008), will not yet have been achieved. These initial steps of big-picture coding assist beginning qualitative researchers to work towards line-by-line coding without being overwhelmed or lost.

Figure 3

Initial Coding Scheme

Preliminary big-picture coding schema

a. Experiences with affected child's diagnosis
b. Reasons for wanting to know carrier status
c. Reactions to receiving results
d. Parent's views on disclosure
e. Carrier status identified incidentally
f. Parents refused testing for unaffected child

Note: List of "big picture" content categories developed after the first round of coding the transcripts of interviews with parents

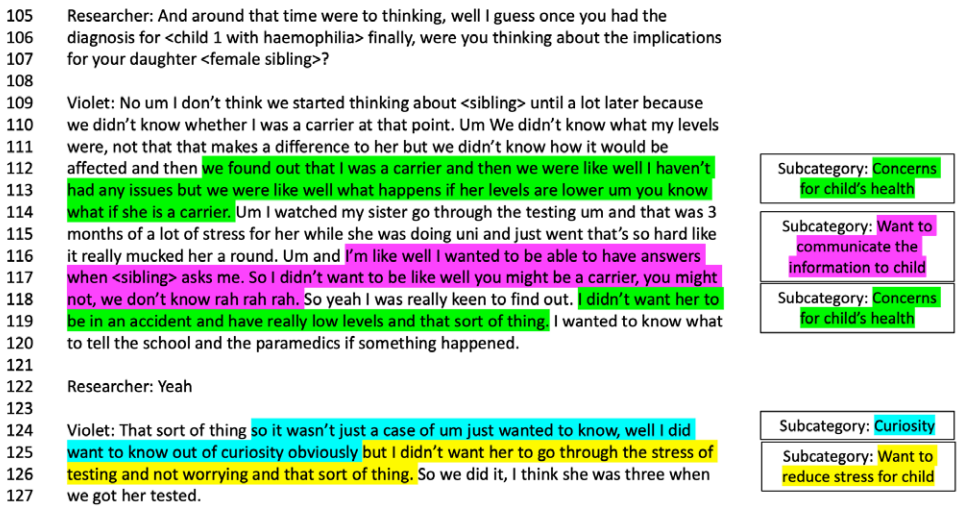
Step 3: Second-round coding—developing subcategories and fine-grained codes

The second coding step involves taking a closer look at the text within each of the sections that you have coded as a particular big-picture category. This step allows the beginning researcher to "break open" the data, to use the helpfully vivid term from Corbin and Strauss (2008), into pieces so that the many rich and varied ideas and meanings within it can be seen—but without totally unmooring these ideas from the context in which the original words were written or spoken.

To do this step, choose a big-picture category and look at all of the sections in all of the transcripts under that category. Each of these sections needs to be coded in a more fine-grained manner, going through the text of the section line by line. Think about the specifics of what is being said by the text and then label it with a new code that describes this succinctly. You could use words directly taken from the text for this type of code or a word or phrase that is close in meaning to the words in the text. Second-round coding, in essence, produces subcategories of the big-picture categories. Referring back to our example, the big-picture category of "reasons that parents want to know child's carrier status" was coded into subcategories, which included "peace of mind", "out of interest/ curiosity", "information seeking as a way of coping", "reassurance of child's health" and "wanting to communicate the information to their child" (Vears et al., 2016). Do this fine-grained coding for each of the big-picture categories. An example is provided in Figure 4.

Figure 4

Example of Second-Round Coding



Note: Using the same segment of transcript 2 shown in Figure 1, the broad category of "reasons for wanting carrier testing in affected child" is now re-coded into subcategories.

In this third step, the idea of subcategories should not be interpreted too rigidly. This step should not be understood as some sort of deductive process of assigning small bits of text to big-picture codes that are now fixed. On the contrary, the process of line-by-line coding within all the sections of text that are labelled with the same big-picture code is likely to suggest refinements or revisions to your initial big-picture categories.

Step 4: Refining the fine-grained subcategories

After you have coded all the text within all the big-picture categories into subcategories, then you need to compare and refine all of these fine-grained subcategories. Some subcategories may be so similar that they should be collapsed into one; others will be sufficiently different from each other that they should remain as stand-alone subcategories. Also look to see if any subcategories actually relate to a different big-picture category from the one you currently have them coded to. This will require comparing all the chunks of text coded in each subcategory from all transcripts or texts.

In carrying out this comparative process, you may realise that some categories are still too broad, or too vague, and need further specification. For example, Danya looked at all of the pieces of text that were coded into the subcategory "wanting to communicate the information to their child" together. She found that a third round of coding was required to fully describe the content of this subcategory. Further coding under this subcategory highlighted aspects that included "wanting to avoid the shock of finding out later", "providing them with future reproductive opportunities" and "allowing gradual integration of the information over time".

It is important to keep in mind that second- (and third-, if required) round coding is an iterative process where subcategories might be grouped and ungrouped in order to best account for the richness and complexity of the data, without being repetitious or cutting the data up into so many tiny categories that the meaning is lost.

The end result of these rounds of coding is a refined coding tree or coding schema. An example is provided in Figure 5. With this fine-grained coding in place, you are now in a position to move to the final step of analysis.

Step 5: Synthesis and interpretation

Interpretation of the data is a critical step in the analysis process. However, in contrast to thematic analysis, which generally involves aiming to produce an explicitly theorised overall interpretation, the aim in ICA is typically to stay closer to the phenomenon you are investigating and produce an interpretation that provides a rich answer to your research question/s that is practically relevant to the research context. We like to think of this as "internal" interpretation, in that it is a process of synthesising and connecting the categories to create a narrative for the reader that gives an overall explanation of the phenomenon.

Figure 5

Refined Coding Schema

a. Experiences with affected child's diagnosis

b. Reasons for wanting to know carrier status
   i. Peace of mind
   ii. Curiosity/interest
   iii. Coping
   iv. Needing to know
   v. Wanting to communicate information to child
      Providing child with future reproductive opportunities
      Wanting to avoid shock of child finding out later
      Allowing gradual integration of information over time
      Better to tell when children are younger
      Better to tell status than just risk
   vi. Health concerns

c. Reactions to receiving results
   i. Satisfaction or relief
   ii. Unconcerned
   iii. Upset initially but pleased and peace of mind
   iv. Perceived utility

d. Parent's views on disclosure
   i. Had already communicated results
   ii. Intended to communicate results
   iii. Age intended to communicate
   iv. Perceived level of confidence to communicate results
   v. Could understand why some parents may forget

e. Carrier status identified incidentally
   i. Unconcerned because not the focus
   ii. Mixed emotions
   iii. Unsure if would have requested
   iv. Happy to know

f. Parents refused testing for unaffected child
   i. Could understand why testing wasn't performed
   ii. Unhappy with decision

Note: List of content categories and subcategories developed after the third round of coding of the transcripts of interviews with parents

To do this in practice, you need to bring together ("synthesise") the content categories and subcategories in a way that makes an overall meaning, showing how the different content categories relate to each other and tell a coherent story. It can be helpful to think "what is the message or story that my data is trying to tell?" Yet in undertaking this high-level interpretation, it is important not to lose the nuances of the data because the purpose of qualitative research (as opposed to quantitative methods) is to explore the complexity of a phenomenon. Within the overall story or interpretation, it is important to represent the complexity contained in the data rather than to make things neat and tidy by omitting bits that don't fit so well. Even if only one participant expressed a particular reason for wanting carrier testing performed in their child, it is an important finding to report because we are trying to investigate the scope/range of reasons not only the most prevalent. When you are writing the results section of your thesis or journal paper, think of the overall story as being like a map. It is important to represent all of the back streets, not just the main road.

While synthesis does not necessarily rely heavily on theoretical concepts external to the data, it should use existing theory and existing interpretations of similar phenomena to help frame and enrich the interpretation. There are many possible theories or conceptual frameworks that can be used for this purpose. Deciding on which are the most appropriate involves considering a range of factors, such as the scope of the research question/s and the aims of the research (Liamputtong, 2020). For example, in relation to the "reasons that parents want testing" category, the interpretation step involved making sense of the reasons that parents give by connecting them into a cohesive message. During the synthesis and interpretation step, Danya noted that these reasons could be viewed in another way, namely as reasons that were of benefit to the parents (such as "out of interest/curiosity" and "peace of mind") and reasons that were of benefit to the children (such as "wanting to avoid the shock of finding out later" and "providing them with future reproductive opportunities") (Vears et al., 2016). This way of looking at the types of reasons was suggested by an ethical framework used to consider whether decisions that parents make for their children should be accepted and acted on by health professionals or not.

In this framework, distinguishing between parents' reasons that relate to themselves and reasons that relate to their child is significant. The distinction draws on a more general distinction between self-regarding and other-regarding reasons, which is standard in the philosophical ethics literature, and is a key factor in assessing the ethical justifiability of some actions. It is important to note that this interpretation did not come from deductively applying that specific theoretical lens to the data from the outset. The analysis was done inductively, with content categories arising out of the data. A theoretical framework was brought in after the coding was complete to assist in interpreting the overall significance of what had been found.

This process of interpretation will be happening as you are trying to figure out what will be included in the results section of your article or thesis and the order in which you will present your results. Typically, however, your account of how the categories connect and how you have made sense of the data to produce an overall interpretation will be written into the discussion section.

Conclusion

ICA is a meaningful approach for analysis of qualitative data. It is particularly suited to practically oriented research of the kind often undertaken by students in health-related courses and health professionals moving into qualitative research for the first time, including in the context of a research-focused degree course. The best way to decide whether ICA is appropriate for your research is to consider what sorts of answers you are seeking for your research question/s, what audience/s you most directly want to address and how you would envisage your findings being used. As we have described, ICA is not "merely descriptive". It involves interpretation and explanation. However, one of the key ways that it differs from thematic analysis is that it does not lend itself well to the more explicitly theoretical interpretation that arises from thematic analysis. So, for audiences or purposes that are more theory focused (such as health sociologists), ICA will not be a good choice.

As we have highlighted, there are some differences between our account and the methods others have described as ICA. Our motivation in writing this article was to reduce confusion for researchers new to using ICA to analyse their data, but we do recognise that adding yet another account of ICA to the literature may, perversely, have the opposite effect. Our experience in teaching and supervising qualitative research is that our approach to ICA does provide clear guidance at a level appropriate to beginning researchers. Our overall advice to those new to qualitative research is to use a method of data analysis that is described in the literature and that you find understandable and useable. We hope that this article provides a source to use for this purpose. Being explicit and transparent about methodology is important for rigour. Having a clear description of what ICA is, when it is useful and how to do it will help new researchers to ensure that they are analysing their data appropriately and increase the methodological rigor of their research.

Conflicts of interest and funding

Danya Vears acknowledges the infrastructure funding received from the Victorian State Government through the Operational Infrastructure Support (OIS) Program and support by the Australian Government through the Medical Research Future Fund, as part of the Genomics Health Futures Mission (Grant number 76749). The authors declare no conflicts of interest.

References

Ahuvia, A. (2001). Traditional, interpretive, and reception based content analyses: Improving the ability of content analysis to address issues of pragmatic and theoretical concern. *Social Indicators Research*, *54*(2), 139–172. https://doi.org/10.1023/a:1011087813505

Bennett, D., Barrett, A., & Helmich, E. (2019). How to analyse qualitative data in different ways. *The Clinical Teacher*, *16*(1), 7–12. https://doi.org/10.1111/tct.12973

Bloor, M., & Wood, F. (2006). *Keywords in qualitative methods*. Sage. https://doi.org/10.4135/9781849209403

Boyatzis, R. E. (1998). *Transforming qualitative information*. Sage.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Cavanagh, S. (1997). Content analysis: Concepts, methods and applications. *Nurse Researcher*, *4*(3), 5–16. https://doi.org/10.7748/nr.4.3.5.s2

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Sage. https://doi.org/10.7748/nr.13.4.84.s4

Corbin, J., & Strauss, A. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage. https://doi.org/10.4135/9781452230153

Denzin, N. K., & Lincoln, Y. S. (2017). *The Sage handbook of qualitative research* (5th ed.). Sage.

Downe-Wamboldt, B. (1992). Content analysis: Method, applications, and issues. *Health Care for Women International*, *13*(3), 313–321. https://doi.org/10.1080/07399339209516006

Elo, S., & Kyngas, H. (2008). The qualitative content analysis process. *The Journal of Advanced Nursing*, *62*(1), 107–115. https://doi.org/10.1111/j.1365-2648.2007.04569.x

Erlingsson, C., & Brysiewicz, P. (2017). A hands-on guide to doing content analysis. *African Journal of Emergency Medicine*, *7*(3), 93–99. https://doi.org/10.1016/j.afjem.2017.08.001

Graneheim, U. H., & Lundman, B. (2004). Qualitative content analysis in nursing research: Concepts, procedures and measures to achieve trustworthiness. *Nurse Education Today*, *24*(2), 105–112. https://doi.org/10.1016/j.nedt.2003.10.001

Hansen, E. C. (2006). *Successful qualitative health research*. Allen & Unwin.

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, *15*(9), 1277–1288. https://doi.org/10.1177/1049732305276687

Kleinheksel, A. J., Rockich-Winston, N., Tawfik, H., & Wyatt, T. R. (2020). Demystifying content analysis. *American Journal of Pharmaceutical Education*, *84*(1), 127–137. https://doi.org/10.5688/ajpe7113

Krippendorf, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). Sage.

Liamputtong, P. (2020). *Qualitative research methods* (5th ed.). Oxford University Press.

Lichtman, M. (2014). *Qualitative research for the social sciences*. Sage. https://doi.org/10.4135/9781544307756

Morse, J. M. (1994). "Emerging from the data": The cognitive processes of analysis in qualitative inquiry. In J. M. Morse (Ed.), *Critical issues in qualitative research methods* (pp. 23–43). Sage.

Polit, D. F., & Beck, C. T. (2004). *Nursing research: Principles and methods*. Lippincott Williams & Wilkins.

Schick-Makaroff, K., MacDonald, M., Plummer, M., Burgess, J., & Neander, W. (2016). What synthesis methodology should I use? A review and analysis of approaches to research synthesis. *AIMS Public Health*, *3*(1), 172–215. https://doi.org/10.3934/publichealth.2016.1.172

Sousa, D. (2014). Validation in qualitative research: General aspects and specificities of the descriptive phenomenological method. *Qualitative Research in Psychology*, *11*(2), 211–227. https://doi.org/10.1080/14780887.2013.853855

Thayer, A., Evans, M., McBride, A., Queen, M., & Spyridakis, J. (2007). Content analysis as a best practice in technical communication research. *Journal of Technical Writing and Communication*, *37*(3), 267–279. https://doi.org/10.2190/TW.37.3.c

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, *8*(1), Article 45. https://doi.org/10.1186/1471-2288-8-45

Thorne, S., Kirkham, S. R., & O'Flynn-Magee, K. (2004). The analytic challenge in interpretive description. *International Journal of Qualitative Methods*, *3*(1), 1–11. https://doi.org/10.1177/160940690400300101

Vaismoradi, M., Turunen, H., & Bondas, T. (2013). Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. *Nursing & Health Sciences*, *15*(3), 398–405. https://doi.org/10.1111/nhs.12048

Vears, D. F., Delany, C., Massie, J., & Gillam, L. (2016). Why do parents want to know their child's carrier status? A qualitative study. *Journal of Genetic Counseling*, *25*(6), 1257–1266. https://doi.org/10.1007/s10897-016-9964-7

Weber, R. P. (1990). *Basic content analysis* (2nd ed.). Sage. https://doi.org/10.4135/9781412983488

Zhang, Y., & Wildemuth, B. (2009). Qualitative analysis of content. In B. Wildemuth (Ed.), *Applications of social research methods to questions in information and library science*. Libraries Unlimited.

Zolnoori, M., Balls-Berry, J. E., Brockman, T. A., Patten, C. A., Huang, M., & Yao, L. (2019). A systematic framework for analyzing patient-generated narrative data: Protocol for a content analysis. *JMIR Research Protocols*, *8*(8), 13914. https://doi.org/10.2196/13914