

## Qualitative research in health care

# Analysing qualitative data

Catherine Pope, Sue Ziebland, Nicholas Mays

**This is the second in a series of three articles**

Department of Social Medicine, University of Bristol, Bristol BSS 2PR

Catherine Pope  
*lecturer in medical sociology*

ICRF General Practice Research Group, University of Oxford, Institute of Health Sciences, Oxford OX3 7LS

Sue Ziebland  
*senior research fellow*

Social Policy Branch, The Treasury, PO Box 3724, Wellington, New Zealand  
Nicholas Mays  
*health adviser*

Correspondence to: C Pope  
c.pope@bristol.ac.uk

Series editors: Catherine Pope and Nicholas Mays

BMJ 2000;320:114-6

Contrary to popular perception, qualitative research can produce vast amounts of data. These may include verbatim notes or transcribed recordings of interviews or focus groups, jotted notes and more detailed “field-notes” of observational research, a diary or chronological account, and the researcher’s reflective notes made during the research. These data are not necessarily small scale: transcribing a typical single interview takes several hours and can generate 20-40 pages of single spaced text. Transcripts and notes are the raw data of the research. They provide a descriptive record of the research, but they cannot provide explanations. The researcher has to make sense of the data by sifting and interpreting them.

### Relation between analysis and qualitative data

In much qualitative research the analytical process begins during data collection as the data already gathered are analysed and shape the ongoing data collection. This sequential analysis<sup>1</sup> or interim analysis<sup>2</sup> has the advantage of allowing the researcher to go back and refine questions, develop hypotheses, and pursue emerging avenues of inquiry in further depth. Crucially, it also enables the researcher to look for deviant or negative cases; that is, examples of talk or events that run counter to the emerging propositions or hypotheses and can be used to refine them. Such continuous analysis is almost inevitable in qualitative research: because the researcher is “in the field” collecting the data, it is impossible not to start thinking about what is being heard and seen.

### The analysis

None the less there is still much analytical work to do once the researcher has left the field. Textual data (in the form of fieldnotes or transcripts) are explored using some variant of content analysis. In general, qualitative research does not seek to quantify data. Qualitative sampling strategies do not aim to identify a statistically representative set of respondents, so expressing results in relative frequencies may be misleading. Simple counts are sometimes used and may provide a useful summary of some aspects of the analysis. In most qualitative analyses the data are preserved in their textual form and “indexed” to generate or develop analytical categories and theoretical explanations.

Qualitative research uses analytical categories to describe and explain social phenomena. These categories may be derived inductively—that is, obtained gradually from the data—or used deductively, either at the beginning or part way through the analysis as a way of approaching the data. Deductive analysis is less common in qualitative research but is increasingly being used, for example in the “framework approach”

### Summary points

Qualitative research produces large amounts of textual data in the form of transcripts and observational fieldnotes

The systematic and rigorous preparation and analysis of these data is time consuming and labour intensive

Data analysis often takes place alongside data collection to allow questions to be refined and new avenues of inquiry to develop

Textual data are typically explored inductively using content analysis to generate categories and explanations; software packages can help with analysis but should not be viewed as short cuts to rigorous and systematic analysis

High quality analysis of qualitative data depends on the skill, vision, and integrity of the researcher; it should not be left to the novice

described below. The term grounded theory is used to describe the inductive process of identifying analytical categories as they emerge from the data (developing hypotheses from the ground or research field upwards rather defining them a priori).<sup>3</sup> Initially the data are read and reread to identify and index themes and categories: these may centre on particular phrases, incidents, or types of behaviour. Sometimes interesting or unfamiliar terms used by the group studied can form the basis of analytical categories. Becker and Geer’s classic study of medical training uncovered the specialist use of the term “crock” to denote patients who were seen as less worthwhile to treat by medical staff and students.<sup>4</sup>

All the data relevant to each category are identified and examined using a process called constant comparison, in which each item is checked or compared with the rest of the data to establish analytical categories. This requires a coherent and systematic approach. The key point about this process is that it is inclusive; categories are added to reflect as many of the nuances in the data as possible, rather than reducing the data to a few numerical codes. Sections of the data—such as discrete incidents—will typically include multiple themes, so it is important to have some system of cross indexing to deal with this. A number of computer software packages have been developed to assist with this process (see below).

Indexing the data creates a large number of “fuzzy categories” or units.<sup>5</sup> Informed by the analytical and theoretical ideas developed during the research, these categories are further refined and reduced in number



LANE PAYNE

by grouping them together. It is then possible to select key themes or categories for further investigation—typically by “cutting and pasting”—that is, selecting sections of data on like or related themes and putting them together. Paper systems for this (using multiple photocopies, cardex systems, matrices, or spreadsheets), although considered somewhat old fashioned and laborious, can help the researcher to develop an intimate knowledge of the data. Word processors can also facilitate data searching, and split screen functions make this a particularly appealing method for sorting and copying data into separate files.

### Software packages designed to handle qualitative data

Several software packages designed for qualitative data analysis enable complex organisation and retrieval of data. Among the most widely used are QSR NUD\*IST and ATLAS.ti.<sup>6 7</sup> This evolution has been welcomed as an important development with the potential to improve the rigour of analysis.<sup>8</sup> Such software can allow basic “code and retrieval” of data, and more sophisticated analysis using algorithms to identify co-occurring codes in a range of logically overlapping or nesting possibilities, annotation of the text, or the creation and amalgamation of codes. Some packages can be used to make theoretical links or search for “disconfirming evidence” (for example, by using boolean operators such as “or,” “and,” “not”). The Hypersoft package uses “hyperlinks” to capture the conceptual links which are observed between sections of the data; this can protect the narrative structure of the data to avoid the problem of decontextualisation or data fragmentation.<sup>9</sup>

Using software to help with the more laborious side of analysis has many potential benefits, but some caution is advisable. The prospect of computer assisted analysis may persuade researchers (or those who fund them) that they can manage much larger amounts of data and increase the apparent “power” of their study. However, qualitative studies are not designed to be representative in terms of statistical generalisability, and they may gain little from an expanded sample size except a more

cumbersome dataset. The sample size should be directed by the research question and analytical requirements, such as data saturation, rather than by the available software. In some circumstances, a single case study design may be the most successful way of generating theory. Furthermore, using a computer package may not make the analysis less time consuming,<sup>10</sup> although it may show that the process is systematic.

### Taking the analysis forward—the role of the researcher

A computer package may be a useful aid when gathering, organising, and reorganising data and helping to find exceptions, but no package is capable of perceiving a link between theory and data or defining an appropriate structure for the analysis. To take the analysis beyond the most basic descriptive and counting exercise requires the researcher’s analytical skills in moving towards hypotheses or propositions about the data.

One way of performing this next stage is called analytic induction. This involves an iterative testing and retesting of theoretical ideas using the data. Bloor described his use of this procedure in some detail (box).<sup>11</sup> In essence, the researcher examines a set of cases, develops hypotheses or constructs, and examines further cases to test these propositions.

#### Inter-rater reliability

Some researchers have found that the use of more than one analyst can improve the consistency or reliability of analyses.<sup>5 12 13</sup> However, the appropriateness of the concept of inter-rater reliability in qualitative research is contested.<sup>14</sup> None the less there may be merit in involving more than one analyst in situations where researcher bias is especially likely to be perceived to be a problem—for example, where social scientists are investigating the work of clinicians. In a study of diagnosis in cardiology, Daly et al developed a modified form of qualitative analysis involving external researchers and the cardiologists who had managed the patients. The researchers identified the

#### Analysis

Stages in the analysis of fieldnotes in a qualitative study of ear, nose, and throat surgeons’ disposal decisions for children referred for possible tonsillectomy and adenoidectomy (with examples)<sup>11</sup>:

- (1) Provisional classification—for each surgeon all cases categorised according to disposal category used (tonsillectomy and adenoidectomy or adenoidectomy alone)
  - (2) Identification of features of provisional cases—common features of cases in each disposal category identified (most tonsillectomy and adenoidectomy cases found to have three main clinical signs)
  - (3) Scrutiny of deviant cases—include in (2) or modify (1) to accommodate deviant cases (tonsillectomy and adenoidectomy performed when only two of three signs present)
  - (4) Identification of shared features of cases—features common to other disposal categories (history of several episodes of tonsillitis)
  - (5) Derivation of surgeons’ decision rules—from the features common to cases (case history more important than physical examination)
  - (6) Derivation of surgeons’ search procedures (for each decision rule)—the particular clinical signs looked for by each surgeon
- Repeat steps (2) to (6) for each disposal category

### Five stages of data analysis in the framework approach

- *Familiarisation*—immersion in the raw data (or typically a pragmatic selection from the data) by listening to tapes, reading transcripts, studying notes and so on, in order to list key ideas and recurrent themes
- *Identifying a thematic framework*—identifying all the key issues, concepts, and themes by which the data can be examined and referenced. This is carried out by drawing on a priori issues and questions derived from the aims and objectives of the study as well as issues raised by the respondents themselves and views or experiences that recur in the data. The end product of this stage is a detailed index of the data, which labels the data into manageable chunks for subsequent retrieval and exploration
- *Indexing*—applying the thematic framework or index systematically to all the data in textual form by annotating the transcripts with numerical codes from the index, usually supported by short text descriptors to elaborate the index heading. Single passages of text can often encompass a large number of different themes, each of which has to be recorded, usually in the margin of the transcript
- *Charting*—rearranging the data according to the appropriate part of the thematic framework to which they relate, and forming charts. For example, there is likely to be a chart for each key subject area or theme with entries for several respondents. Unlike simple cut and paste methods that group verbatim text, the charts contain distilled summaries of views and experiences. Thus the charting process involves a considerable amount of abstraction and synthesis
- *Mapping and interpretation*—using the charts to define concepts, map the range and nature of phenomena, create typologies and find associations between themes with a view to providing explanations for the findings. The process of mapping and interpretation is influenced by the original research objectives as well as by the themes that have emerged from the data themselves

main aspects of the consultations that seemed to be related to the use of echocardiography, and they developed criteria which other analysts could use to assess the raw data. The cardiologists then independently assessed each case using the raw data in order to produce an account of how and why a test was or was not ordered and with what consequences. The assessments of the cardiologists and researchers were compared statistically and the level of agreement was shown to be good. Where there was disagreement between the original researchers' analysis and that of the cardiologist, a further researcher repeated the analysis and any remaining discrepancies were resolved by discussion between the researchers and the cardiologists. Although there was an element of circularity in part of this lengthy process (in that the formal criteria used by the cardiologists were derived from the initial researchers' analysis) and it involved the derivation of quantitative gradings and statistical analysis of inter-rater agreement that are unusual in a qualitative study, this process meant that clinical critics could not argue that the findings were simply based on the subjective judgments of an individual researcher.

### Applied qualitative research

The framework approach has been developed in Britain specifically for applied or policy relevant qualitative research in which the objectives of the investigation are typically set in advance and shaped by the information requirements of the funding body (for example, a health authority).<sup>15</sup> The timescales of applied research tend to be short and there is often a need to link the analysis with quantitative findings. For these reasons, although the framework approach reflects the original

accounts and observations of the people studied (that is, "grounded" and inductive), it starts deductively from pre-set aims and objectives. The data collection tends to be more structured than would be the norm for much other qualitative research and the analytical process tends to be more explicit and more strongly informed by a priori reasoning (box).<sup>6</sup> The analysis is designed so that it can be viewed and assessed by people other than the primary analyst.

### Conclusions

Analysing qualitative data is not a simple or quick task. Done properly, it is systematic and rigorous, and therefore labour-intensive and time-consuming. Fielding contends that "good qualitative analysis is able to document its claim to reflect some of the truth of a phenomenon by reference to systematically gathered data," in contrast, "poor qualitative analysis is anecdotal, unreflective, descriptive without being focused on a coherent line of inquiry."<sup>16</sup> At its heart, good qualitative analysis relies on the skill, vision and integrity of the researcher doing that analysis, and as Dingwall et al have pointed out, this requires trained, and, crucially, experienced researchers.<sup>17</sup>

### Further reading

- Bryman A, Burgess R. eds. *Analysing qualitative data*. London: Routledge, 1993
- Miles M, Huberman A. *Qualitative data analysis*. London: Sage, 1984

The views expressed in this paper are those of the authors and do not necessarily reflect the views of the New Zealand Treasury, in the case of NM. The Treasury takes no responsibility for any errors or omissions in, or for the correctness of the information contained in this article.

- 1 Becker HS. *Sociological work*. London: Allen Lane, 1971.
- 2 Miles M, Huberman A. *Qualitative data analysis*. London: Sage, 1984.
- 3 Glaser BG, Strauss AL. *The discovery of grounded theory*. Chicago: Aldine, 1967.
- 4 Becker HS, Geer B. Participant observation: the analysis of qualitative field data. In: Burgess RG, ed. *Field research: a sourcebook and field manual*. London: Allen and Unwin, 1982.
- 5 Perry S. *Living with multiple sclerosis*. Aldershot: Avebury, 1994.
- 6 Richards T, Richards L. *QSR NUD\*IST, version 3.0*. London: Sage, 1994.
- 7 Muhr T. *ATLAS.ti for Windows*. Berlin: Scientific Software Development, 1997.
- 8 Kelle U, ed. *Computer-aided qualitative data analysis: theory, methods and practice*. London: Sage, 1995.
- 9 Dey I. *Qualitative data analysis: a user friendly guide for social scientists*. London: Routledge, 1993.
- 10 Lee R, Fielding N. User's experiences of qualitative data analysis software. In: Kelle U, ed. *Computer aided qualitative data analysis: theory, methods and practice*. London: Sage, 1995.
- 11 Bloor M. On the analysis of observational data: a discussion of the worth and uses of inductive techniques and respondent validation. *Sociology* 1978;12:545-52.
- 12 Daly J, McDonald I, Willis E. Why don't you ask them? A qualitative research framework for investigating the diagnosis of cardiac normality. In: Daly J, McDonald I, Willis E, eds. *Researching health care: designs, dilemmas, disciplines*. London: Routledge, 1992:189-206.
- 13 Waitzkin H. *The politics of medical encounters*. New Haven: Yale University Press, 1991.
- 14 Armstrong D, Gosling A, Weinman J, Marteau T. The place of inter-rater reliability in qualitative research: an empirical study. *Sociology* 1997;31:597-606.
- 15 Ritchie J, Spencer L. Qualitative data analysis for applied policy research. In: Bryman A, Burgess R, eds. *Analysing qualitative data*. London: Routledge, 1993:173-94.
- 16 Fielding N. Ethnography. In: Fielding N, ed. *Researching social life*. London: Sage, 1993:155-71.
- 17 Dingwall R, Murphy E, Watson P, Greatbatch D, Parker S. Catching goldfish: quality in qualitative research. *J Health Serv Res Policy* 1998;3:167-72.

This article is taken from the second edition of *Qualitative Research in Health Care*, edited by Catherine Pope and Nicholas Mays, published by BMJ Books