



## Organizing, storing, and analysing qualitative research information in a computer database

Jan Mainz & Frede Olesen

To cite this article: Jan Mainz & Frede Olesen (1997) Organizing, storing, and analysing qualitative research information in a computer database, Scandinavian Journal of Primary Health Care, 15:1, 7-9, DOI: [10.3109/02813439709043421](https://doi.org/10.3109/02813439709043421)

To link to this article: <https://doi.org/10.3109/02813439709043421>



© 1997 Informa UK Ltd All rights reserved:  
reproduction in whole or part not permitted



Published online: 12 Jul 2009.



Submit your article to this journal [↗](#)



Article views: 6652



View related articles [↗](#)

# Organizing, storing, and analysing qualitative research information in a computer database

Jan Mainz and Frede Olesen

The Research Unit for General Practice, University of Aarhus, Denmark.

Received June 1995. Accepted November 1995.

Scand J Prim Health Care 1997;15:7-9. ISSN 0281-3432

**Objective** – To describe a method to handle data in qualitative research.

**Design** – Description of the data handling process in a personal computer database.

**Setting** – The method was used in a Danish qualitative study with more than 100 interviews on quality of care.

**Outcome measures** – The author's own description and assessment of the system.

**Results** – Storing information in a database makes it possible to store and retrieve information from qualitative research divided in themes and subthemes. The themes and subthemes are based on an interview guide and the succeeding development of theory and new themes. The comprehensive sorting

systems in a database enable the researcher to keep an overview of his data while analysing and reporting.

**Conclusion** – Storing, analysing, and retrieving data from qualitative research demands careful planning. A computer database may be a good help in this process if the researcher will not use more complex and comprehensive software for his analysis.

**Key words:** qualitative research, research methods, software, data analysis.

*Frede Olesen, MD, PhD, The Research Unit for General Practice, University of Aarhus, Høegh-Gulbergs Gade 8, DK-8000 Århus C, Denmark.*

Qualitative research is often based on data from interviews (1). Before the analysis, the interview is often transcribed to paper word by word, or the analysis is performed by repetitive listening to tapes from the interview (2,3).

During the analysis the researcher describes themes, possible subthemes, and associated memos in terms of quotations from the interviewer, partially based on the interview guide, which was the basis of the research interview. The themes are further developed through the analysis. Following the initial analysis, the qualitative data are analysed in depth and categorized. Qualitative theories and hypotheses are built, based on evolved themes, subthemes, and quotations.

The research report states developed theories and hypotheses supported by themes, subthemes, and quotations. All data are sorted, based on themes and subthemes. In the research report, the developed theory is supported by one or more quotations from the interview (4).

The quality of the analysis depends, among other things, on the researcher's ability to structure his raw data in a retrievable manner that can easily be used through the writing and reporting phases of the research.

Standard computer software has been developed to help researchers in this process, but first generations almost all demanded that qualitative data were transcribed to paper (5,6). Newer generations of standard software have been developed, but they may have the

disadvantage of great complexity. Some researchers have used normal text editors to handle data including themes and quotes, but these programs do not have sufficient possibilities for sorting and storing in an easily retrievable form.

This paper aims at describing a method to handle qualitative research data in a standard computer database software.

## METHODS

A computer database has indata fields which can eventually be sorted and transcribed to paper in a user-defined random manner, and most new databases also have a large text field for longer qualitative quotations. A commonly used program is Paradox<sup>R</sup>.

Using this program we have defined the following fields: patient identification, theme, subtheme 1, subtheme 2, and a field for quotations. The first fields are normal database fields that we have defined with space for a maximum of six letters. The last field is for quotations of arbitrary length.

Through macro-facilities in the program, we have chosen some standard ways of sorting the data, based on patient identification, themes, or subthemes.

It is essential for the researcher to use themes in a well-defined and reproducible way. Researchers, therefore, often have supplementary definitions and an auxiliary text. The program helps researchers to store these

as themes and subthemes, including different definitions and auxiliary text. They are stored in a so-called reference file, in which during the analysis you can look up a list of your themes while making input of research data in the main file.

This reference file can be brought onto the screen in a pop-up picture, and the file's contents can be defined partly before the analysis or can be worked out during the analysis.

When the researcher has finished the input phase, the program can sort the data. This process is the main advantage of the program because the sorting makes it possible to transcribe data in a hierarchic manner, which is essential to many qualitative research projects.

#### DATA INPUT

Data input can be done from tape-transcribed data, typed interview data, or interview data solely based on listening to tapes. Using typed interview data, page number and line number are written on the indata. Using tapes solely by listening, tape number and position are based on the tape counter.

Data can be entered randomly with regard to which patient, theme or quotation you want to put into the database because the sorting allows for later necessary data structuring. At the beginning of each quote, you refer to the page and line number concerning typed paper data, or to the number of the tape and the tape recorder's marked position on the tape. This makes it possible to retrieve the original data. Fields with theme and subtheme are then notified. An infinite number of quotations per patient can be quoted in terms of patient interview quotation and/or researcher comments. Editing and correction can be done in any step of the process.

Data input can be done in two ways:

- with the screen set up in normal form, i.e. all data attached to one quotation are on the same screen (Fig. 1),
- in a table form, i.e. all data from one quotation are placed on one long line and data from identical fields are in columns.

During data entry you always start with the patient identification (number), a theme, perhaps a subtheme, and then a quote. As mentioned, these themes can be caught up from a reference file.

#### DATA ANALYSIS, SORTING, AND REPORTING

Data may be transcribed for any given theme in one group, or you may transcribe all data from a given patient.

We have been working with the following sortings:

Fig. 1. Copy of the data entry screen.

- a) data sorted by patient, theme, subtheme 1, and subtheme 2
- b) data sorted by theme, subtheme 1, subtheme 2, and patient
- c) data sorted by subtheme 1, subtheme 2, theme, and patient
- d) data sorted by subtheme 2, subtheme 1, theme, and patient
- e) data sorted by the first letter in the quote, patient, theme, subtheme 1, and subtheme 2.

After categorizing the data, the researcher can use this directly on the screen or it can be transcribed to paper copies.

During the analysis the researcher can use the screen set up in a formula form or in a table form as wished, and data can be browsed through with the page-up, page-down key.

Apart from the described fields, we have used several additional database fields in a specific project (7). These fields have been used for more quantitative descriptions of details in the material and as part of good data documentation. A database program is here again a good working tool because fields can be defined and deleted as wished.

Our paper transcriptions have been predefined in four forms:

- a) line no 1 gives patient number, theme, subtheme 1, and subtheme 2. The next line gives the quote, but not more than 3 lines per quote. Afterwards comes the next data, etc.
- b) the same as a, but not more than 10 lines per quote per theme
- c) the same as b, but not more than 20 lines per quote per theme

- d) the same, but the quote has infinite length in the transcribing.

The advantage of the short quotes with few lines is that it is easy to get an overview regarding data which is an advantage in the report or paper writing.

#### *Backup procedure*

After each working session we make a backup of the data onto a floppy disk, which is stored physically in a different place from the computer. A special macro routine for this function is prepared. The database allows you to limit admission to the database by a password.

#### *Necessary computer power and software*

The program needs an IBM compatible computer with an Intel-386 processor or more. We have used the UK version 4.0 of the database Paradox, which can be bought for DOS, Windows and OS2. Apart from our purpose, the program is very appropriate for storing all types of information files in both qualitative and quantitative projects.

#### FINAL COMMENT

One of the authors (JM) has used the program and has tried to use other assistant programs to analyse his qualitative data based on nearly 100 patient interviews. The described model proved to be the most flexible, safe and appropriate for handling the data. The well-structured indata and reporting functions in the program were of great help in the reporting phase of the study.

However, it must be strongly stressed that a large group of very sophisticated tools for analysing qualitat-

ive data have been developed. In our research milieu we have experience in using the program Textbase Alpha, and with the advanced Australian software package NUD.IST. However, it is our experience that the described method is good for qualitative projects with relatively simple hierarchic structure in the data. In any case, electronic storage of qualitative data means better documentation of the original research data than can be managed if manual handling of the data is used.

#### REFERENCES

1. Kvale S. Det kvalitative forskningsinterview. Ansættelser til en fænomenologisk hermeneutisk forståelsesform. Kvalitative modeller i samfundsforskningen (The qualitative research interview. Tendencies to a phenomenological hermeneutic way of understanding). In: Qualitative models in social research. Crach T, Larsen PK, Ruper O, editors. København: Nyt fra samfundsvidenskaberne, 1979.
2. Kvale S. Om tolkningen af kvalitative forskningsinterviews (Interpretation of qualitative research interviews). Tidsskr Nor Foren Paed Forsk 1984;4:55-6.
3. Lunde IM. Patienters egenvurdering (Patients' self evaluations). Århus: FADL, 1990.
4. Strauss AL, Corbin J. Basics of qualitative research. Grounded theory procedures and techniques. London: Sage Publications, 1990.
5. Miles M, Huberman M. Qualitative data analysis: a sourcebook of new methods. Beverly Hills: Sage Publications, 1994.
6. Miles M, Weitzman E. Computer programmes for qualitative data analyses. Beverly Hills: Sage Publications, 1995.
7. Mainz J. Problemidentifikation og kvalitetsvurdering i sundhedsvæsenet. Teori. Metode. Resultater (Problem identification and quality evaluation in the health system. Theory. Method. Results). København: Munksgaard, 1996.