Qualitative Research Methods: A Data Collector's Field Guide

Module 5

Data Documentation and Management: Organizing and Storing Your Data

DATA MANAGEMENT

F A M I L Y   H E A L T H   I N T E R N A T I O N A L

Data Documentation and Management: Organizing and Storing Your Data

Because they capture the thoughts and experiences of individual people, every set of qualitative data collected (from every participant observation event, interview, and focus group) is distinct. In addition, individual researchers inevitably have differences in style that affect how data are managed in the field, and different locations have unique logistical constraints. Thus, systematically comparing and analyzing qualitative data in raw form is challenging.

Organizing data in a rigorous, standardized way is essential to their security and to the validity of the study results, however. Consistency is important for every study but is especially crucial for large, team-based projects involving extensive amounts of data located in multiple sites.

Your study's local principal investigator or data manager will set up a data management system that is specific to your site. This module suggests a general method for the complex task of systematically managing qualitative data, covering the following topics:

- Converting Raw Data to Computer Files
- Organizing Data Storage
- Suggested Reading
- Data Archiving Steps
- Data Management Checklist

For a specific study, the procedures outlined in this module may be followed exactly or adapted, as the principal investigator or data manager sees fit. The Tools for Data Managers appendix, page 105, contains a detailed example of a transcription protocol and data archive model, which should be of interest to your study's data management personnel but is likely less relevant for field workers.

Converting Raw Data to Computer Files

What do we do with audio recordings?

In-depth interviews and focus groups are tape-recorded whenever possible. Preparing these recorded data for analysis requires transcribing all tapes and typing the transcriptions into computer files. Before transcription, the tabs on the tapes should be punched to prevent them from being recorded over. Next, backup copies of the tapes should be made. The backup copies should be securely stored in a separate location from the original tapes.

What happens to the handwritten field notes?

Participant observers, focus group facilitators, and interviewers take handwritten notes to document a wide range of information, including:

- casual and structured observations
- verbatim quotes
- paraphrases of participant responses
- interview and focus group backup documentation
- the researcher's questions
- questions, conclusions, and observations discussed during the staff debriefing sessions

These notes are written on standardized forms, the interview or focus group question guide, or field notebooks, according to the situation.

For focus groups and interviews, after transcribing all relevant recordings, the transcriptionist types up the interviewer's or focus group moderator's corresponding handwritten field notes. These typed field notes could either be appended to the transcript within the same file or kept in a separate file (the principal investigator or data manager would make this decision). Whatever the case, the typed field notes provide contextual information that could enhance the researchers' understanding of the transcript and therefore need to be easily identifiable as part of the same data collection event. Expanded notes from each participant observation event should be typed as separate computer files.

When should we begin typing the data?

Transcription of recordings and typing of field notes should begin as soon as possible after the data collection event. Tapes of interviews and focus groups should be processed as soon as they are archived, rather than allowed to accumulate. Field notes should be typed as soon as the data collector has expanded them (if they are not typed while they are being expanded).

What does transcription involve?

To transcribe an audio recording, the transcriptionist listens to the tape and simultaneously writes down or types everything that is said on the tape. Nonverbal sounds (such as laughter, sirens, someone knocking on the door) are also often noted on the transcript. Transcription is performed by either the data collectors themselves or other staff hired specifically as typists or transcriptionists. When the transcriptionist is not the person who collected the recorded data, the interviewer or focus group moderator who did collect it should review completed transcripts for accuracy. Transcripts may also need to be translated into the language(s) of the organization sponsoring the study.

Does it matter what format is used for transcription?

Whether one person or many staff members are involved in transcription, it is highly recommended that everyone involved in a given study use a common format for transcribing all recordings. That is, there should be standard conventions for identifying the researcher and individual participants throughout the transcript. There should also be a uniform way of presenting information on the location, date, and type of data collection event.

These conventions should be detailed in a project-specific transcription protocol which precisely outlines procedures and formats for transcribing recorded data. A sample transcription protocol is provided in the Tools for Data Managers appendix, page 108, for your use or adaptation. Note, however, that this protocol reflects the specifications of the data-analysis software frequently used by the authors. The transcription protocol you develop for your own study should similarly reflect any formatting or other requirements of the software that you will use. Whether the research team or some other entity is in charge of data analysis, it is important to determine such requirements before even the first transcription is done. Doing so will save much work later. If you find that the software does not require any specific format, it is still important for you to design a systematic and consistent transcription protocol.

> **Naming transcript files**
>
> To avoid confusion, consider using the archival number as the computer file name for the transcript. As with the archival number, the file name should indicate the site name, method of data collection, participant category, and sequential number.
>
> For example, for your fifth individual interview with a health care service provider at your study's Capital City site, the naming convention might be:
>
> File name: CCIISP05.doc
>
> CC = Capital City
>
> II = In-depth interview
>
> SP = Health care service provider
>
> 05 = Fifth data collection event (either in this category, or sequentially for all events)

What do we name the computer files?

You should name the computer file for each interview and focus group transcript according to a standard convention specified in the transcription protocol. As previously noted, such conventions are project-specific. To avoid confusion, consider using the archival number as the computer file name for the transcript. As with the archival number, the file name should indicate the site name, method of data collection, participant category, and sequential number. (See the box above for a sample.)

What other information should we include in the computer files?

Each transcript or set of notes should begin with a standard header that indicates the archival number, site, data collector's name(s), date of data collection, data collection method, transcriber, translator, typist, and date of computer entry. (See the sample on page 86.)

Organizing Data Storage

Who organizes the data?

The site coordinator or principal investigator typically assigns a member of the research team or other staff member to be the data manager. This person is in charge of creating a system of organizing and archiving (that is, storing) the data, and for ensuring that all staff members follow these procedures for the duration of the study. This helps ensure the security and integrity of the data. It is also a good idea to have one person who always knows where all data are located.

How should the data be organized?

In the Tools for Data Managers appendix, page 105, a detailed sample protocol for organizing data (called the Model Data Archive Protocol) is provided for the data manager to use as a model. However, it is important that each project have a set of procedures that are appropriate for the site and that make sense to the research team. If a project has multiple sites, one system should be used in all sites to avoid confusion.

Sample Header for Typed Field Notes or Transcripts
Archival #: Site: Data collector(s): Date of data collection: Data collection method: Transcriber: Translator: Typist: Date of computer data entry: *Note: This is the same header information that you will have written at the top of your field notes, the interview and focus group guides, note-taker form, and debriefing form.*

What steps should we take before data collection?

A strategy that works well for keeping data organized, and which should be implemented even before data are collected, is to create packets of all necessary forms for each kind of event using large, heavy-duty envelopes, sometimes called “archival envelopes.” Then, when you are ready to do the data collection – for example, a focus group – you can go to the study office and get a focus group envelope containing the focus group guide, note-taker form, informed consent forms, debriefing form, reimbursement form, and all other materials you will need. At that time, the data manager should assign an archival number to the data collection event. (See the section on page 87, “What are an archival number and archival log?”) This strategy enables researchers to label all materials with the appropriate archival number before the event.

Where should we put the data after collection?

Centralizing data, or keeping them all together in one place, is a key element of managing data in an organized and systematic fashion. All physical data (notes, tape recordings, transcripts, etc.) should be kept in a locked filing cabinet or equally secure location. We recommend keeping all documents related to a given data collection event in a secure location at each field site, in one large, heavy-duty archival envelope per event with an archival information sheet. (See the Case Study Sample, page 88.) The envelope’s contents will include typed transcripts, expanded field notes, debriefing notes, handwritten versions of the notes, and possibly cassette tapes.

The data manager may also decide to store the original tapes in the envelope, but backup copies of tapes should be stored separately (or vice versa). Eventually, both copies of the tapes will be destroyed for reasons of confidentiality. Once you are instructed by the project manager or coordinator to destroy the tapes, it will be necessary to put documentation of this in the archival envelope.

Copies of all electronic files should be sent to the sponsoring organization and also maintained at the field site.

What are an archival number and archival log?

All data, both electronic and paper, are organized and identified according to archival numbers – numbers assigned in sequential order to each data collection event. The archival number is used to label all documentation related to a particular data collection event. The archival log is the list of sequential numbers assigned to each data collection event and is used to track data. (See the Case Study Sample, page 89, for an example of an archival log created to track interviews. For more information, see the sample Model Data Archive Protocol in the Tools for Data Managers appendix, page 105.)

Suggested Readings

McLellan E, MacQueen KM, Niedig J. Beyond the qualitative interview: data preparation and transcription. *Field Methods* 2003;15(1):63-84.

For additional information on this topic, refer to Chapter 5: Logistics in the Field and Chapter 6: Qualitative Data Analysis in these companion guides:

Qualitative Methods in Public Health: A Field Guide for Applied Research

Qualitative Methods: A Field Guide for Applied Research in Sexual and Reproductive Health

CASE STUDY

Sample Archival Information Sheet

Archival Information Sheet

Study: FP VCT Integration Archival # CCIISP01

Data type: (circle one)	Interview	Focus group	Participant observation
Indicate type of participant(s) or observational setting:	X FP Provider __ HIV Clinic Staff __ FP Users __ FP Non-users	__ FP Users (#________) __ FP Non-users (#________)	__ FP Clinic __ HIV Clinic
Gender (circle one): Female Male	Age: 50	Ethnicity: Bimou	Language of Data: Bimouti

Data in this envelope:		**Staff:**	**Name**	**Date**
Cassette tape	X	Data collector	Beatrice B.	7-6-04
Interview guide notes	X	Transcriber	Beatrice B.	11-6-04
FG note-taker notes		Translator	Samuel D.	14-6-04
FG guide notes		Typist	Samuel D.	15-6-04
FG debriefing notes				
Handwritten field notes	X			
Expanded field notes	X			
Handwritten transcript	X			
Translation	X			
Hardcopy of Electronic	X			

Data Sign-out:

Name	Item(s)/Purpose	Date Out	Date In
Beatrice B.	Handwr. transcript & Typed translation / Review translation	16-6-04	16-6-04

Sample Archival Log

Interview Data Tracking Form

Archival #	Category*	Interviewer	Transcriber	Translator	Typist	Date of Interview	Language of Interview	Sex	Age	Date to Manager
CCIISP01	SP	Beatrice B.	Beatrice B.	Samuel D.	Samuel D.	7-6-04	Bimouti	F	50	

* SP = Health care service provider
HP = Provider for HIV/AIDS-related services
FPU = Family planning users
FPN = Family planning non-users

DATA MANAGEMENT

STEPS

Data Archiving Steps

In the Field

1 Label all materials (tapes, guide, field notes, forms).

2 Punch the re-record tabs on each cassette after the interview/focus group is completed.

3 After data collection, return **all** materials to the large, heavy-duty envelope.

At the Research Office

4 Return all materials, including field notes and expanded field notes, to field office as soon as possible (no later than 48 hours).

5 Obtain the archival number if you did not do so before data collection.

6 Write the archival number on the archival envelope, **all** materials, and **all** pages of the question guide.

7 Complete the archival information sheet and place it in the envelope with the other items.

8 Duplicate tapes as soon as possible (within 24 to 48 hours).

9 Store materials in a secure location.

Translating and Transcribing Data

10 Sign out the duplicate audiotape (on the large envelope itself) and other necessary materials for transcription or translation (e.g., notes). **Do not take the entire envelope from the secure location.**

11 Conduct the translation or transcription.

12 Write the archival number on the translated or transcribed data.

13 Record new materials (e.g., diskette and hard copy of transcription or translation) on the archival information sheet.

14 Return **all** materials to the envelope.

15 Return the envelope to the secure data area and record the return of data per the data manager's procedures.

Data Transfer to Project Manager or Coordinator

16 Verify the accuracy of the translation and that transcription protocol has been followed.

17 Remove identifying information from the transcript (if necessary).

18 Send the transcript to the project manager or coordinator and send a copy to individuals designated in the study protocol. If you send the transcript electronically, we recommend that you password-protect it and place a hard copy in the archival envelope.

19 The project manager or coordinator should confirm receipt of data. Document this in writing and put it in the archival envelope.

20 Make sure **all** materials are returned to the archival envelope and return it to the secure data storage area.

21 Print any correspondence regarding a particular archival number and place it in the archival envelope.

Maintain security and confidentiality of data at all times.

CHECKLIST

Data Management Checklist: Items to Be Labeled and Placed in the Archival Envelope

Participant observation

- ❑ Handwritten notes, including maps and diagrams
- ❑ Expanded notes
- ❑ Archival information sheet
- ❑ Any additional documentation relevant to the event

In-depth interviews

- ❑ Consent form signed by the interviewer (and participant if obtaining written consent)
- ❑ Audiotape(s) of the interview
- ❑ Handwritten notes
- ❑ Expanded notes
- ❑ Transcripts (handwritten and typed, with translations)
- ❑ Archival information sheet
- ❑ Correspondence related to the data collection event

Focus groups

- ❑ Consent form signed by the focus group moderator (or participants if obtaining written consent)
- ❑ Audiotape(s) of the focus group
- ❑ Handwritten notes
- ❑ Seating chart
- ❑ Expanded notes
- ❑ Debriefing notes
- ❑ Archival information sheet
- ❑ Correspondence related to the data collection event