Data Management Plan Template:
Qualitative Health Sciences Research

Abstract

This data management template is meant to be used by health sciences researchers conducting qualitative research on human subjects. It includes guidance on data management best practices beginning with data collection through to data sharing.

Administrative Details

Template Author(s): Kevin B. Read, Shelley Peacock, Tracie Risling, University of Saskatchewan

Published: March 2, 2021

DOI: 10.5281/zenodo.4564537

Contact: Portage Network - portage@carl-abrc.ca, portagenetwork.ca

Version:

Version	Date	Changes
1.0	2021-03-02	Formatted for inaugural publication.

Data Collection

Outline the processes and procedures you will follow during the data collection process of your study.

What types of data will you be collecting?

Describe all types of data you will collect throughout the research process, with special attention paid to participant-driven data (e.g., written transcripts, video files, audio recordings, journals, art, photographs, etc.).

Will you be using any existing data from external sources or previous research?

If you will be combining original research data with existing, or previously used research data, describe those data here. Provide the name, location, URL, and date of the dataset(s) used. Describe any end-user license assigned to the data, or terms of use you must abide by.

What data collection instrument or scales will you use to collect the data?

Provide a description of any data collection instruments or scales that will be used to collect data. These may include but are not limited to questionnaires, interview guides, or focus group procedures. If using a pre-existing instrument or scale, provide the citation(s) in this section.

Is your data collected longitudinally or at a single point in time?

Describe the frequency in which you will be collecting data from participants. If you are performing narrative inquiry or longitudinal data collection for example, how often will you gather data from the same participants?

What is the time frame over which you are collecting data?

Provide an estimate of when you will begin and conclude the data collection process. List this information in the following format: YYYY/MM/DD - YYYY/MM/DD. If you do not know the exact dates, list YYYY/MM - YYYY/MM instead.

What is the geographic location within the context of the phenomenon/experience where data will be gathered?

Provide a description of the environment and geographic location of where data will be gathered, within the context of the study (e.g., hospital setting, long-term care setting, community). Include national, provincial, or municipal locations if applicable.

What steps will be involved in the data collection process?

Summarize the steps that are involved in the data collection process for your study. This section should include information about screening and recruitment, the informed consent process, information disseminated to participants before data collection, and the methods by which data is gathered.

In this section, consider including documentation such as your study protocol, interview guide, questionnaire, etc.

What software programs will you use to collect the data?

Include a description of any software that will be used to gather data. Examples may include but are not limited to word processing programs, survey software, and audio/video recording tools. Provide the version of each software program used if applicable.

What file formats will you be generating during the data collection phase?

List the file formats associated with each software program that will be generated during the data collection phase (e.g., .txt, .csv, .mp4, .wav).

Data Analysis

Outline the steps, materials, and methods that you will use to document how you will analyze the data collected in your study.

How will you document the changes you make to your data on a regular basis?

Provide a description of how you will track changes made to any data analysis files. Examples might include any audit trail steps, or versioning systems that you follow during the data analysis process.

What software will you be using to support your data analysis?

Include any software programs you plan to use to perform or supplement data analysis (e.g., NVivo, Atlas.ti, SPSS, SAS, R, etc.). Include the version if applicable.

What file formats will your data analysis files be saved in?

List the file formats associated with each analysis software program that will be generated in your study (e.g., .txt, .csv, .xsls, .docx).

What coding scheme or methodology will you use to analyze your data?

Include the coding scheme used to analyze your data -- consider providing a copy of your codebook. If other methods of analysis were performed, describe them here.

What quality assurance measures will be implemented to ensure the accuracy and integrity of the data?

Outline the steps that will be taken to ensure the quality and transparency during the data analysis process. In this section, describe procedures for cleaning data, contacting participants to clarify responses, and correcting data when errors are identified. Consider the principles of credibility, dependability, confirmability, and transferability as described in Lincoln and Guba, 1985 when completing this section.

Documentation and Metadata

This section is designed for you to provide information about your data, so that others will be able to better understand, interpret, and potentially re-use your data for secondary analysis.

What information about your research would someone need to know to reuse or interpret your data?

Consider what information might be useful to accompany your data if you were to share it with someone else (e.g., the study protocol, interview guide, codebook, information about software used, questionnaires, user guide for the data, etc.).

Are there metadata standards which you could use to describe your data?

Metadata standards can provide guidance on how best to document your data. If you do not know of any existing standards in your field, visit this website to search for available standards: https://fairsharing.org/.

Who is the target population being investigated?

Describe the participants whose lived experiences/phenomena are being studied in this project.

How is the population being sampled?

Provide a brief description of the sampling process undertaken in the study (e.g., purposive sampling, theoretical sampling).

Is the population being weighted?

Outline any weighting or representative sampling that is being applied in this study.

Are there any acronyms or abbreviations that will be used within your study?

Provide a glossary of any acronyms or abbreviations used within your study.

Storage and Backup

This section will ask you to outline how you will store and manage your data throughout the research process.

What are the storage requirements needed for your data?

Provide an estimate of how much data you will collect in the form of terabytes, gigabytes, or megabytes as needed. Include estimates for each data type if possible (e.g., 2 GB for video files, 500 MB for interview transcripts).

Where will your data be stored during the *data collection* phase?

Describe where your data will be stored while data is being gathered from participants (e.g., in a secure, password protected computer file, hard copies stored in locked filing cabinets, or institutional computer storage).

Where will your data be stored during the *data analysis* phase?

If different from the above, describe where your data will be stored while performing data analysis.

What backup measures will be implemented to ensure the safety of your data?

If different from the above, describe where your data will be stored while performing data analysis.

If your data contains confidential information, how will your storage method ensure the protection of this data?

Outline the procedures that will safeguard sensitive data collected during your study. This may include storing identifying data (consent forms) separately from anonymized data (audio files or transcripts), keeping files password protected and secure, and only providing access to investigators analyzing the data.

What file naming conventions will be used to save your data?

Provide examples of a consistent file naming convention that will be used for this study. Examples of file names might include the type of file, participant number, date of interaction, and/or study phase. Follow this guide for more information on file naming.

Preservation

Describe the steps that will ensure that your data will be available and usable for the foreseeable future after your study is complete.

Where will data be stored *after* the project is complete?

Describe where your data will be stored after project completion (e.g., in an institutional repository, external data repository, in secure, institutional computer storage, or external hard drive).

Who is responsible for managing the data after the study is complete?

Name the person(s) responsible for managing the data at the completion of the project. List their affiliation(s) and contact information.

Will your data be migrated to preservation formats?

Many proprietary file formats such as those generated from Microsoft software or statistical analysis tools can make the data difficult to access later on. Consider transforming any proprietary files into preservation-friendly formats to ensure your data can be opened in any program. Describe the process for migrating any data formats here.

How long do you intend to keep your data after the project is complete?

Provide details on how long you plan to keep your data after the project, and list any requirements you must follow based on Research Ethics Board guidelines, data use agreements, or funder requirements.

What procedures are in place to destroy the data after the retention period is complete?

Describe what steps will be taken to destroy study data. These steps may include shredding physical documents, making data unretrievable with support from your Information Technology department, or other personal measures to eliminate data files.

Ethics and Legal Compliance

Outline any ethical and legal implications placed on your research data.

How is the informed consent process carried out in your study?

Outline the information provided in your Research Ethics Board protocol, and describe how and when informed consent is collected during the data collection process. Examples include steps to gain written or verbal consent, re-establishing consent at subsequent interviews, etc.

Who holds the intellectual property rights to your data?

Provide the name, institutional affiliation, and contact information of the person(s) who hold intellectual property rights to the data.

What ethical guidelines or restraints are applicable to your data?

Describe any ethical concerns that may be associated with the data in this study. For example, if vulnerable and/or Indigenous populations are included as participants, outline specific guidelines that are being followed to protect them (e.g., OCAP, community advisory boards, etc.).

What legal restraints are applicable to your data (e.g., ownership)?

Provide details describing the legal restrictions that apply to your data. These restrictions may include, but are not limited to details about how your research data can be used as outlined by funder, institutional, or community agreements, among others.

What methods will be used to manage the risk of disclosure of participant information?

List all the steps that will be taken to remove the risk of disclosing personal information from study participants. Include information about keeping data safe and secure, and whether certain information will be removed from the data. If data is being anonymized or de-identified, specify the information type(s) being altered (e.g., names, addresses, dates, location).

Responsibilities and Resources

Indicate who will be working with the data at various stages, and describe their responsibilities.

What financial resources will you require for data management in this study?

Describe any financial resources that may be required to properly manage your research data. This may include, but not be limited to personnel, storage requirements, software, or hardware.

Who is the main contact and steward for the data collected in this study?

Provide the name(s), affiliation(s), and contact information for the main study contact.

Who will have access to your data throughout the project?

Provide the name(s), affiliation(s), contact information, and responsibilities of each study team member in relation to working with the study data. If working with institutional Information Technology members, statisticians, or other stakeholders outside your immediate team, provide their information as well.

Will any new members be added or responsibilities be transferred over the course of the study?

Describe the process by which new collaborators/team members will be added to the project, if applicable. Include the type(s) of responsibilities that may require new team members to be added during, or after the project is complete.

Data Sharing

Provide information about how you will make your data available and/or discoverable to the broader community.

Who are the likely users/benefitters of your data?

Describe the intended audience for your data.

What data can/will be shared at the end of the study?

Describe what data can/will be shared at the end of the study.

What restrictions are placed on your data that would prohibit it from being made publicly available?

Restrictions on data may include, but are not limited to, the sensitivity of the data, data being acquired under license, or data being restricted under a data use agreement. Describe what restrictions (if any) apply to your research data.

Where will you share your data?

Provide the location where you intend to share your data. This may be an institutional repository, external data repository, or through your Research Ethics Board, among others.

If you have collected restricted data, what steps would someone requesting your data need to follow in order to access it?

If your data is restricted, describe how a researcher could access that data for secondary analysis. Examples of these procedures may include completing a Research Ethics Board application, signing a data use agreement, submitting a proposal to a community advisory board, among others. Be as specific as possible in this section.

What license will you apply to your data?

Select a license that best suits the parameters of how you would like to share your data, and how you would prefer to be credited. See this resource to help you decide: https://creativecommons.org/choose/.