SOP 11: Data Management, Collection and Storage

The South African College of Applied Psychology Research Ethics Committee	
Title	SOP 11: Data Management, Collection and Storage
SOP No	SOP 11_SACAP REC_11.1
Date of approval	January 2022
Revision date	October 2021
Pages	6

1. COMPILATION AND AUTHORIZATION

Action	Designated person	Date	Signature
Compiled for SACAP by	Dr Malesa Kgashane	6 August 2021	
Amended for SACAP by	REC Office	26 October 2021	K. J. Young
Checked by	Academic Manager	11 November 2021	C.E. Ford
Authorised by	Academic Dean	22 January 2022	J.O. Lotter

2. DOCUMENT HISTORY

Date	Version no	Reason for revision
10 November 2017	1	Development of the document

3. ABBREVIATIONS AND DEFINITIONS

Abbreviation	Definition
SACAP REC	The South African College of Applied Psychology Research Ethics Committee
SOP	Standard Operating Procedure/s

4. PURPOSE OF THE SOP

The purpose of this SOP is to provide guidelines on the management of data as well as on the storage of such data. Data management includes design, collection, cleaning and management of all participants and other information, observations and measurements that are part of the research project. Through efficient and effective data management the risk of data error can be addressed, as data management should be considered during the planning, design and finalization of the research protocol. It should be fully understood that data management does not only address data collection and data capturing but includes the preparation of data, data analysis and publication, data archiving and lastly the destruction of the data. Key considerations to data management is that:

4.1. Appropriate scientific data gathering instruments should be used to provide plausible and reliable data;

4.2. The recorded data should be durable and appropriately referenced by the researcher;

4.3. The data must be retained for the period specified in the *SACAP Ethics Research Policy* Part 4 paragraph 1.3.3; and

4.4. Data reported in publications should be available for discussion and interrogation without breaching the confidentiality or anonymity of the respondent(s).

5. SCOPE

The scope of this document covers the establishment of the procedures to follow when initiating data management during a research project and the subsequent procedures to follow when data is stored and destruction or data banking thereof. This SOP needs to be read together with SACAP's Research and Ethics Policy which provide guidelines on research data storage.

6. RESPONSIBILITIES

All SACAP REC members and researchers issued with ethics approval and the REC Administrator should be aware of the procedure during the continuous review and recertification process.

7. PROCEDURE(S)

7.1. Identification and Description of Data:

7.1.1. The identification of the data the researcher wants to work with is important as it will address questions such as what type of data will be collected, why it is needed and very importantly how it will be used. Figure 1 is a decision route for the identification of the data.

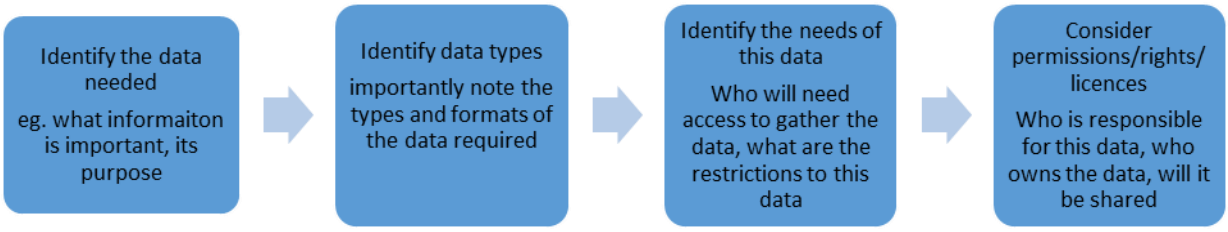

Figure 1: Identifying the data (Source: SOP-4.0.0-DM-121203 from Psycho-oncology Cooperative Research Group, The University of Sydney, Australia).

7.1.2. During the decision about the data that is required the lifecycle of the data or lifespan of the data should also be taken into consideration.

7.1.3. Identification of the data will also include consideration for the types and formats in which the data should be obtained such as numeric data or verbal transcripts, etc.

7.1.4. All questionnaires to be used should embody professional representation of SACAP and standard format.

7.1.5. All questionnaires or interview guides should have been approved by the supervisor or research group and submitted to the SACAP REC for review.

7.1.6. The different types and formats of the data that will be obtained have to be considered in terms of the sample and the ability of the selected participants to complete the instruments with truthful and reliable data that captures the situation at that moment in time.

7.1.7. The ability of the researcher to execute a particular research instrument should also be taken into consideration as the inability to correctly administer the instrument will lead to unethical data capturing sessions wasting respondents' time and resources.

7.1.8. Further consideration should be given to what the data will be needed for in particular who will need access to the data and what restrictions are there in terms of using the data.

7.1.8.1. It should be made very clear in the informed consent form what the data will be used for and the researcher should not go beyond this stipulation without further permission to do so.

7.1.8.2. It should also be indicated who will be working with the data and that access to the data will not be made available to others not authorised by the participant to access the data.

7.1.8.3. Should the data come from a particular source the time limitation on using the data should be adhered to, and if the data should be returned to the source.

7.1.9. Lastly in identifying the data, consideration should be given to the necessary permission to gather such data, but also to consider who owns the data and if the data will be shared in future with other researchers. Informed consent will need to be sought and obtained from the participants concerned.

7.2. **Identifying the Mechanisms to Capture the Data:**

7.2.1. The data collection methods should be outlined in detail in the methodology of the research proposal. These methods should include a step-by-step procedure of how the data will be collected and who will do what during this process. For example, if questionnaires are used, these will be distributed through a fieldworker to the sample of participants and collected from the participants by the fieldworker. The researcher on the first Monday of each month will then collect them from the fieldworker.

7.2.2. The procedures for each data collection instrument should be clearly stipulated.

7.2.3. If more than one instrument is used in different phases of a study the same procedure should be followed as described in 7.2.

7.2.4. The sequence of data gathering and instrument completion should be clear to form a picture of how the data gathering will be executed in the study.

7.3. **Outline the Infrastructure and Mechanisms to Store the Data:**

7.3.1. A researcher should have a clear set of guidelines in terms of how variables for numeric data were coded, which is part of the data management, collection and storage system applied by the researcher.

7.3.2. Other data storage systems that would be used such as spreadsheets, text documents, specific computer driver space (local drive), etc. should be specified.

7.3.3. When mechanisms for data storage are considered, the following questions should be asked for both physical and digital data storage:

7.3.3.1. Will storage of data be centralized or on-site where data is being gathered?

7.3.3.2. What is the timeline for data collection and storage?

7.3.3.3. How much data storage is needed?

7.3.3.4. How the system is secured (locked filing cabinets, password protected, etc.?)

7.3.3.5. Has a log system been created for cataloguing the movement of data?

7.3.3.6. Where is the data stored?

7.3.3.7. In what format will the data be stored and why is this the chosen format?

7.3.3.8. Is any specific software required to read, analyse or process the data?

7.3.3.9. Who is responsible for the data?

7.3.3.10. What are your institution's data management policies?

7.4. **Data Security Imperatives Include the Following Actions:**

7.4.1. Build and maintain a secure network system in which confidential passwords and documentation of an audit trail to capture changes to information is clearly articulated in the research proposal.

7.4.2. Protect participant information through de-identification of personal information and use of participant pseudonyms or participant IDs where necessary.

7.4.3. Maintain vulnerability management programmes, which include anti-virus software and ensuring that regular backups of the research data is made.

7.4.4. Implement a strong access control measure, which restricts access, and use unique IDs for each person who has permission to access the data and criteria for electronic signatures.

7.4.5. Regularly monitor and test networks.

7.4.6. Specific security and data management procedures for informed consent:

7.4.6.1. Signed participant consent forms are to be kept by the researcher or with the supervisors of the student. Should the student complete the study the informed consent forms should be handed to the supervisor if not already with the supervisor.

7.4.6.2. The signed participant consent forms should be stored separately and securely from the de-identified data for a period of 5 years.

7.4.6.3. If verbal consent is given by the participant, it should be recorded on the consent form by the researcher or fieldworker.

7.4.6.4. If delayed informed consent applied, this should also be noted on the database.

7.4.6.5. The person responsible for the management of the informed consent records should be clearly identified.

7.4.7. De-identification of participant data:

7.4.7.1. During the de-identification process, the researcher should remove all identifying information from the data to protect the anonymity, privacy and confidentiality of the participants that may be necessary when data is published or shared.

7.4.7.2. The researcher should retain enough information to confirm who the participant is if necessary.

7.4.7.3. The use of unique participant numbers issued at recruitment will allow the researcher to re-identify the data if needed.

7.4.7.4. The researcher should retain a master-file of names and other identifiable data to be stored securely and separately from the study data in locked password-protected databases with passwords kept separately.

7.5. **Strategy for Backing-Up Data:**

7.5.1. The strategy for backing up data should be clearly stated such as once a week or daily.

7.5.2. The researcher should specify how frequently data will be backed up on the systems or manually.

7.5.3. The researcher should also ask how disaster recovery will be dealt with should anything go wrong.

7.6. **Data Analysis:**

7.6.1. The data cleaning decisions might influence the analysis and should be considered and checked.

7.6.2. Revision of missing values should be considered in numeric data to determine if a pattern of missing data can be identified.

7.6.3. Member checking should be considered in any qualitative study.

7.7. **Archiving and Destruction of Data:**

7.7.1. Data should be stored for 6 years as per the *SACAP Research Ethics Policy*.

7.7.2. Storage of this data should be done in such a way that it is easily retrieved.

7.7.3. Data should be kept de-identified and consent forms should also be kept.

7.7.4. Preferably the information should be boxed and labelled with the project title, data on which the data was transferred to storage, name of the researcher who was responsible for the data, date for destruction and number of boxes for this particular project.

7.7.5. When the data is destroyed it should be done in such a way that the information is completely destroyed.

7.7.6. Confidential data and records in paper format should be shredded.

7.7.7. Confidential data and records in electronic format should be destroyed by reformatting or overwriting. A delete instruction is not sufficient to ensure that all system software has been destroyed.

7.7.8. Audio-visual tapes should be degaussed through a magnetic field bulk eraser.

7.7.9. When destroying confidential data and records, the researcher should ensure that the most effective method is applied.

7.7.10. Data that may be permanently kept includes but not limited to:

7.7.10.1. Controversial or of high public interests;

7.7.10.2. Would be costly or impossible to reproduce;

7.7.10.3. Relates to the use of or supports the development of an innovative intervention;

7.7.10.4. Supports a patent application or other services;

7.7.10.5. Has long-term heritage, historical or cultural value; and/or

7.7.10.6. Is of significant value to other researchers.

8. REFERENCE DOCUMENTS

- Department of Health. (2015). *Ethics in health research: Principles, processes and structures* (2nd ed.)
- SACAP. (2021). *Research ethics policy.*
- The University of Sydney. (n.d.). *Study documentation and data management, psycho-oncology co-operative research group* (SOP_4.0.0-DM-121203).