SAMAREC DATA MANAGEMENT FOR CLINICAL TRIAL RESEARCH SOP

Version 1, Nov 2024

TABLE OF CONTENTS

Final Version	Reason for Amendment	Effective Date
1	Developed and published for implementation	November 2024

DEFINITIONS

Data – Refers to all information collected, recorded, or generated during a research study, including clinical, personal, and administrative data. This encompasses both raw and processed data from patient records, research observations, medical devices, and statistical analyses, in physical or digital formats."

Database – Refers to a structured system or repository used to store, organise, and manage data collected during a research study. It enables efficient retrieval, analysis, and sharing of data, often containing patient information, clinical trial results, or research records in both raw and processed formats.

Data Curation – the process of creating, organisation and integration and maintaining of data sets so that the value of the data is maintained over time, and the data remain available for reuse and preservation. It involves collecting, annotating, structuring, indexing, cataloguing, publication and presentation of the data for users.

Data Sharing – the process of sharing available data resources with multiple third parties, facilitating data access without compromising data integrity

Data Management Plan (DMP) – a roadmap for how data will be handled throughout the research process. It details how data will be collected, stored, accessed, processed and shared.

PURPOSE & SCOPE

The *Department of Health: SA Good Clinical Practice - Clinical Trial Guidelines (3rd edition, 2020)* states that the aim of data management is to translate the information collected from a participant into data that can be organised in a database efficiently and without errors such that statistical analysis of data can be performed. All steps involved in this process must be documented to allow for step-by-step retrospective assessment of data quality and study performance, i.e. an audit trail.

This Standard Operating Procedure (SOP) outlines the minimum requirements for a data management plan (DMP) that applies to all research studies submitted to SAMAREC, as set out by the *National Department of Health: SA Ethics in Health Research Guidelines (3rd Edition, 2024)* and the *Department of Health: SA Good Clinical Practice - Clinical Trial Guidelines (3rd edition, 2020)*. This SOP does not refer to the research study's Statistical Analysis Plan.

Establish guidelines for the proper management, storage, processing, and sharing of data within clinical trials to ensure compliance with South African Protection of Personal Information Act (POPIA). Compliance with POPIA mandates that personal information be processed lawfully, with appropriate consent, and measures are in place to protect data from unauthorized access or breaches. This SOP also covers data retention and destruction practices in line with regulatory requirements to safeguard trial participant information and maintain the ethical standards of data management.

DATA COLLECTION

There are several types of data in clinical research which can include demographic data, clinical measurements, laboratory results, imaging data and patient-reported outcomes. The DMP should clearly detail which types of data will be collected.

The Investigator is responsible for collection, quality, recording, maintenance and retrieval of source data arising from the clinical study.

- The Investigator should maintain adequate and accurate source documents and trial records that include all pertinent observations on each of the site's trial participants.
- Source documents should adhere to the ALCOA-CCEA standard (attributable, legible, contemporaneous, original, accurate, complete, consistent, enduring and available).
- The source document must be signed and dated by the clinician identified in the protocol, or designated person, on a visit-by-visit basis and stored securely.
- Changes to source data must be traceable, must not obscure or delete the original entry, and must be explained if necessary (e.g. via an audit trail).

Data collected by direct entry on a computer should be entered only by the Investigator and/or an appropriate designee.

- The computer system must be virus-protected, access-restricted and must record a data trail of all changes made to the **Case Report Forms (CRFs)**.
- The data collection system should be designed so that data changes are documented automatically, and so that no deletion of entered data can occur in order to maintain, audit and edit the data trail.
- Once a 'hard copy' of the computer stored data has been made, procedures for editing are as for paper CRF.

DATA VERIFICATION

The Investigator should ensure the accuracy, completeness, legibility and timeliness of the data reported to the Sponsor by means of CRFs. The design of the CRF should facilitate observation of the participant and be consistent with the study protocol.

- The protocol should specify which data will be entered directly into the CRF and will not be supported by other source data.
- Corrections to CRFs may be made only by the Investigator or an appropriate designee.
- Changes to CRFs must be traceable, must not obscure or delete the original entry, and must be explained if necessary (e.g. via an audit trail).

Quality control should be applied to each stage of data handling to ensure that all data are reliable and have been processed correctly. SOPs outlining the data collection process may be developed to ensure consistent and accurate data is collected.

The DMP should detail the frequency and degree of data verification checks.

DATA HANDLING

The Sponsor must use appropriately qualified individuals to handle data.

When using electronic trial data handling and/or remote electronic trial data systems, the Sponsor must:

- Ensure and document that the systems conform to established requirements for completeness, accuracy, reliability and consistent intended performance (i.e. validation).
- The Sponsor must base its approach to validation of systems on a risk assessment that considers the intended use of the system and its potential to affect human participant protection and the reliability of trial results.
- Maintain SOPs for using these systems. Such SOPs should:
    - Cover system setup, installation and use;
    - Describe system validation and functionality testing, data collection and handling, system maintenance, system security measures, change control, data backup, recovery, contingency planning and re-commissioning;

    - Describe the responsibilities of the Sponsor, Investigator and other parties with respect to the use of these computerised systems; and
    - Make provision for training by all users.
- Ensure that the systems are designed to document data changes without deleting previously entered data (i.e. maintain an audit trail).
- Maintain a security system that prevents unauthorised access to the data.
- Maintain a register of persons authorised to make data changes.
- Maintain adequate data backup.
- Ensure that blinding, if any, is maintained during data entry and processing.
- Ensure the integrity and confidentiality of data, including any that describe the context, content and structure of the data – especially when making changes to computerised systems.

If data are transformed during processing, it must be possible to compare the original data with the processed data.

The Sponsor must use an unambiguous participant identification code that allows identification of all data reported for each participant.

The DMP should address who will handle the data as well as the processes involved in data handling.

DATA STORAGE, SHARING AND ACCESS

Data should be stored using systems that are reliable administratively and technically, and secure from unwarranted exploitation or cyber-attacks. Regular data backup procedures should be implemented. The DMP should also detail how long the data will be stored and describe how data backups will be stored and accessed.

The Sponsor is responsible for securing agreement from all involved parties to ensure direct access to all trial-related sites, source data/documents, and reports for the purpose of monitoring and auditing by the Sponsor, and inspection by domestic and foreign regulatory authorities. Any transfer of ownership of the data must be reported to SAHPRA.

The Sponsor must ensure that the protocol or other written agreement specifies that the Investigator must provide direct access to source data/documents for trial-related monitoring, audits, REC review and regulatory inspection.

Data curation is essential for the long-term preservation and accessibility of data. Researchers must implement proper curation practices, including cataloguing, annotating, and structuring data, to facilitate future reuse and maintain the integrity of the dataset.

DATA MONITORING

Ongoing monitoring processes to identify and rectify data issues promptly should be established. Centralised monitoring processes provide additional monitoring capabilities that can complement and reduce the extent and/or frequency of on-site monitoring by such methods as:

- Routine review of submitted data;
- Identification of missing data, inconsistent data, data outliers or unexpected lack of variability and protocol deviations that may be indicative of systematic or significant errors in data collection and reporting at a site or across sites, or may be indicative of potential data manipulation or data integrity problems;
- Using statistical analyses to identify data trends such as the range and consistency of data within and across sites.

ROLES AND RESPONSIBILITIES

Researchers must be suitably qualified and technically competent (suitably trained and supervised, in the case of student researchers) to carry out the proposed research. The principal investigator (PI) or research leader has primary responsibility to ensure the safety and wellbeing of participants, the scientific integrity of the protocol, research data management, and responsible implementation of that protocol. For international multi-site research, at least one (co-PI must be physically present in South Africa.

Ideally, before submission to SAMAREC, persons with discipline specific expertise and experience should assess whether:

- the proposed methodology and study design of the protocol are sound and align with the relevant disciplinary scholarly standards
- the study is feasible for the circumstances, considering the available resources
- the importance and novelty of the scientific question are appropriate
- the stated aims and objectives are achievable and will likely produce valid outcomes
- the evaluation of relevant literature and previous studies are thorough and appropriate
- the researchers are suitably qualified
- the suggested research DMP appears appropriate. (All members of the research team who are involved in the data management process should be trained on the DMP).
- there is a results dissemination plan
- potential or existing conflicts of interest are addressed.

SAMAREC REVIEW

An agreement should exist between the host research institution and the collaborating institution(s) regarding all aspects of the research, including (a) management of the research itself, (b) research data management that includes the fate of the data and samples after completion of the study, (c) financial arrangements, (d) approach to research output publications, (e) infrastructure development, (f) allocation of intellectual property rights, and (g) dispute resolution mechanisms.

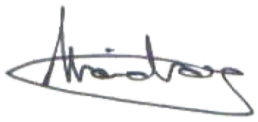
Therefore, SAMAREC is required to consider the following when reviewing all clinical trials

a) What are the data's availability and quality, and are they applicable to the local context?
b) How were the data sourced? Are the data in aggregate format, how are the rights of affected populations protected?
c) How has the proposed selection of data sets been identified? Has potential selection bias been addressed? Has the limitation of the data been demonstrated?
d) Informed consent: how is this facilitated, implemented and documented for project-specific studies?
e) How are the participants' right to withdraw their stored data facilitated and protected?
f) Where secondary data is being used, how were the provisions for informed consent for the original data collection facilitated?
g) Where big data is used and consent for the use of the stored data is not available, how does the researcher balance issues related to participant privacy while acting in the public interest?
h) What measures are in place to uphold individual privacy and confidentiality? What measures are put in place to mitigate possible risks to reidentification of study participants? Is there potential for data to be re-linked to groups of individuals?
i) Are there processes in place for regular evaluation of biases especially in big data research?
j) Has the researcher considered the entire data cycle, from the source to where the data go, where and how it is stored, who has access, with whom and how data will be shared, how the data will eventually be disposed of? Was the data processing lawful, fair and transparent? For this purpose, researchers should submit a DMP,

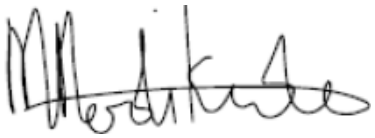
which can include how data will be collected, stored, accessed, shared and disposed of or retained. The DMP should indicate how data security will be maintained and the processes for dealing with possible data breaches.

k) Has maintenance of confidentiality of and accountability for the data been discussed?
l) Does the DMP indicate how data security will be maintained as well as the processes for dealing with possible data breaches?
m) Where stored data are re-purposed, how is unauthorised sharing of data with third parties avoided? Similarly, would unauthorised sharing of data inadvertently stigmatise affected individuals?
n) Does transfer and processing of data meet POPIA requirements?
o) If data sharing options include the use of open-access databases, do the selected databases meet the minimum legal, ethical, and security requirements?
p) Will the data be used for commercial purposes?

Approved by:



Dr N Naidoo
SAMAREC Chairperson
Signed: dd mmmm yyyy



Dr M Nodikida
SAMA CEO
Signed: 21 November 2024