Data Management in Clinical Trials

This SOP has been adapted from RIES/F-SOP/007/01 and reproduced with kind permission from Cardiff University

POLICY DETAILS	
Document Type	Standard Operating Procedure
Document name	KHP-CTO/CT/SOP18.0 Data Management in Clinical Trials
Version	Final v2.1 - 27/Mar/2023
Effective from	28/Mar2023
Review date	10th January 2025
Owner	King's Health Partners Clinical Trials Office
Original Prepared by	Helen Critchley, Quality Manager
Reviewed by	Amy Holton, KHP-CTO Quality Manager
Approved by	Jackie Pullen, Director KHP-CTO
Superseded documents	Final v2.0 - 05/Jan/2021
Relevant regulations/legislation/guidelines	Statutory Instrument 2004 no 1031 Statutory Instrument 2006 no 1928 *(as amended from time to time)* Data Protection Act 2018 General Data Protection Regulation NHS Data Security and Protection Toolkit Federal Information Processing Standard Publication 140-2

CHANGE HISTORY			
Date	Version Number	Change details	Approved by
05/Jan/2021	2.0	Scheduled review, Update to section 4.1.1, “Subject” wording replaced with “participant” in section 3.	Jackie Pullen
27/Mar/2023	2.1	Scheduled review, minor amendments to section 4.6	Jackie Pullen

Table of Contents

1.0 GLOSSARY

Case Record Form (CRF) - a printed, optical, or electronic document designed to record all of the protocol required information to be reported to the sponsor on each trial subject.

Chief Investigator (CI) - A Registered Physician, Dentist, Pharmacist or Registered Nurse/Midwife who has overall responsibility for the conduct of the trial.

Clinical Research Associates (CRAs) – Part of the KHP-CTO Quality Team. Ensure compliance with the Regulations, GCP and SOPs, by monitoring clinical trials.

Clinical Trial of an Investigational Medicinal Product - Any investigation in human subjects, other than a non-interventional trial intended to discover or verify the clinical, pharmacological or other pharmacodynamic effects of one or more medicinal product or to identify any adverse reactions to one or more such products and to trial absorption, distribution metabolism and excretion in one of more such products with the object of ascertaining the safety or efficacy of those products.

Data - Facts, figures and statistics collected together for reference or analysis.

Data Management Plan (DMP) – A document, detailing how the data management activities for the trial will be carried out.

Good Clinical Practice (GCP) - as defined in the Regulations.

Investigational Medicinal Products (IMP) - means a pharmaceutical form of an active substance or placebo being tested, or used as a reference in a clinical trial. This includes a medicinal product which has a marketing authorisation but is, for the purposes of the trial -

(a) used or assembled (formulated or packaged) in a way different from the form of the product authorised under the authorisation,

(b) used for an indication not included in the summary of product characteristics (or equivalent document) under the authorisation for that product, or

(c) used to gain further information about the form of that product as authorised under the authorisation

KHP-CTO Standard Operating Procedures (SOPs) - "detailed, written instructions to achieve uniformity of the performance of a specific function," SOPs are the base on which Quality Systems and Processes are conducted and monitored against.

King's Health Partners (KHP) - King's Health Partners Academic Health Science Centre is a pioneering collaboration between one of the King's College London (University) and three of London's most successful NHS Foundation Trusts – Guy's & St Thomas', King's College Hospital and the South London & Maudsley.

King's Health Partners Clinical Trials Office (KHP-CTO) - Established in 2006 by King's College London, Guy's & St Thomas' NHS Foundation Trust, South London and Maudsley NHS Foundation Trust and King's College Hospital Foundation Trust to provide a streamlined approach for all aspects of trial administration.

Personal Data and Processing Personal Data - The term 'personal data' refers to "[A]ny recorded information about a living individual who can be identified from that data or from that data and other available data.

Principal Investigator (PI) - A Registered Physician, Dentist, Pharmacist or Registered Nurse/Midwife who has responsibility for the conduct of the trial at a host site.

Protected Personal Information - Protected personal information is:

a) any information that links an identifiable individual with information that, if released, would put them at significant risk of harm or distress;
or

b) any source of information relating to 1,000 or more individuals not in the public domain, even if the information is not considered likely to cause harm or distress.

Research & Development Dept. (R&D) – NHS department responsible for confirmation of capacity and capability for all clinical research.

Source data - *"*all information in original records and certified copies of original records of clinical findings, observations, or other activities in a Clinical Trial necessary for the reconstruction and evaluation of the Clinical Trial."

Electronic Source (eSource) Data - Source data captured initially into a permanent electronic record.

Source documents - Source documents are original documents, data and records.

eSource documents - The electronic record used to aggregate a particular instance of eSource data items for capture, transmission, storage, and/or display, and serving as a source document for a CTIMP.

Source Data Verification (SDV) - Source Data Verification (SDV) refers to the process of checking the reliability of the data recorded in the data collection tool against the source documents.

The Regulations - Medicines for Human Use (Clinical Trial) Regulations 2004 transposed the EU Clinical Trials Directive into UK legislation, as Statutory Instrument 2004 no 1031. This became effective on the 1st May 2004**.** An amendment to implement Directive 2005/28/EC was made to the Regulations as Statutory Instrument 2006 no 1928. As amended from time to time.

Trial Master File (TMF) - a standard filing system which allows the effective storage and location of essential documents, that is the large volume of regulatory documents and approvals needed for clinical research. The filing system can be in the form of a single project file or a number of files/filing cabinets, depending on what is deemed most appropriate for a particular clinical trial given its size and complexity. The regulatory documents and approvals within the TMF will be maintained alongside case report forms and source documentation.

2.0 BACKGROUND AND PURPOSE

The purpose of this SOP is to provide the minimum standards required to ensure all Clinical Trial data, from the point of collection from source documents up to the point of archiving, excluding the requirements for statistical analysis, are managed, collected and verified in the appropriate manner. This SOP is to ensure the data are recorded correctly in order that Clinical Trials conducted within the partner institutions comply with UK and European Law. These laws comprise; Statutory Instrument 2004/1031 – the Medicines for Human Use (Clinical Trials) Regulations 2004 which transposed the European Union Directive 2001/20/EC for Clinical Trials into UK law effective from the 1st May 2004. The original UK regulations were amended in August

2006 to incorporate the EU Good Clinical Practice Directive (2005/28/EC) as Statutory Instrument 2006/1928 and as amended at any time.

Efficient data collection and management is an essential component of a Clinical Trial. Only data that is relevant for the purpose of the Clinical Trial should be recorded.

3.0 SCOPE

All clinical trials sponsored by one or more of King's Health Partner Organisations or clinical trials where the sponsor responsibilities are managed by the KHP-CTO will conduct data management as described in this SOP. The standards set out in this SOP are considered as the minimum required for a Clinical Trial.

The data management process typically encompasses: the design and production of the data capture tool for the collection of participant data from an investigator site(s); the design and construction of databases; the processing of the data, database lock and the production of the final data set(s) for analysis. This SOP does not apply to the statistical analysis of Clinical Trial data.

Where Data Management is outsourced to an external organisation this SOP will not apply. However, vendor oversight will be maintained by the KHP-CTO on behalf of the Sponsor. This will include confirmation that the Vendor has, and adheres to, organisational SOPs that detail the principles described within this SOP to ensure compliance with GCP and UK clinical trial legislation.

4.0 PROCEDURE

Appropriately trained and delegated individuals are responsible for handling the data, verifying the data, conducting the statistical analyses, and preparing the Trial reports. The Sponsor may delegate management of the trial data to the Chief Investigator (CI) or a specialist function group such as a Clinical Trials Unit. Any delegation of data management will be clearly documented.

4.1 Data Management Process

- Data management is the collection, storage, preparation for evaluation, extraction and archiving of data collected according to a clinical trial protocol, during the conduct of a clinical trial.

- A Case Report Form is a printed, optical, or electronic document (eCRF) designed to capture or record all of the protocol required information

4.1.1 Data Management Plans

How the data is to be managed both during and after the conduct of a clinical trial must be documented, this may be detailed within the trial protocol or as a separate document – a Data Management Plan (DMP) (see section 5.1). It is recommended that for large multi-centre studies a separate DMP is in place.

The DMP will be finalised and authorised by the appropriate member of staff prior to the start of trial recruitment.

A DMP (or the data management section of a protocol) should cover an overview of the data management process including: data flow, database specifics and systems used, validation and SDV processes, query processes, QC data checks, protocol non-compliance definitions and handling, PV reconciliation, training of personnel, database lock process, data extraction procedures, data location and archiving, process for return of the trial data and documentation to the sponsor.

4.2 Data Management databases and software

4.2.1 Databases

- The CRF(s) will be designed in accordance with the protocol. Once approved, the KHP-CTO on behalf of the Sponsor will ensure that a regulatory compliant database is available to store the relevant Clinical Trial data collected. The database specifics will be detailed within the trial DMP.
- A fully validated eCRF and/or database is required for all trials within the scope of this SOP.
- After full risk assessment conducted by the Quality Manager or delegate, on rare occasions, the use of MS Excel for data storage may be permitted for trials with minimal data collection on less than 10 participants. If it is decided that MS Excel is appropriate, each spreadsheet must be reviewed against source data printed and the sheet signed and dated by the KHP-CRA or delegate that has conducted the source data verification. The CI or delegate conducting the data extraction (for analysis) will be required to provide documentation to demonstrate validation of the data extracted against the printed signed spreadsheets.
- Where data management is contracted out to a third party an appropriate agreement must be put in place and vendor oversight maintained.
- All computer systems (hardware and software) being utilised for the collection and analysis of clinical trial data, must undergo full user acceptance testing, validation and sign off by the KHP-CTO Clinical Trials Systems Executive or delegate, prior to use.
- It is expected that all databases to be used in a Clinical Trial will have undergone final validation, testing and sign-off prior to the trial opening to recruitment. Where extenuating circumstances prevent the database(s) (or parts thereof) from being finalised before

opening to recruitment, appropriate alternative manual systems must be implemented and the database not released for use until full validation has been obtained

- If any modifications are made to the database during its lifecycle these must be tested and re-validated prior to implementation.

4.2.2 Software

- All software will be subjected to the appropriate validation and testing processes. This process will usually require formal statistical input.
- The results of all testing and validation will be filed in the Trial Master File (TMF).
- The Quality Manager or delegate must consider the longevity of the file formats which the selected software uses, to ensure that the data remains accessible for as long as required.
- Use of well documented file formats which have undergone thorough testing, are non-proprietary and usable on different hardware and software platforms will reduce the risk of inaccessibility. If non-proprietary file formats are not an option, the use of widely adopted file formats (for which manufacturers are more likely to provide some backward compatibility) or the use of software which can export data into non-proprietary open file formats should be considered.
- It should be noted, that computerised Laboratory Information Management Systems which capture analytical results of tests conducted during a clinical trial, are considered as part of the data management. The Quality Manager or delegate will seek assurance at the trial planning stage that the accreditation status of the computerised system in the chosen laboratory is suitable.
- Where local arrangements have been made, consideration will need to be given to the expected lifespan of the storage media used as media degradation and technological advances may render the records inaccessible in the future.

4.2.3 Freeware and open source

Freeware and open source software use remains subject to the normal vigorous validation and testing procedures.

4.3 Data entry, data processing and data validation

- In accordance with GCP, quality control (QC) checks must be implemented for each stage of the data handling process, to ensure that all data are reliable and have been processed correctly.

- The frequency, level and point at which a QC check is conducted will vary according to the complexity and risk level of the trial and should be determined in accordance with the trial risk assessment.
- The frequency and level of data QC checks are to be included in the trial DMP.
- For trials not using an eCRF, the CRA will conduct a review of the database before the end of the trial at time points specified in the trial specific monitoring plan.

For electronic Clinical Trial data, the following should be maintained:

- A security system that prevents unauthorised access to the data;
- A list of individuals who are authorised to make changes;
- Adequate back up of the data;
- A fully auditable trail of data queries and corrections.

If data are transformed during processing, it should always be possible to compare the original data and observations with the processed data.

Chief Investigators should ensure that the data management team are fully aware of any amendments to a trial which may affect the data collection/management process.

4.3.1 Data entry (for paper-based data collection)

- The Principal Investigator (PI) at the site is responsible for ensuring the accuracy, completeness, legibility, and timeliness of the data reported to the CI/ Data Manager in the CRFs and in all required reports.
- Data reported on the CRF that are derived from source documents should be consistent with the source documents or the discrepancies should be explained.
- Data must be entered into the CRFs and database in a timely manner.

4.3.2 Data processing

The following procedures should be followed when processing Clinical Trial data:

- All transactions to the database (insert, update, delete) must have a clear and complete audit trial. For some software (e.g. Excel) this may necessitate the printing of data and the certification and dating of the data as an accurate record of the previous and current versions of the database;
- Data should only be accessible to authorised personnel;
- The Data Handler must comply with GCP and is responsible for keeping data secure and confidential at all times;

- Coding should be performed using appropriate dictionaries;
- Where autocoding is not possible, manual coding may be performed.

4.3.3 Data querying

A process for raising and resolving data queries throughout the Clinical Trial will be included in the DMP.

4.3.4 Data validation/ Source Data Verification (SDV)

- An essential aspect of data management is the process of data verification/validation. This process must be designed to ensure that the most accurate validated set of data is provided for statistical analysis.
- The type and level of data validation will vary according to the trial risk assessment.
- The data validation and SDV processes should be decided upon prior to the trial opening and should be documented in the DMP (or protocol) and trial specific monitoring plan.
- Data validation can take place at certain points during the trial but must be completed before the data is released for statistical analysis.

Data validation should continue until all missing values and inconsistencies are corrected or clarified.

4.4 Data Back-up

- Robust back-up systems are required to guard against loss of data due to software or environmental disasters.

4.5 Data Protection

- Throughout the data management process, it is vital that all trial data are kept in a secure location and in accordance with the terms of the General Data Protection Regulation and the Patient Information Leaflet and Consent Form.
- Participant confidentiality must be maintained, and data stored in (e)CRFs be pseudonymised at all times.
- Protected Personal Information must not be accessed via remote access (e.g. personal laptops, home PCs or PCs in public locations such as libraries) or synchronisation facilities (i.e. facilities which copy the data and store them locally on the machine/device from which it is accessed), unless the machine/device used to access the data is fully encrypted.

- If electronic CRFs, or any other data defined as Protected Personal Information, are transferred or stored off site using removable media (including laptops, memory sticks, smart phones, CDs, and portable hard drives) they must be encrypted using software which meets a current industry standard such as FIPS 140-2. The transferred material on the removable media will be deleted as soon as transfer is successfully affected.
- Paper CRFs will be kept in locked filing cabinets in rooms which should be locked and made accessible to authorised personnel only. Where the filing cabinets are stored in open-plan offices, appropriate levels of security must be adopted (e.g. restricted access).
- Paper CRFs must be transferred to a coordinating centre for data entry, they must be sent by contracted secure courier service or fax to minimise loss of data. A log of documents sent and received at each centre involved must always be maintained by the Trial Manager or delegate.
- If electronic data transfer is used, this will be via a secure system, password protected and encrypted using software which meets a current industry standard, such as FIPS 140-2. If transferred via email, the encryption password must never be transferred in the accompanying email.
- The database management system should also be password protected, with each member of the research team responsible for data entry/management having their own password.
- Where a suspected or actual data security breach occurs, the Sponsor and clinical site institution must be notified immediately.

4.6 Database Lock

- The final database will be “locked” to ensure access to the final dataset is permanently restricted after extraction or final analysis.
- The process used to lock the database will be described in the DMP.
- Depending on the trial, a staged locking process may be appropriate. For example, a two-stage locking process may be employed, where access is initially restricted to a (pre-defined) limited group, before the final lock is performed. If staged locking is utilised the procedure for staged locking processes will be described in the DMP.
- Evidence of when and how the lock(s) was/were performed will be documented. The same standards of security will be upheld for any final dataset files.
- Confirmation that database lock has occurred must be confirmed in writing, email is acceptable, to CI or delegate prior to the release of treatment blinding codes.
- The final dataset and analysis codes will be archived as detailed in KHP-CTO Archiving SOP 4.0.

4.7 Release and Access to the Database/Dataset

- It will be documented in the DMP or protocol how the final data are to be provided/extracted for statistical analysis after database lock has occurred.
- Extractions from the database for final analysis will only be undertaken after the database is locked.
- Occasionally it may be necessary to correct previously missed data errors or inconsistencies after the data has been released for analysis. This is limited to important corrections (that will have a significant impact on the reliability of the results) and will only occur in extraordinary conditions. Unlocking will only be undertaken in consultation with the statistician, and written approval with justification will be documented and filed in the TMF prior to unlocking. The approval document will contain details of the authorised and delegated individuals who can make the required changes and the agreed process to restrict access to unlock and re-lock the database after all the intended changes have been made. A documented review of the audit trail will be conducted at re-locking. The new final dataset must not overwrite any analysis sets that were created at the original lock. Both data sets will be stored and archived as detailed in this SOP and KHP-CTO SOP 4.0.
- All details of any changes made to the locked database will be documented in the Clinical Study Report.

5.0 RELATED TEMPLATES

5.1 Data Management Plan template

6.0 APPROVAL AND SIGNATURE

27 March 2023

Jackie Pullen
Director,
King's Health Partners Clinical Trials Office

Date