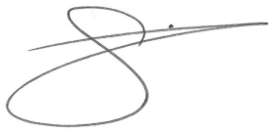
Standard Operating Procedure (SOP)

Data Management

SOP ID Number:	JRESGOVSOP0038	**Effective Date:**	04/11/2021
Version Number and Date:	Version 3.0 23/04/2021	**Review Date:**	04/11/2023
Author:	Georgia Bullock	**Title:**	Research Development and Governance Manager
Approved by:	Subhir Bedi	**Date:**	19/07/2021
Signature of Authoriser			

This is a controlled document.

The master document is posted on the JRES website and any print-off will be classed as uncontrolled.

Researchers and their teams are responsible for checking the JRES website for the most recent version. They may print off this document for training and reference purposes.

SOP Chronology		
SOP Version Number:	**Reason for Change:**	**Author:**
v1.0	Original Version	Lucy H H Parker
V2.0	Updated trust logo and reference added to Data Management resource on University website	Debs Rolfe
V3.0	Update to reflect the change in job titles and insertion of associated JRES documents and major re-write of SOP	Ali Alshukry / Georgia Bullock

Associated JRES documents

SOPs	WPDs	Docs	LOGs
JRESGOVSOP0024 Computerised Systems JRESGOVSOP0051 Trial Oversight Committees	JRESWPD0023 General Research Definitions JRESWPD0029 CRF Design	JRESDOC0126 CRF Template	

Contents

1. Background 4

2. Joint Research and Enterprise Services (JRES) Policy 4

3. Scope 4

4. Definitions 4

5. Responsibilities 5

6. Procedure 5

6.1 Data Management Process 5

6.2 Data Management System 5

6.3 Data Capture and Entry 6

6.3.1 Electronic Data Capture 7

6.3.2 Paper-based Data Capture 7

6.4 Database Lock 8

6.5 Data Analysis 8

6.6 Data Protection 9

6.7 Data Security 9

7. References 10

8. Appendices 10

1. Background

An essential element of conducting a clinical trial is efficient data collection and management. Only data that is essential for the purposes of the study should be collected. It is advisable for investigators to seek advice from a trial statistician as early as possible in the trial design process to facilitate this.

ICH GCP guidelines specify that appropriately qualified individuals should supervise the trial data handling, verify the data, and conduct the statistical analyses.

2. Joint Research and Enterprise Services (JRES) Policy

All JRES SOPs will be produced and approved in accordance with the JRES SOP on SOPs and must be used in conjunction with local NHS Trust and University policies and procedures.

The JRES acts as the representative of both St George's University of London (SGUL) and St George's University Hospitals NHS Foundation Trust (SGHFT). St George's will be the official name used on all SOPs to represent either institution acting as Sponsor.

3. Scope

This SOP describes the process for data management for St George's sponsored clinical studies. Specifically, it describes the processes involved in collecting, validating and analysing the data.

This SOP describes the full data management process, including data entry, data cleaning and resolving data queries.

This SOP must be used in conjunction with any relevant SGHFT and SGUL policies and procedures.

4. Definitions

For general research-related acronyms used in this SOP, refer to General Research Definitions Working Practice Document (JRESWPD0023).

5. Responsibilities

This SOP should be read and understood by the Chief Investigator and all those in the research team that will be responsible for data management, in order to understand their responsibilities in relation to the management of data during the lifecycle of the study.

6. Procedure

6.1 Data Management Process

The process of data management involves converting the data collected using data collection tools, most commonly Case Report Forms (CRFs), into electronic data that can then be statistically analysed.

This SOP needs to be followed according to the size and complexity of the specific trial being conducted, as smaller trials may not require some of the processes described.

6.2 Data Management System

Once the CRF has been designed in accordance with the protocol, the database to store the information collected should be designed. If an electronic data capture (EDC) system is used, it is important to ensure that the database has been validated and has been deemed fit for purpose and that this has been documented. This is essential for EDC systems used for CTIMPs.

Depending on the size and type of study, this database could be a standard spreadsheet or a more technical Data Management System. When developing a database, the following points should be considered:

- Ease of setting up and maintaining data entry screens.
- The ability for more than one user to use the system at the same time.
- The ability to store and retrieve all data required for the study efficiently.
- A full audit trail can be maintained.

A useful resource that SGUL investigators should access together with this SOP is http://library.sgul.ac.uk/researchers/research-data-management

The Chief Investigator (CI) should decide whether to have the database developed in-house or whether to purchase a system. If a decision is made to purchase a system, the CI must ensure the system is fit for purpose.

REDCap is one EDC provided by SGUL for SGUL researchers. However, unless validated prior to use, REDCap cannot be used for CTIMPs which are sponsored by St George's. The JRES governance team can provide advice on the use of REDCap and on alternative databases for CTIMPs. There are costs associated with the set-up and management of REDCap and for other databases. There is further information on REDCap and other data capture systems here: SGUL Guidance on Databases for Research Projects V1.0_06.05.2021

Under ICH GCP there should be a specific process for managing the study database in place. The database should:

- Allow changes to be made to the data in a documented manner.
- Not delete the original data entry to ensure an audit trail for the data is maintained.
- Be secure, with appropriate password-protected access to prevent unauthorised access to the data.
- Have a list identifying those individuals permitted to make changes to the data.
- Have adequate backup for the data.
- Allow for blinding to be maintained if this is involved in the study.

For CTIMPs, there must be an automatic electronically-generated audit trail, rather than a handwritten log of database modifications.

6.3 Data Capture and Entry

The trial team must be trained in the completion of the CRF (paper or electronic) prior to recruitment of the first participant and their role must be captured on the Delegation of Duties Log. Only suitably trained personnel can undertake data entry/capture.

The design of the CRF is the responsibility of the CI or their designated person. For guidance on the design and content of a CRF, see JRESWPD0029 CRF Design and JRESDOC0126 CRF Template (this can be modified for each individual study). The CRF must be consistent with the protocol and the design dependent on the data to be collected. No information will be collected that cannot be justified by the protocol or standard safety procedures; this is to be confirmed by the CI and the delegated clinical trial managing organisation/facility and signed off by the Sponsor's Clinical Research Associate (CRA).

Any missing data must have an explanation in the source documents. Values outside of normal or expected reference ranges will have a comment (e.g. the significance) noted in the source data. All required fields must be completed. If a procedure is not carried out, then this must be indicated.

Laboratory numbers must be entered without conversion from printed reports. If conversions are necessary, eg: multi-centre sites where units of measurement may differ, space must be made available on the CRF for the original figure, the conversion factor, and the converted result. CRFs will indicate which unit of measurement is to be used for any measurements captured.

6.3.1 Electronic Data Capture

The following requirements must be met by the CI and research team when using electronic data capture:

- All electronic systems used must include data validation, range checks and consistency checks, where agreed with the trial team, to ensure good data quality.
- Actual dates/times of events or data collection points must be recorded.
- Data queries or data anomalies will be handled following trial-specific guidelines.

6.3.2 Paper-based Data Capture

The following requirements must be met by the CI and research team when using paper-based data capture:

- Paper CRFs must be version controlled and paginated.
- Paper CRFs must be completed in permanent ink.
- Any error or alteration will be crossed through with a single line so that the incorrect entry is still legible. **Correction fluid or eradication of entries made on the CRF is not permitted**. The correct data must then be added, and the change initialled and dated. Where appropriate, an explanation of the correction must accompany the change.
- The CRF must be signed, where indicated, by the CI or delegate to confirm that the CRF entries are complete and accurate.
- Data queries must be addressed quickly and an audit trail must exist.

6.4 Database Lock

The CI will ensure that all data queries are raised as soon after data collection as is practicable. Data queries may include, but are not limited to, missing, inconsistent or implausible data. All data queries must be resolved prior to data lock.

At the end of a trial, data must be locked to prevent further additions or changes to the data, using the data management system's file-locking facility if such a facility exists, otherwise a snapshot of the data must be taken for analysis and access to the data denied to anyone except the Statistician or Data Manager.

Occasionally there may be a need to unlock the data after it has been released for analysis to correct missed data errors or inconsistencies. Unlocking must only be undertaken in consultation with the Statistician and where the omitted/erroneous data would have a significant impact on the reliability of the study.

The CI must ensure that all details of the changes made to the database whilst unlocked are documented. The justification, the written approval and the effect of the unlocking on the statistical outcome must be documented in the TMF prior to unlocking. If possible a partial opening of one section of the locked data will be sought rather than a full opening. An audit of the changed data must be carried out by the CI or their designee at the point of re-locking. The CI or their designee must ensure the new version is not saved over the previously locked version and that both versions are made available for audit and inspection. The database lock must be documented.

6.5 Data Analysis

The CI is responsible for ensuring that a quality check of data is carried out prior to the interim and final analyses. Quality checks must include checking for outliers, missing data, date checks and any inconsistencies. A study data monitoring plan shall describe any data entry checking process to be used (e.g. single data entry and verification, double data entry), an acceptable data entry error limit and action to be taken if the error rate exceeds this limit.

When data is to be exported for statistical analysis, the dataset must be fully anonymised. Any files containing identifiable sensitive data must be encrypted and password-protected before transfer. Passwords shall be communicated to the recipient separately from the file containing the sensitive data.

6.6 Data Protection

All research study staff must ensure that they are familiar with, and adhere to, current Data Protection legislation and SGUL/SGHFT Information Governance policies. The CI will ensure that any individual working on the trial is handling personal data in an appropriate, confidential manner.

Paper or computer records will only be accessible to the minimum number of people needed to ensure the smooth delivery of the trial. Personal information must not be sent by email - encrypted files must be used unless sent via the secure 'NHS.net' email system.

Personal data must not be transmitted in a way that could cause loss of the data or allow interception by unauthorised parties.

6.7 Data Security

The CI is responsible for the data security for the trial, and must ensure the following requirements are met:

- Electronic data must only be stored on devices that that are backed up in a secure and timely manner.
- Data used for trial analysis shall be anonymised. Data used for trial management will use the minimum number of personal identifiers necessary for study conduct and to ensure patient safety.
- Personal data must not be stored or transmitted on removable media or laptops without encryption. Datasets shall be encrypted prior to transmission, or transfer (e.g. on CD), using an industry-standard encryption mechanism (e.g. AES-128), or transferred securely using the 'NHS.net' email system.
- Any machines used to enter or access trial data shall have appropriate security software installed.
- Access to the data will be limited to authorised personnel and each user of the system shall have an individual account.
- A record must be kept of authorised users and the access levels that apply to each user. The list of authorised users will be kept in the Trial Master File.
- The study database must have an audit trail for any changes made to electronic data after initial entry. This shall include the original value together with the date of the change and who made the change. If possible, a brief explanation of why the change was made must also be recorded.

7. References

ICH GCP.

During Your Project (sgul.ac.uk)

https://www.sgul.ac.uk/about/our-professional-services/information-services/information-governance/about-information-governance/what-is-information-governance

https://www.hra.nhs.uk/planning-and-improving-research/policies-standards-legislation/data-protection-and-information-governance/

https://mhrainspectorate.blog.gov.uk/2017/04/20/computer-system-validation-gcp/

SGUL Guidance on Databases for Research Projects V1.0_06.05.2021

8. Appendices

None associated with this SOP.