RD&I	NHS University Hospitals Plymouth NHS Trust **STANDARD OPERATING PROCEDURE**

DO NOT USE THIS SOP IN PRINTED FORM WITHOUT FIRST CHECKING IT IS THE LATEST VERSION

The definitive versions of all UHPNT RD&I Dept SOPs appear online, not in printed form, to ensure that up to date versions are used. If you are reading this in printed form check that the version number and date below is the most recent one as shown on the Trust's website: https://www.plymouthhospitals.nhs.uk/research-sops

Data Management & Independent Data Monitoring Committee

SOP No: P16

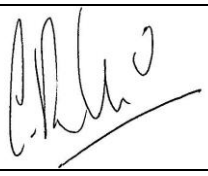
Version No: 2.1

Effective Date: Jan 2019

Supersedes: Version 2.0, Sep 2017

Page: **1** of **9**

Last Review Date: Jan 2019 Next review date: Sep 2022

	APPROVED BY
Name:	Chris Rollinson
Job Title:	Research Governance Manager
Signature:	
Date:	21st Jan 2019

1 Purpose and Scope

This Standard Operating procedure (SOP) describes the process for data management for University Hospitals Plymouth NHS Trust (UHPNT) sponsored clinical trials. Specifically the processes involved with collecting (data entry), validating (resolving data queries) and analysing such data. This SOP also describes the use of a Data Monitoring Committee (DMC) for assessing data during interim analyses, and how such a committee should operate.

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 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 (ICH 5.5).

In scope: research sponsored by UHPNT.

Definitions

PI	Principal Investigators
CI	Chief Investigator
CTIMP	Clinical Trial of an Investigational Medicinal Product
GCP	Good Clinical Practice
HCA	Health Care Assistants
TSC	Trial Steering Committee
MHRA	Medicines and Healthcare products Regulatory Agency
REC	Research Ethics Committee
RD&I	Research Development & Innovation
IMPs	Investigation Medicinal Products
ICH GCP	International Council on Harmonisation of Good Clinical Practice
CRFs	Case Report Forms
SOP	Standard Operating Procedure

WHO-ART	The World Health Organisation Adverse Reaction Terminology
MedDRA	Medical Dictionary for Regulation Activities
SDV	Source Data Verification
ECS	Edit Check Specification
UHPNT	University Hospitals Plymouth NHS Trust
ICT	Information and Communication Technology
IDMC / DMC	Independent Data Monitoring Committee / Data Monitoring Committee

2 Who should read this document?

All staff involved in setting up and conducting research e.g. Chief Investigators (CI), Principal Investigators (PI), Research Nurses & Midwives, Health Care Assistants (HCA), RD&I Managers and Clinical Trial Administrative staff.

The Chief Investigator is ultimately responsible for the study's data integrity.

3 Procedure to Follow

3.1 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. The processes described in this SOP may need to be adapted according to the size and complexity of the trial being conducted, as smaller trials may not require full use of the processes described.

3.1.1 Data Management Software

Once the CRF has been designed in accordance with the protocol; the database to store the information collected should be designed. Depending on the size and type of study this database could be a standard spreadsheet, or a more technical Data Management System may be required. When developing a database points to consider include:

- ease of setting up and maintaining data entry screens;
- the ability for more than one user to use the system at the same time; and
- the ability to store and retrieve all data required for the study efficiently.

Under ICH GCP there should be a specific SOP for managing the study database in place. The database should allow changes to be made to the data in a documented manner, and should not delete data entry to ensure an audit trail for the data is maintained (ICH GCP 5.5.3). The database should be secure, with appropriate

password-protected access to prevent unauthorised access to the data, with a list identifying those individuals permitted to make changes to the data. ICH GCP also requires that there is adequate backup for the data, and that if there is blinding involved in the study, that the data entry and processing systems allow this to be maintained (ICH GCP 5.5.3).

For clinical trials involving Investigation Medicinal Products (IMPs) there must be an automatic electronically generated audit trail, rather than a handwritten log of database modifications.

3.2 Coding CRF Responses

Before data entry to the database, the responses from the CRFs may need to be coded, using either a numerical or alphabetical code that can then be used for analysis. These codes should be decided before data entry begins e.g. codes 0 or 1 for Yes or No. Codes should also be in place for answers such as 'not known' or 'not applicable' e.g. 999 to show missing data. It is important to make sure that whatever value is chosen to represent missing data, that value would be unfeasible as an actual response.

Clinical data also needs to be coded for recording of all adverse events. The World Health Organisation Adverse Reaction Terminology (WHO-ART) and Medical Dictionary for Regulation Activities (MedDRA) both have a system of coding to assist with this categorised by System Organ Class. A code is assigned for each disease and adverse reaction. You can access WHO-ART and MedDRA through http://www.umc-products.com/graphics/3149.pdf and https://www.meddra.org/ respectively.

The coding can be done at various stages of the trial such as: during the initial data collection from the participant by the investigator or research nurse; after the data collection, but prior to entering the data on the database; or when the data are entered onto the database.

3.3 Data Entry

On initial receipt of a completed CRF (or CRF page), it should be date stamped and checked for initial missing or incomplete responses. If any inconsistencies are found these should be queried with the investigator and a record should be kept of all queries sent out. Instructions for sites to respond to data query responses should include **no use of Tippex**, **not to obscure the original data entry** and to **initial and date any amendments made**.

Once the paper CRFs are completed, the data must then be entered onto the database. Data entry should be done by trained data entry staff. For multicentre studies where the CRFs are being sent to a coordinating centre for data entry, a copy of the CRF should be retained by the Investigator, with the originals (usually 2 copies from No Carbon Required Paper CRFs) going to the coordinating centre. The coordinating centre must keep a log of all CRFs received, maintained by the Data Manager for the study.

All stored CRFs should be kept in a secure environment such as a locked filing cabinet in a locked room. Secure also means protection against environmental damage such as water, fire or rodent and insect damage.

3.3.1 Double and Single Data Entry with Control Checks

During data entry by trained staff, an average of 5% of errors is expected to occur. Two methods can be used to reduce the risk of errors: Double Data Entry or Single Data Entry with control checks.

i. Double Data Entry

This method involves two people entering the same CRF data onto the database independently of each other. Depending on the software used, the data may be entered twice onto the database on two separate files, which are then compared by the system for accuracy. If the two entries do not match this would be flagged up by the database. Alternatively when the second data entry person enters the data, if it differs from that entered by the first person, a message immediately appears on screen and the original data can be checked. This method depends on the availability of a technically capable database.

ii. Single Data Entry with Control Checks

This method may be more suitable for smaller single centre studies with less staff available for data entry and/or less sophisticated database software. Once the data has been entered, a visual check can be done between what is recorded on the paper CRF, and what was entered on screen.

3.4 Data Cleaning and Validation

An integral part of the data management process is validation; to ensure the most accurate 'clean' set of data is provided for the statistical analysis. Data validation can be carried out at three stages during the trial:

a. When CRFs are completed by the investigator

To improve accuracy at this stage all staff completing CRFs should be sufficiently trained in their completion. A CRF completion manual would assist with this. Validation should also be carried out as part of the on-going monitoring of the study; either by members of the research team or by independent monitors. Validation *via* monitoring is done through Source Data Verification (SDV). SDV involves checking the data entered into the CRFs against that in the original source records e.g. patient's hospital files for accuracy.

b. When data are entered in the database by data entry staff

During data entry the two methods for validation described above (3.3.1) can be used i.e. data entry checks or double data entry. Where data entry checks are used, if the study database has software enabled for automatic data entry checks, an Edit Check

Specification (ECS) document should be put together by the clinicians/statisticians/data staff involved with the study. The ECS should provide full details of the data entry checks that have been set up, and all checks should be tested before the trial begins.

Depending on the database software it is also advisable to set up warnings to alert data entry staff when values are entered outside of the expected range, or if the type of value entered is incorrect e.g. a numeric value entered rather than text. It is also useful to set up alerts for missing values where possible.

At this stage it is advisable to carry out systematic post-entry computer tests. Lists should then be created (either through automatic database software system, or manually) of the following data queries:

- All missing values will be listed
- All values outside of pre-defined range

Logical checks should also be performed to ensure consistent reporting between relevant fields and that there are no implausible difference between fields e.g. male and pregnant.

All checks should be defined before the study starts, and should be described in the Edit Check Specification document described previously. Data validation should continue until all missing values and inconsistencies are corrected or clarified.

6.5 Data Protection

During the entire data management and validation process it is essential that all study data are kept in a secure location and in accordance with the terms of the Data Protection Act 1998. Participant confidentiality must be maintained at all times and all study records should be kept in pseudonymised form identifying participants by their study code rather than name, or hospital number.

Any paper CRFs should be kept in locked filing cabinets in locked rooms only accessible by authorised personnel. The key to the participant code list should be kept separately to these documents, again in a locked, secure location. If paper CRFs must be transferred to a coordinating centre for data entry, they should be sent either by courier or registered post to minimise the risk of losing data. A log should always be maintained of documents sent and received at each centre involved. If electronic data transfer is used, this should be via a secure system, password protected and encrypted where possible.

The database itself should be password protected, with each data entry staff member having their own password. If data entry is performed at the investigator site it is essential that the investigator does not have access to the whole database, to protect against biases occurring due to investigators making decisions based on interim data.

Data that is stored on UHPNT networked computers should be stored when possible in an anonymised form with no identifiable information. For further guidance on data

protection please refer to the UHPNT Data Protection Policy or contact the Trusts Caldicott Guardian.

3.6 Data Processing SOP

Before the study starts, it is essential that a procedure for data processing; management; and validation is put together, and updated as necessary throughout the study. The procedure should contain information on the following:

i. Contact details for all study staff
ii. Details of the flow of data from the investigator site to archiving
iii. Procedures on how to complete the CRFs
iv. Monitoring plans e.g. frequency, how Source Data Verification will be done, expected ranges for data values
v. Data Entry
    a. How to use the data entry system
    b. Double or single entry
    c. Roles and responsibilities of study staff with regard to data
    d. Procedures in case of discrepancies
vi. Details of edit checks
vii. Description of Post Data Entry Validation System
    a. Who checks the consistency of the data?
    b. Who queries the Investigator?
    c. What is the format of the query form?
    d. How many days are allowed to answer a query?
    e. Who decides that a query is resolved?
viii. Data Protection procedures, including a back-up system

Although the above list is not exhaustive it provides a basis for the Data Management SOP that can be adapted and expanded as necessary.

3.7 Data Backup Systems

Whatever the format of the database software used to manage the study data, there should always be a back-up system in place to guard against loss of data due to software or environmental disasters. The UHPNT ICT service has a data backup service that provides a reliable means of protecting data held on the Trust servers. It is important to remember that ICT does not backup files on the local desktop of computers so data must be saved to an appropriate drive.

3.8 Independent Data Monitoring Committee (IDMC)

All Clinical Trials of an Investigational Medicinal Product (CTIMPs) and device studies which are blinded studies should have an IDMC, it is also recommended for large, complex trials that an Independent IDMC is set up to carry out reviews of trial data at staged intervals during the study. The role of the IDMC is review interim results and determine whether or not there are any safety issues or any reason why the study should not continue e.g. if interim results are showing strong evidence that the treatment/intervention is superior or inferior to the control.

The data reviewed by the monitoring committee should be as up to date as possible and should be validated up to the point of the interim analysis to ensure it experienced trial investigators, statisticians and clinicians; all of whom must be independent to the research team. The results should be reviewed at regular intervals as sufficient data accumulate. If there is a Trial Steering Committee (TSC) for the study, the IDMC would normally make their recommendations for action through them.

4 Document Ratification Process

The review period for this document is set as ***default of three*** years from the date it was last ratified, or earlier if developments within or external to the Trust indicate the need for a significant revision to the procedures described.

This document will be approved by the ***RD&I Manager or their Deputy***.

Non-significant amendments to this document may be made, under delegated authority from ***a Senior RD&I manager***, by the nominated author. These must be ratified by ***a Senior RD&I manager***.

Significant reviews and revisions to this document will include a consultation with ***appropriately knowledgeable staff***. For non-significant amendments, informal consultation will be restricted to ***staff*** who are directly affected by the proposed changes.

Dissemination and implementation

4.1. Dissemination of this SOP

4.1.1. New SOPs and new versions of existing SOPs: The Research Governance Manager will be responsible for ensuring authorised SOPs are uploaded on the RD&I intranet site. Internal Trust Staff are expected to use the RD&I intranet site to access latest versions of SOPs and to check the website regularly for updates.

Notice of new or amended procedural documents that have undergone a major amendment will be given *via* the following routes:

- Inclusion in the Trust weekly e-bulletin Vital Signs
- Direct email to Trust Researchers and or teams

4.2. Training in this SOP

4.2.1. All staff whose activities are subject to this SOP should ensure that they read and understand the content of the SOP.

5 Reference material

International Conference on Harmonisation (ICH) of Good Clinical Practice (and subsequent addendums).

Data Protection Act 2018.

6 Amendment History

Version Number:	2.1
Date Of Amendment:	Jan 2019
Details Of Amendment:	Updated Trust and Dept. name. Reduce signature requirement to single senior RD&I Manager. Updated references to the Data Protection Act 2018.

Version Number:	2.0
Date Of Amendment:	Aug 2017
Details Of Amendment:	Updated SOP template and number system. Reviewed and updated SOP.

Version Number:	1.1 (minor amendment)
Date of Amendment:	Mar 2012
Details of Amendment:	Cover page - Change of SOP location address.