Reference Number: UHB 449 **Version Number: 2**	**Date of Next Review:** *06/07/2025* **Previous Trust/LHB Reference Number:** *UHB039*

Data Management for Clinical Trials: Standard Operational Procedure

Introduction and Aim

This SOP complies with the UK Policy Framework for Health and Social Care Research 2017 and the requirements of the Medicines for Human Use (Clinical Trials) Regulations 2004 and International Conference for Harmonization – Good Clinical Practice (ICH-GCP). It aims to provide guidance on the 'minimum' standards for collecting, maintaining, verifying, correcting transferring and analysing data generated in trials Sponsored and Co-Sponsored by C&V UHB.

Within C&V UHB, all sponsored CTIMP trials require involvement of a clinical trials unit (CTU) and data management within the CTU should meet the requirements of the standards detailed herein.

In accordance with ICH GCP, the Sponsor should ensure appropriately qualified individuals are responsible for the overall conduct of the research study, handling the data, verifying the data, conducting the statistical analyses, and preparing the study reports. For C&V UHB sponsored trials, the Sponsor responsibility, in this regard, is delegated to the Chief Investigator (CI).

The standards in this SOP may also be considered 'best practice' for data management in C&V UHB hosted commercial and non-commercial trials.

Objectives

- To provide research staff with the standards for collecting, maintaining, verifying, correcting transferring and analysing data generated in trials sponsored and co-sponsored by C&V UHB. Thereby, ensuring that such documentation is complete, legible and easily accessible for monitoring, audit or inspection.
- To ensure that data is robust and accurately reflects the effects of investigational medicinal products on human subjects and allows the retrospective reconstruction of the trial if required.
- To fulfil the Sponsor's compliance with UK Policy Framework for Health and Social Care 2017; the Medicines for Human Use (Clinical Trials) Regulations 2004 and ICH GCP with regards data management for CTIMPs.
- Provide 'best practice' standards for data management for non CTIMP C&V UHB sponsored studies and all trials hosted by C&V UHB.

Scope

This SOP should be used by the CI of C&V UHB sponsored CTIMPs as well as any research staff involved in collecting, entering, checking, correcting, transferring and analysing data for C&V UHB sponsored CTIMPs. It is expected that the CI will ensure clear trial specific plans for data management in addition to these guidelines. Whilst this SOP is aimed at C&V sponsored CTIMPs, the principles herein are also applicable for all research undertaken at C&V UHB.

Equality Health Impact Assessment	An Equality Impact Assessment has been carried out on the Research Governance Policy (UHB 099), under which this SOP falls, and there was no impact.
Documents to read alongside this Procedure	UK Policy Framework for Health and Social Care Research 2017. Research Governance Policy (UHB 099) Archiving of Clinical Trial and Research Study Data (UHB 121) Managing amendments for UHB Sponsored Research (UHB 032) Notifications of serious breaches of GCP or study protocol (UHB 247)
Approved by	*Research Governance Group*

Accountable Executive or Clinical Board Director	Executive Medical Director
Author(s)	Commercial Trial Manager, R&D Office

Disclaimer

If the review date of this document has passed please ensure that the version you are using is the most up to date either by contacting the document author or the Governance Directorate.

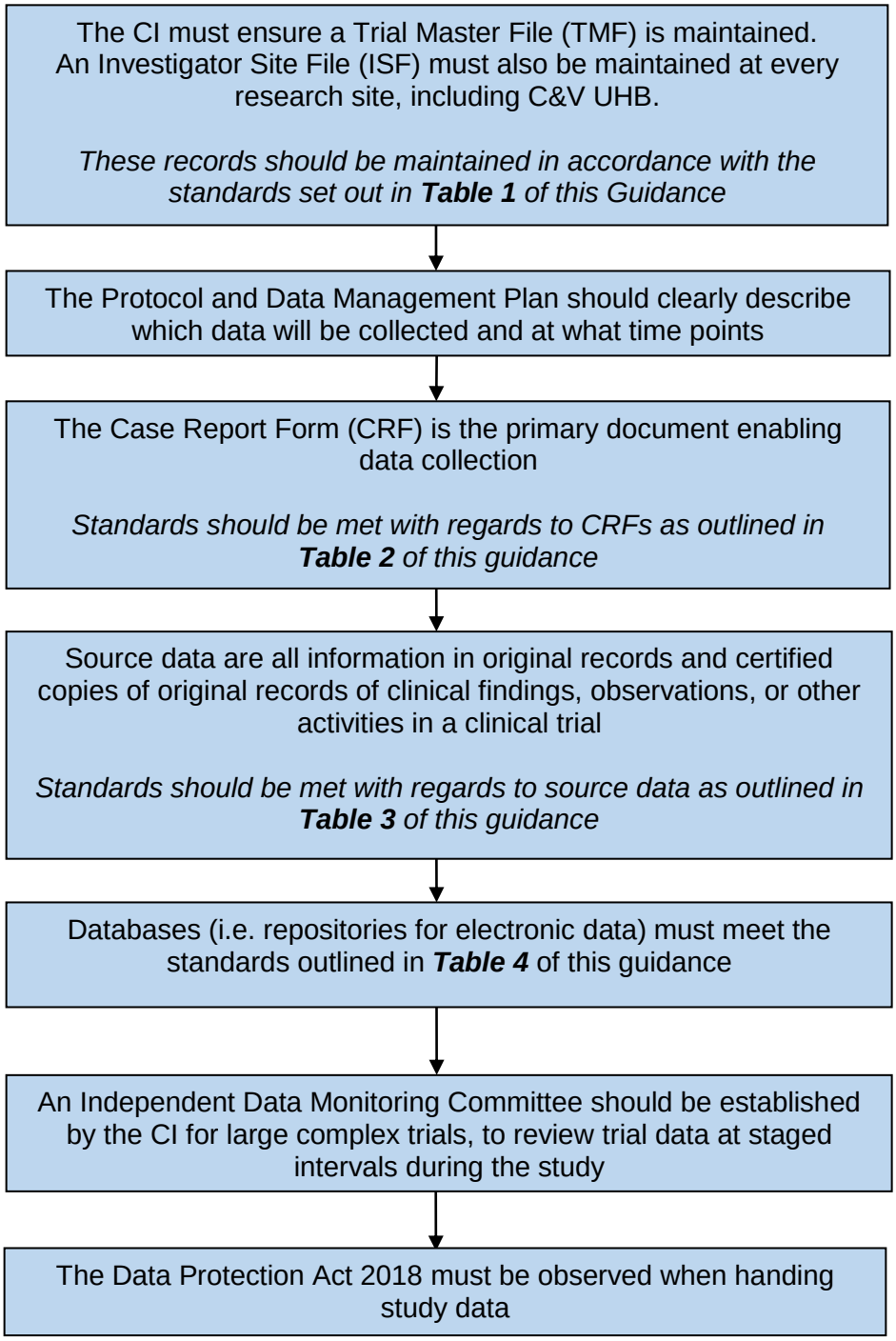
Summary of reviews/amendments

Version Number	Date of Review Approved	Date Published	Summary of Amendments
1.0	17/07/2021	29/04/2019	This document was previously a guideline (UHB139). Document is now a SOP and has totally redrafted with inclusion of flow diagram and tables to achieve a shorter

			more user-friendly document. The new document is adapted from the North Bristol NHS Trust SOP Data Management with their kind permission.
1.1			Minor spelling mistakes and formatting corrected
1.2	08/06/2022		No updates required however minor formatting corrected

CONTENTS

1. SOP FLOW CHART



2. DEFINITIONS/ABBREVIATIONS

CI	Chief Investigator
CRF	Case Report Form
CTIMP	Clinical Trial of an Investigational Medicinal Product
C&V UHB	Cardiff and Vale University Health Board
DMP	Data Management Plan
ICH GCP	International Conference on Harmonisation Guidelines for Good Clinical Practice
IDMC	Independent Data Monitoring Committee
ISF	Investigator Site File
HRA	Health Research Authority
HCRW	Health and Care Research Wales
PI	Principal Investigator
R&D Office	Cardiff and Vale University Health Board Research & Development Office
Sponsor	The individual, company, institution or organisation, which takes on ultimate responsibility for the initiation, management (or arranging the initiation and management) of and/or financing (or arranging the financing) for that research
TMF	Trial Master File
TSC	Trial Steering Committee

3. PURPOSE AND SCOPE

The purpose of this SOP is to describe the 'minimum' standards required for collecting, maintaining, verifying, correcting, transferring, and analysing data generated by C&V UHB sponsored CTIMPS.

For CTIMPs, the standards defined in The Medicines for Human Use (Clinical Trials) Regulations 2004 apply, along with those described in ICH GCP. The data generated through research may be used to influence or drive changes in clinical practice. The standards are in place to ensure that both robust data are generated and patients are safe. The need to be able to robustly defend the source of the data and the systems through which it passes until publication is thus paramount, and robust systems to document the effects of investigational medicinal products on human subjects must be in place.

A substantial amount of documentation is generated before, during and after undertaking any research project. It is important that such documentation is complete, legible and easily accessible at any time for monitoring, audit or inspection. In accordance with ICH GCP the Sponsor should ensure appropriately qualified individuals are responsible for the overall conduct of the research study, handling the data, verifying the data, conducting the statistical

analyses, and preparing the study reports. For C&V UHB sponsored studies, these responsibilities are delegated to the CI.

Within C&V UHB, all Sponsored CTIMP trials require involvement of a Clinical Trials Unit (CTU) and data management within the CTU should meet the 'minimum' requirements of these guidelines.

All documents for C&V UHB sponsored studies should be created and used in line with C&V UHB research SOPs. Moreover, the development of patient facing documentation should be in line with national HRA and Health and Care Research Wales guidance.

http://hra.nhs.uk/resources/before-you-apply/consent-andparticipation /consent-and participant-information/

Who should use this SOP?

This SOP should be used by the CI of C&V UHB sponsored CTIMPs as well as any research staff involved in collecting, entering, checking, correcting, transferring and analysing data for C&V UHB sponsored CTIMPs.

Although this SOP is aimed principally at C&V UHB sponsored CTIMPs, the principles are also applicable for all research undertaken at C&V UHB.

For studies sponsored by C&V UHB, it is expected that the CI will confirm clear trial specific plans for data management in addition to these guidelines.

When should this SOP be used?

This SOP should be used before, during, and after conducting CTIMPs sponsored by C&V UHB, to determine the standards required for collecting, maintaining, verifying, correcting, transferring and analysing data generated.

4. DATA MANAGEMENT 'MINIMUM' REQUIREMENTS

4.1 Maintaining the Trial Master File (TMF) and Investigator Site File (ISF)

ICH GCP defines documentation as: "all records, in any form (including, but not limited to, written, electronic, magnetic and optical records, and scans, x-rays, and electrocardiograms) that describe or record the methods, conduct and/or results of a trial, the factors affecting a trial, and the actions taken."

Essential documentation is "documents which individually and collectively permit the evaluation of the conduct of a study and the quality of the data produced". The 'essential documents' are therefore the minimum required documents to be maintained during any research project.

There are essential documents collected before, during, and after a clinical trial**.** The specific documents to be maintained for each project will vary and it is

therefore important that the specific essential documents for individual projects are considered on a case-by-case basis by the CI before initiating the trial.

The Trial Master File (TMF) and Investigator Site File (ISF) represent the standard filing system for the storage of essential documentation. The CI should set up and maintain the TMF (at C&V UHB, the TMF and C&V UHB ISF can be one and the same file) and the PI at each trial site should set up and maintain a separate ISF. A Site File Template for Non-CTIMP studies, is available on the HRA internet site and should be used to determine the appropriate content/structure of these files. Standards for maintaining the TMF and ISF are outlined in Table 1. For Non-CTIMPs, a study master file index is available, FR-RG-015 on the UHB intranet.

The R&D Office will retain electronic records relating to research governance and sponsorship, but are not required to retain Case Report Forms (CRFs) or other source documentation.

Storage	Documents contained in the TMF/ISF may include original regulatory approvals and confidential information. The files should therefore be stored in a secure place with restricted access. Documents may be kept in separate folders, files or cabinets but the TMF/ISF must indicate specifically where these are stored. Currently, there are no guidelines relating to the storage of documents in electronic format. It is good practice to print and retain hard copies of this information. Electronic copies should be password-protected or stored in a password-protected folder or drive for backup purposes. Direct access to all data must be provided for quality control and quality assurance reviews (e.g. monitoring and audit) or regulatory inspection.
Quality	All documents must be complete and legible so that they may be easily accessed and understood by monitor, auditors and inspectors.
Version Control	A system should be in place for version control of documents. It is recommended that a chronology of amendments is kept on file that records all the amendments submitted and the documents that they relate to. Old version of documents should be retained on file alongside the new versions and old versions clearly marked as no longer being used. Information on amendments can be found in SOP UHB 032 Managing amendments for UHB Sponsored Research

Table 1: Standards that should be met when maintaining the TMF and ISF

4.2 The Protocol

The protocol must clearly describe which data will be collected, and at what time points. Further requirements for protocol design are provided within the clinical trials resources on the HRA website.

https://www.hra.nhs.uk/planning-and-improving-research/research-planning/protocol/

The R&D Office must approve the protocol prior to any regulatory submissions, including amendments. On receipt of necessary regulatory approvals the CI is required to sign off the protocol either through 'wet' signature or electronically to certify that it is the current authorised Protocol.

4.3 Data Management Plan

A robust Data Management Plan (DMP) is required for all CTIMPs. The DMP should be approved by the Sponsor prior to study initiation and as required during the course of the trial.

Data management should be included as a standing agenda item at the trial management meetings, and should be checked as part of monitoring. The DMP should be reviewed regularly, at the frequency agreed between the CI and the R&D Office.

In the case of non-CTIMPs, the Sponsor will take a proportionate risk-based decision regarding the need of a separate formal DMP, taking into consideration whether there is sufficient information in the Protocol.

The DMP should specify, but not be limited to, the following information:

- Where the Reference Safety Information (RSI) is located, i.e. Investigational Brochure or SmPC.
- How data will be collected, clarified, stored and analysed with reference to the database lock process and audit trails.
- How eligibility of participants is assessed and documented.
- How trial related information is provided to participants and documented.
- The requirements for baseline data to be obtained and documented, and the implications if this is not adhered to.

4.4 Case Report Forms (CRFs)

A CRF is *"a printed, optical or electronic document designed to record all of the protocol-required information to be reported to the Sponsor on each [research] subject."* The CRF is the primary document enabling collection of data. Standards that should be met with regards to CRFs are outlined in Table 2.

Design	The CRF should be designed by the CI (or delegated individual(s)) before the research begins. It should be designed to only collect data as required in the protocol and nothing more, to comply with data protection requirements. CRFs may be paper or electronic. Original paper CRFs form part of the trial master file, as an essential document. Any verification of data which must be done by particular members of the research team (e.g. inclusion/exclusion criteria and safety data by medically qualified staff) must be evidenced. For paper CRFs, this would usually take the form of a signature and date; for electronic CRFs, this may be carried out by means of audit software incorporating particular logins, or documented separately within the source data. The CRF and any amended versions must be signed off by the CI prior to implementation.
Validation	The CRF should be reviewed by a range of staff to ensure ease of data collection and inputting, and all data is collected in line with the planned aims and objectives of the trial. Any changes to the CRF should be documented clearly and stored in the TMF. The amended CRF should be signed off by the CI before implementation. All electronic eCRFs should be validated before use (see Appendix 1)
Completion	The recording of data on the CRF should be performed by the PI at each trial site, however this responsibility may be delegated to other members of the team if appropriately trained. Paper CRFs should be completed in ink and data fields should not be left blank. Where there is no data to record in certain fields they should be marked Not Applicable (N/A) or No Data (N/D). The precise completion of CRFs is vital to preserving confidence in the findings of the project and therefore any discrepancies between the data required and the data collected or the source data should be minimised and explained. Any changes or corrections to a CRF or entries within them should be dated, initialled and explained (where necessary) and should not obscure the original entry. There should be an agreed system and process in place for authorising changes to data. The CI should agree with the research team what changes are acceptable, and document this. For example, the CI may agree that any member of the research team can amend clear transcription errors where the source data have not been transcribed correctly into the paper CRF. Other items, such as medical assessments and safety data changes

	should be authorised by the CI. The investigator should ensure the accuracy, completeness, legibility, and timeliness of the data recorded in the CRFs. CRFs should be completed at the earliest opportunity, contemporaneously with collection of source data wherever possible.
Review	The PI should regularly review the CRFs and source documents to identify any discrepancies or deviations from protocol and record such reviews. Protocol deviations should be documented and explained and all serious breaches notified to the R&D Office as per SOP UHB 235 (Managing Breaches of GCP or the study protocol).
Storage	The CI should ensure provision is made for trial sites to retain a copy of the CRF at site. Direct access to data must be provided for quality control and quality assurance reviews (e.g. monitoring and audit) or regulatory inspection.

Table 2: Standards that should be met with regards to CRFs

4.5 Source Data

Source data are 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. Source data are the first place that a piece of information is recorded.

Source data are contained in source documents (original records or certified copies). Source documents are considered essential documents that serve to certify compliance with ICH GCP and regulatory requirements.

Standards that should be met with regards to source data are outlined in Table 3.

Identifying source documentation	Prior to a trial commencing at a trial site, the PI at that site should identify what constitutes source documents at that site, and must record this on a Source Data Form. Source documentation may include (but is not limited to): hospital records, clinical and office charts, laboratory notes, memoranda, subjects' diaries or evaluation checklists, pharmacy dispensing records, recorded data from automated instruments, copies or transcriptions certified after verification as being accurate copies, microfiches, photographic negatives, microfilm or magnetic media, x-rays, subject files, and records kept at the pharmacy, at the laboratories and at medico-technical departments involved in the clinical trial. On occasion, the CRF may act as the source. Source data may be captured initially into a permanent

	electronic record. In this context 'permanent' means that all changes in the data are recorded in an audit trial (the minimum standard for this is a record of who made the change and when).
Source documentation verification	It should be possible to establish that the source information for data collected in the CRF existed at the appropriate point in time. That is, electronic systems used to record source data should have appropriate audit software providing the date, or be saved in a version controlled manner, and paper records should include a date and signature.
Source data retention	Source data must remain at the location at which it was generated. The location of all source data records should be documented to allow quick access. Direct access to data must be provided for quality control and quality assurance reviews (e.g. monitoring and audit) or regulatory inspection.

Table 3: Standards that should be met with regards to source data

4.6 Databases

A database is a repository for electronic data. Databases vary widely, depending on the size, type and complexity of the research being carried out. For a simple, small study, an excel spreadsheet may be used. At the other end of the spectrum are complex databases which have automated audit software and consistency checking capability, as well as the ability to generate data queries. Standards that should be met with regards to databases are outlined in Table 4. The CI is delegated responsibility for setting up and managing databases for recording trial data.

Design	A database should reflect the CRF so that the data required by the protocol can be collected. The CI should check that the database meets the needs of the study by reviewing and testing it, and documenting that the database meets the required specifications (user acceptance testing). The database should have suitable audit trail functionality. See Appendix 1 for basics of computer system validation and backup 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. If there is blinding involved in the trial, the system should allow this to be maintained.
Data entry	CRF data queries should be raised and resolved before entering data in the database. The process for managing data queries should be

	specified in the DMP. Data entry should be completed by trained delegated staff. Data should be entered in a format that allows for analysis, e.g. coded. Clinical data also needs to be coded for recording of all adverse events. Plans for coding must be incorporated into the DMP. To reduce errors, data should be checked. This may involve double entry checking (if sophisticated systems that allow for this are in use) or visual checking the database against CRFs. Records of checks and audit trails must be retained as part of the essential documents. Predicted timelines for inputting the data across sites should be included in the DMP, with a contingency plan should these timelines not be met. The Sponsor should be informed if the contingency plan is drawn upon, with regular reviews of the action.
Validation reports	Post-entry computer tests should be undertaken, for example by running lists of all missing values will be listed, or all values outside of pre-defined range. Logical checks should also be performed to check for implausible data. Post-entry checks should be defined before the study starts in the DMP. Records of checks and audit trails must be retained as part of the essential documents
Change control	As more data are entered, or changes are made it is important that an audit trail of the changes is available, so that previous versions of the datasets can be accessed if necessary. For sophisticated databases, the mechanism may be by using the database software to record changes to data fields and the associated logins that carried out the change(s); or a simpler database this might be by saving subsequent copies with a version number and date and a form of identification of the person who modified the file (e.g. initial and last name).
Management	The CI should ensure there is a trial specific SOP (or agreed plan for non CTIMPS) for managing the study database. There must be adequate backup for the data.
Access	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.
Data Lock	Database locking is the process by which it is declared and identified as final. No changes to the data should be made once the database has been locked, and arrangements should be put in place to control access to the data and protect it. The files should be protected from editing and deleting, and the decision about the approach to doing this should be made in a risk-based way. Unlocking the database should take place only under exceptional circumstances, and requires agreement from Sponsor and trial statisticians. Written approval for data unlocking, the justification, the changes that will be made and the impact on the analysis must be recorded in the trial master file prior to unlocking.

Data Release	Data should be extracted from the locked database to carry out the final analysis. The process to do this should be adequately described. Test extracts may be made, and these must be stored in a separate location to the extracted datasets on which the analysis will be performed.

Table 4: Standards that should be met with regards to databases

4.7 Publication Plan

A publication plan should be created as a standalone document or included within the DMP to ensure suitable dissemination of the results. This should include the publication of non-significant results and lessons which have been learnt from any errors during the development or delivery of the study.

4.8 Independent Data Monitoring Committee (IDMC)

For large complex trials, the CI should establish an Independent Data Monitoring Committee (IDMC) to carry out reviews of trial data at staged intervals during the study. The role of the IDMC is review the interim results and determine whether or not there are any safety issues or any reason why the study should not continue.

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 is of sufficient quality.

The membership of the committee should include 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.

4.9 Data Protection

During the entire data management and validation process it is essential that all study data are kept in accordance with the terms of the Data Protection Act 1998 and the General Data Protection Regulation 2018

5 RELATED SOPS AND DOCUMENTS AND USEFUL ADDITIONAL READING

The following R&D documents are available on the C&V UHB intranet

UHB099	Research Governance Policy
UHB121	Archiving of Clinical Trial and Research Study Data
UHB032	Managing amendments for UHB Sponsored Research
UHB235	Managing breaches of GCP or the Study Protocol

A useful review article is Matkar S et al., Int. J. Clin. Trials. 2017 Feb; Feb 4 (1) : 1-6. An outline of data management in clinical research.

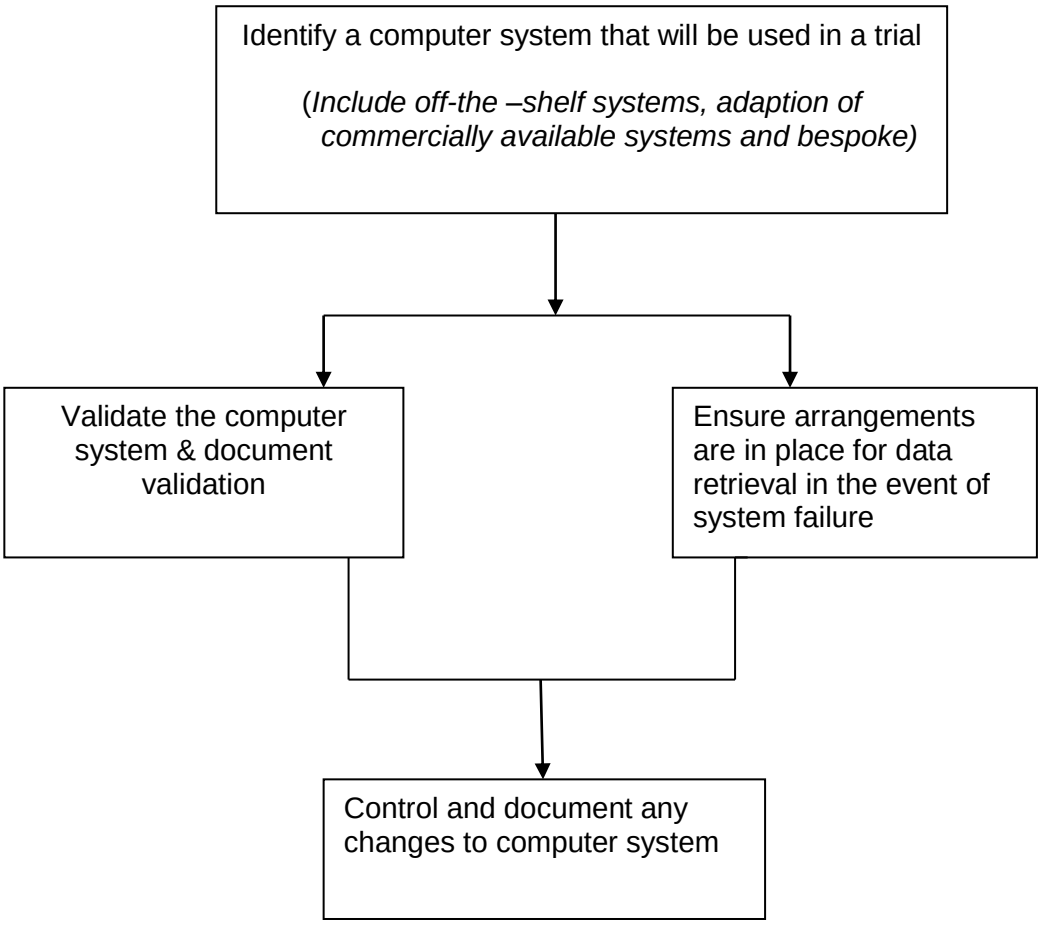
Good Clinical Practice Guide (ISBN978011 7081079) Compiled by the MHRA, Chapter 8 Data Management

6 DISTRIBUTION

This SOP will be made available on the R&D Office pages of the C&V UHB intranet.

APPENDIX 1: Computer systems validation, change control and back up

Over View Flow Chart



Validation

Validation should take the form of robust controls throughout the system's use, supporting documentation (as part of the DMP) and demonstrable evidence that a computer system in use is fit for purpose. This would include the use of test data to assess:

- Ease of system navigation and data entry.
- Ease of report generation.
- Accuracy of reports generated, including appropriately identifying 'red flags' e.g. appropriately identifying which fields should be populated.

There should be clear documented evidence of the process undertaken for system validation including the iterations and changes after feedback. The final version of the system should be approved by the CI and the Sponsor to confirm that the system is fit for purpose, prior to entering any trial data. For activities that are carried out by a third party, evidence of validation of relevant systems must be provided prior to their use. The final version of the system should also be approved by the CI and the Sponsor to confirm that the system is fit for purpose, prior to entering any trial data.

Examples of systems and the levels of validation required

System Type	Example System	Validation Required
Off the shelf	Microsoft excel	Cell formatting and formulae should be checked to ensure the required specification is met, and the checks made should be documented. For example, confirm that columns intended to receive a date are appropriately formatted; confirm the required number of decimal places is captured, confirm that values calculated from a number of cells are correct.
Trial specific adaptation of a commercially available off the shelf package	Microsoft access data, eCRFs	Document the agreed and approved specification, how the system will be tested (both by the users and the developers), that any issues with the system identified through testing have been resolved and the specification is met (validation report), instructions for use and how users will be trained, training records, how the final system will be released.
Bespoke systems	Any purpose built system solely for use in the trial	Document the process by which the decision to use a bespoke system was made and the risk assessment conducted as part of that decision making process, the agreed and approved specification (functional and user requirements), validation plan, code-testing documentation, that any issues with the system identified through testing have been resolved and the specification is met (validation report), instructions for use and how users will be trained, training records, how the final system will be released

Change control

There should be suitable and proportionate audit trail functionality for the computer system. The Sponsor's approval of trial databases is conditional on the ability of the database to produce audit trails for monitoring, for example, who made the change, time, date and rationale for changes. Any change to the system must be controlled and documented. The following information should be considered and documented:

(a) Reason for changes and person requesting changes.

(b) Risk assessment.

(c) Assessment of the changes and what actions are required.

(d) Approval of the changes (e.g. by CI and/or Sponsor).

(e) Testing.

(f) Validation report.

(g) Release documentation.

System backup

Arrangements should be in place to ensure that data can be retrieved if there is a computer system failure. Computer systems should be located within an infrastructure which provides for routine backups and disaster recovery in order to protect against accidental loss. Confirmation of this should be documented and local copies of different versions of data sets/databases should be retained if there is not audit software in place. These must be subject to organisational back-up.