University of Bristol | Cardiff University / Prifysgol Caerdydd | University of Exeter

Introduction to Data Management

Data management refers to planning, collecting, organising, processing, storing and maintaining data within research trials for the duration of the study.

Good data management practices and processes are essential when conducting research to ensure high quality research output and data integrity i.e. ensuring that data is accurate, consistent and complete.

Data management begins at study set-up where data management planning takes place, through the live study period involving the day-to-day management of data and its collection, to closure and beyond where data is analysed, preserved and shared for use and publication.

Data Management tasks during trial/ study set up

Creation of the study CRF

A case report form (CRF) is used to collect protocol-defined data in a trial for planned analyses. CRFs can be paper (pCRFs) or electronic (eCRFs). Data on pCRFs are usually entered by trials unit staff into the database, whereas eCRF are entered onto the database by site research staff.

Creation of Data Matrix/Metadata

A data matrix/metadata is a spreadsheet used to capture every field that will make up the CRFs used for a particular study and will form the basis of the database.

Creation of Database Specification

A database specification document is written prior to the build of a database, to ensure that the study requirements and communication process between the study team and database developer are clearly documented.

Database Testing

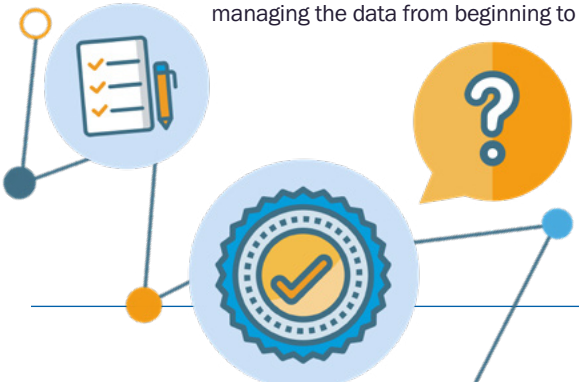
User Acceptance Testing is a key task in the setup, as it involves rigorously testing the study database ahead of its release to live.

Database training documentation

Documentation such a database manual and training slides are created to aid training sites/data collection staff on the trial database.

Creation of a Data Management Plan

A Data management plan is a document that outlines all the details and information about the processes that will be involved in managing the data from beginning to end.



Data Management tasks when a trial/ study is open

Querying Data

A data query is used to clarify or resolve any discrepancies or inconsistencies found in a data set to ensure high quality, reliable and accurate data.

Developing Reports

A range of internal and external reports are routinely required which involve data management input. These can include: updates on recruitment, screening, participant withdrawals, participant death, follow up, CRF completion, protocol non-compliances.

Monitoring

Monitoring is defined as the act of overseeing the progress of a clinical trial, and of ensuring that it is conducted, recorded and reported in accordance with the protocol. Monitoring can occur either centrally or on-site. This will include eligibility checks, quality checking data and ensuring any changes to data have been completed robustly.

Data Management tasks during trial/ study close down

Finalising data for analysis

Ensure the accuracy and completeness of all data by verifying that all data has been entered without errors and that missing data has been resolved or a reasonable reason for why it is missing has been provided.

Database Lock

Database lock is the process of approving and finalizing the completed data set, and "locking it" to further changes.

Returning data to sites

If the trial is using eCRFs then all data must be returned to sites at the end of the trial so they are able to retain their data.

GW4