ACDM GUIDELINES TO FACILITATE PRODUCTION OF A DATA HANDLING PROTOCOL

BACKGROUND

The need was identified by the Electronic Data Transfer Special Interest Group (SIG) for each company or organisation to have a document which describes and defines all data management activities for a study. The Working Party identified and named this document as a “Data Handling Protocol” (DHP). Other companies may know it as a Data Handling Manual, Data Handling Plan or similar name.

The ACDM Main Committee established a Working Party to develop a set of guidelines to facilitate production of a DHP. The Working Party consisted of representatives from Pharmaceutical Companies and Contract Research Organisations (CROs). The original document, first released in November 1994, has now been reviewed and updated based on experience gained, to create this new revised version.

CONTENTS OF THE DATA HANDLING PROTOCOL GUIDELINES

The DHP Guidelines provide the framework around which company-specific DHPs can be written. The DHP as described in the following sections of this document refers predominantly to in-house use, ie within the Pharmaceutical Company. However, additional issues will need to be addressed if a Pharmaceutical Company is collaborating with a CRO (see section on Contract Research Organisations).

It should be noted that this document is a guide only and can be customised by adding or removing sections according to company specific requirements, particularly in those areas where procedures are covered by company SOPs. It is envisaged that development of DHPs will help organisations to develop and to standardise data management procedures, therefore improving internal and external relationships through better understanding of roles and responsibilities.

The DHP Guidelines contain an outline and brief description of the fundamental requirements for company or organisation-specific DHPs.
USE AND BENEFITS

- Data Management staff will use the DHP guidelines to define or further develop the procedures that describe how they manage data, to what standards and why. These will be documented in DHPs and will enable Data Management staff to increase standardisation.

- Through reference to DHPs and involvement in their development, Clinical Research staff will be able to understand data management requirements so that they can apply them at investigator level, for example, which items of data will be queried and why.

- The DHP is a useful document for auditors and an important reference for retrospective retrieval of data.

- The development of a DHP will encourage forward planning.

- The DHP should lead to improved communication between all parties concerned, thereby leading to more efficient and accurate data collection, for example, it could be used to communicate database structures to Statisticians.

- CRQs could use DHPs provided by Pharmaceutical Companies (or vice versa) for more technical information, such as structure of data for transfer between them.

- DHPs could be used to document study-specific deviations from company SOPs.

- Most of the elements of these guidelines can be adopted whether a paper CRF system or electronic data capture system is used.

RECOMMENDATION

We strongly encourage organisations to consider the ACDM DHP guidelines as a reference, to help define, further develop and document company-specific data management procedures, to the benefit of Data Management staff, their colleagues and customers.
1. DATA HANDLING PROTOCOL ADMINISTRATION

1.1 Ownership

The DHP should be initiated and subsequently owned by Data Management staff, who will be responsible for production, review and maintenance, until trial completion.

1.2 Cover Page

Each DHP should have a cover page, which should include:

- full title of study protocol
- study number and all identifying information, e.g. dates, version numbers
- effective date of DHP.

1.3 Signatures And Distribution List

Each DHP should be signed-off to confirm that it has been reviewed and is accepted by appropriate individuals, e.g. Clinical Research, Data Management and Statistics staff. Signature lists will need to be defined within each individual company or CRO. A distribution list should be included to ensure all DHP recipients are known and can be issued with document updates and amendments.

1.4 Contents List

Each DHP should include a contents list to ensure ease of use of the document. It should include all relevant sections.

1.5 Page Numbers

Page number should be indicated on every page of the DHP, if possible.

1.6 Document Status Identifiers

Each draft DHP should be clearly identified as a draft, for example, draft version 1. If possible the draft/version number should appear on each page. The final DHP should be clearly identified as the final document. Any amendments to the DHP after it has been issued should be reflected in the document status identifier, for example, by updating the document version number. Document control procedures should be put in place to ensure all appropriate people have and are working to the correct version of the DHP.
1.7 Amendment And Change Procedures

It is recommended that each DHP includes:

• A statement of how the DHP will be followed

• Details of how any deviations from or modifications to the DHP will be dealt with in terms of:
  - who needs to agree them
  - how the changes will be made

• Details of how minor amendments will be made after finalisation.

2. STUDY PERSONNEL

Everybody involved in the study should be listed, for example:-

  - Clinical Research staff with contact information e.g. phone number
  - Data Management staff with contact information e.g. phone number
  - CRO staff with contact information

This section may also include organisation charts.

3. STUDY OBJECTIVES AND DESIGN

Study objectives and design and/or a study summary sheet should be defined in the study protocol. They should either be transferred as written into the DHP or cross-referenced.

4. TIMESCALES AND KEY ACTIVITIES

This section of the DHP should include a schedule of activities for the Data Management Group. Milestone dates should be agreed between all relevant personnel e.g. study start date, date of receipt of first CRFs (see also section 7).

Timescales should be for Data Management staff but this will need to be an agreed agenda with Clinical Research staff. Timepoints at which patient recruitment can be reviewed should be agreed early in the study so that Data Management staff are aware of any shifts forward or backward in the original schedule (see also Section 7, Data Flow and Tracking).

A mechanism should be identified for informing appropriate people in case changes to study schedules occur, e.g. who should be informed, who should make decisions about proceeding, and what actions should be taken.
5. DATABASE DESIGN

Note: For DB design, we’ve assumed that CRF design is complete and that design of the database includes design of data entry screens and data structures.

It is important to document how clinical data are stored on the database. Benefits of doing this include long term ease of reference to and retrievability of the data. Specification of data structures may also need to be provided to, or supplied by, CROs, if aspects of data management are contracted out, in order to facilitate electronic data transfer. The DHP should therefore include the following elements that describe how the data are stored.

5.1 General

- **Software**
  - name of package/system used
  - version number

- **Hardware**
  - description of platform, e.g. Novell Network, PCs running Windows

- **Location of Database**
  - detailed description of where the database is located on the system

5.2 Details

- **Detailed description of database structure, for example:**
  - list of data files with brief description of contents (cross-referencing appropriate CRF pages). See Appendix I.
  - list of all fields/variables/items on each file with details of all attributes e.g. type (numeric, text, date, time), length, formats, labels etc. See Appendix 1.
  - method of associating database fields/variables/items with questions asked on the CRF e.g. annotated CRF.

- **Detailed description of data entry system (see Section 8)** include:
  - design of windows/screens/forms with hard copies, if required
  - relationships between these windows and data files
  - list of all validation carried out on entry, e.g. range checks, formats, etc (see section 6.4) (Data Validation Checks)
  - type of entry to be employed (see Section 8 Data Entry Procedures)
e.g. single with manual checking double with file comparison
double with verification on second entry
- access to query tools
- data access privileges.

- For electronic data, details of facilities for or methods of import/export of these data.
  Details of how import/export of data are validated, if appropriate.

6. **MONITORING/VALIDATION GUIDELINES**

Monitoring/validation guidelines should be produced with the following aims:

- To avoid unnecessary duplication of checks
- To standardise the clinical review of data
- To standardise the data validation procedure
- To achieve a consistent database.

Joint input is required from Clinical, Data Management, Statistics and all departments who have a stake in data consistency and quality. These guidelines should be written before any monitoring starts, including investigator site monitoring.

The monitoring/validation guidelines may include the following sections:

6.1 **Randomisation And Inclusion/Exclusion Issues**

A system for uniquely identifying all patients during the screening visit should be defined, whether or not patient numbers are issued at randomisation visits. There should also be a link between patients who are screened and those who are eventually randomised. It is important for organisations, for example CROs and laboratories who handle pre-study data to be able to link the data to patient numbers once they are issued.

Details of how to handle the inclusion and exclusion criteria and randomisation issues, should be defined before any data are collected. For example, a central laboratory may be required to send messages directly to investigational sites on the randomisation or dose titration. This information is usually included in the study protocol but is not detailed enough for data handling.

6.2 **Standard Procedures**

The monitoring guidelines should state clearly the rules and principles to be followed. It is often sufficient to list relevant SOPs and ensure availability to everyone involved. Quality Assurance personnel should be involved in developing the guidelines, as appropriate.
Subsections may include:

- who may write on CRFs, and on which copies
- how corrections should be made
- a reminder to sign and date all changes
- how queries are to be resolved
- a reminder to document deviations and amendments to the guidelines.

This section should describe monitoring checks that will be carried out by Clinical Research staff. It should be the joint responsibility of the Data Management and Clinical Research managers to ensure that unnecessary duplication of effort is avoided and appropriate training provided. The section should include checks done at the study site by Clinical Research staff, agreement on required/mandatory items (those that must be collected or followed up if the data are missing), and translation of text where applicable.

6.4 Data Validation Checks

Automated checks should be specified in enough detail to enable set up of data entry screens and validation programs. The detail should be sufficient to enable production of a comprehensive programming specification document for use by another individual.

Checks that can be done automatically during or after data entry should be clearly identified. Decisions on how and when these checks are to be done may be influenced by the length of time between clinical monitoring and automated data validation and whether errors detected will influence recruitment or the further collection of data.

6.5 Quality Checks

It is assumed that quality assurance will be dealt with by an independent individual/department, and is therefore not covered by this document.

Ongoing and final data quality checks to be done by Data Management staff, before the database is finalised, should be covered in this section.

This section may include subsections on the following:

- audit trail checks
- sample checks of critical data
- data review checks to ensure monitoring has been performed consistently
planned external audits, either by a different department or specialised CROs.

Fig 7.1 Summary Of Data Flow

7. DATA FLOW AND TRACKING

The sequence of events and responsibilities for handling data need to be clearly understood by everyone involved in the team working on a study or project.

Figure 7.1 is an example summary data flow.

7.1 Responsibilities

Responsibilities must be clearly assigned for monitoring the status of data flow and how it is tracked. Responsibility must also be assigned for authorising a clean database. These should be documented in the DHP.

7.2 Planning And Preparation

People working on a project should be made aware of how all data in a trial are to be collected. This may be electronically or as a hard copy (CRF). Hard copies of the data may be transmitted in a variety of formats, e.g. by “no carbon required” (NCR) paper, fax, courier etc. If more than one copy of the data are received, distribution of the copies should be documented. The format and method of data collection should be documented in the DHP.

Clinical Research staff should inform Data Management staff of the regularity of their monitoring visits and rate of recruitment. Data Management can then prepare queries in advance and predict the volume of data to be received, thereby facilitating workload management.

7.3 Tracking

A tracking system should be used to track each CRF and other associated patient data through the data management process, for example:

- date of receipt of a CRF and individual pages
- date entered onto first and second databases and by whom
- date compared
- date data validated
- date first queries sent out and date all queries returned
• status of individual queries
• date database declared clean.

Tracking systems usually allow detailed status reports to be produced electronically. These status reports can be made available to all personnel working on a project. Details of which items will be tracked and which status reports produced should be recorded in the DHP.

8. DATA ENTRY PROCEDURES
(INCLUDING CODES AND CODING DICTIONARIES)

Responsibility for each data entry procedure needs to be agreed and specified. The following aspects need to be documented in the DHP.

8.1 Identify Method Of Data Collection
• Case report forms
• External databases
  • e.g. centralised laboratory
• Electronic data capture equipment
  • e.g. ECG monitoring equipment, electronic diary cards

8.2 Define Type Of Data Entry
• From case report forms
  • Double/Single entry, Off/On line
• From external databases
  • define means of access to data e.g. electronic transfer from disk, tape, modem or printouts
• Electronic data capture equipment
  • transcription
  • customised report entered onto database
  • transferred direct from monitoring equipment
• Optical technology
  . e.g. use of bar codes

8.3 Data Preparation

• Define how data will be prepared for entry onto the database.

• Define pre-entry checks e.g. what to do if patient ID is incomplete or there is a signature missing.

• Provide instructions for data entry
  . document all data handling conventions e.g. how symbols are handled
  . any special instructions e.g. if a date is incomplete
  . handling of illegible, inconsistent or missing data.

• Determine volume
  . size of each batch
  . number of forms, over what time period, at what rate
  . set target timings e.g. turnaround time for data entry.

8.4 Handling Of Text

• Define how text will be handled e.g. use of coding or key text. Consult statisticians to determine how data will be reported.

• Document handling of foreign text/translations.

• Use of dictionaries
  . state which, e.g. COSTART, ICD9, WHO, MEDDRA or in-house
  . any in-house dictionary will need to have a well defined set of rules/principles, a request procedure and a dictionary administrator
  . if coding is used, when will this be performed e.g. prior to data entry.

9. SPECIFICATION FOR CLINICAL LABORATORIES

To enable laboratory data to be integrated into the study database, it is necessary to specify the data to be collected and their format. If the laboratory has the facility to transfer data electronically, we recommend that the ACDM Standard for the Electronic Transfer of Laboratory Data be used in conjunction with this document.
9.1 Information To Be Supplied

The following information should be specified to clinical laboratories via the DHP:

- Format for transfer of laboratory data to client company
- Format for any demographic data to be included
- Definition of special codes for missing or invalid results
- Textual data format/coding
- How extra visits will be handled.

9.2 Issues To Be Agreed

Agreement should be reached with clinical laboratories on:

- Identification of laboratory data e.g. header information, patient and visit identifiers and dates
- Reference range information, including:
  - name of test
  - reference range, with units
  - relationship to age and sex
  - how the ranges were derived
  - dates when effective
  - format of reported values and ranges, e.g. number of decimal places
  - system for notifying changes to the reference ranges during the study

- Provision of a sample report prior to study commencement
- The system for flagging out of range values
- Information outlined in Section 10, Electronic Data Transfer, if the transfer is going to be automated
- Method and frequency of transfer.

10. ELECTRONIC DATA TRANSFER

If data for a study are to be transferred electronically from any external source to the study database, we recommend that the following should be specified by the client and agreed between client and external supplier of the data. All agreements should be documented.
10.1 Specifications

- Specification of formats for all data to be transferred (section 5.2, Database Structure). Consider use of standard formats e.g. ACDM Standard for the Electronic Transfer of Laboratory Data.
- Definition of codes to be used for the entry of results, including codes for missing data, non-quantifiable data.
- Formulae for calculated variables.

10.2 Transmission

- Method of validating conversion of data from the supplier’s database/data format to the client’s required data format prior to transmission.
- Transmission details e.g. disk (size and type), modem.
- Identification of compression software, if data are compressed before transmission from supplier to client.
- Transmission frequency and criteria e.g. all data at each transmission, completed patients only, all data at end of study. The need for and timing of a test transmission should be agreed if appropriate.
- Method of confirming that data transfer has been completed successfully i.e. that contents of transmitted files are received unaltered. Agreement on how unsuccessful transmission will be dealt with.
- Virus checking.
- A log of what has been transferred and when.

10.3 Hard Copies

- Procedure for transfer of hard copies of electronically transferred data.

10.4 Amendments To Source Database.

- Identification of source database for correction and amendment of data, if it is necessary. If external supplier holds the source database, agreement on re-transmission of data to the client following correction or amendment.
11. QUERY HANDLING

Figure 7.1 in section 7 (Data Flow and Tracking) gives a description of query flow during the course of the study. For all stages in the process at which queries may be addressed, e.g. for query handling by the monitor at the investigator site or in-house, for data clarification following data entry, data validation and during Statistical analysis etc., the following should be agreed and documented in the DHP:

• identification of how queries will be tracked

• identification of who is responsible for checking and tracking queries

• expected turn-round time for queries to be answered

• details of how and where queries and the answers to queries will be documented

• identification of who is responsible for generating queries at each stage of the process after data is in-house

• identification of who is responsible for ensuring the investigator has copies of all queries in the investigator file

• identification of who is responsible for making the required changes to the data and/or to the query database

• identification of who is responsible for ensuring that answers to queries are received before the database is locked

• details of how relevant parties will be made aware of any outstanding queries e.g. when the database is locked, and how they will be handled (if appropriate).

12. BACK-UP AND RECOVERY PROCEDURES

The security and safety of the clinical data when it is on the database is a primary concern. Procedures should exist which ensure that the data will be safe and intact if anything goes wrong with any element of the database system. These procedures must be documented or referenced in DHPs for individual protocols or in general procedures that apply to all protocols e.g. SOPs.

• Procedures should fit in with corporate policies and not be a departmental responsibility.

• Procedures should exist to ensure that system back-up and recovery can be carried out. The procedures may include, for example:

  • how often back-ups are taken
how long back-ups are kept
  time within which the system must be up and running again
  responsibilities e.g. for different hardware.

• Procedures must be documented and tested, the detail of which may vary depending on
  hardware, software, existence of SOPs etc.

13. ARCHIVING AND SECURITY

13.1 General Considerations

• Procedures should fit in with corporate policies and not be a departmental responsibility.

• All procedures and agreements should be documented in DHPs for individual protocols
  or in general procedures eg SOPs that apply to all protocols.

13.2 For Documents (e.g. CRFs) The following need to be agreed:

• How and where data will be stored and how they will be kept secure e.g. in a lockable,
  fire resistant cupboard

• When and by whom data will be archived.

13.3 For Data

• Consider using corporate archive procedures.

• Agree timing for archiving of electronic data e.g. at study end, once all reports are
  written.

• Ensure that responsibility for electronic archiving of data is assigned and it is carried out
  to the agreed time schedule.

• Use self-explanatory structures e.g. sub-directories as follows:
  • trialno\data
  • trialno\screens
  • trialno\progs
  • trialno\rand
  • trialno\list.

• Compress data if storage space has to be considered. The compression software should
  be saved.

• Consider stability of electronic archiving medium.
• Retain hardware that can use the electronic archiving medium
• Test retrieval of data from archive.
• Archive more than once and retain secure copies in more than one location e.g. store one copy on-site and another off-site.

14. **CONTRACT RESEARCH ORGANISATIONS**

This section aims to highlight issues to be addressed if a Pharmaceutical Company is collaborating with a CRO.

14.1 **Responsibilities**

• All data management responsibilities should be set out clearly in the contract and the CRO and client are both responsible for fulfilling the terms of this contract.
• In order to maintain consistent data interpretation the CRO should allocate a named contact and changes in project personnel should be notified to the client.
• In order to maintain consistent data interpretation the client should allocate a named contact and changes in project personnel should be notified to the CRO.
• All documentation should be agreed between the client and CRO, e.g. DHP, Data Entry Conventions etc. There should be written agreement on the SOPs to be used.

14.2 **Database Design**

• The client must be made fully aware of the software and version that the CRO uses and its capabilities.
• The client may specify their own format or choose to accept the CRO format. The requirement must be clearly stated at the outset.

14.3 **Electronic Data Transfer**

• All electronic data must be virus checked before being loaded onto the system.
• Formats agreed for electronic data transfer must be tested before ‘real’ data arrives.

14.4 **Monitoring Of CRFs**
• It must be clearly stated at the outset whether the client or the CRO is going to monitor the CRFs and to what level.

14.5 Data Flow And Tracking

• It may be appropriate for a data flow diagram to be provided, particularly for a complex data flow.

• The client must try to maintain a regular flow of CRFs and keep the CRO informed of the rate of receipt of data and last CRFs.

• Each CRF should be tracked to a level that has been agreed, e.g. by page.

• The client and CRO should agree a timescale for receipt of data through to issue of queries and return of responses.

• The expected regularity of the monitoring visits should be documented so that data flow occurs at the times most beneficial to both client and CRO and time can be managed efficiently.

• Progress must be closely monitored by the CRO. If any delays are anticipated this should be discussed with the client and corrective action taken before target dates are threatened.

• There should be regular status reports to the client at agreed time intervals, so the client is kept informed of what is happening. The format of the status reports should be agreed at the outset by the client and the CRO.

14.6 Data Entry Procedures

• All specific data handling conventions should be documented and approved by the client.

• The action the CRO is expected to take when data are illegible, illogical or missing must be documented. Anything which requires interpretation should be referred back to the client.

• Procedures for coding text should be agreed. Choice of coding system will normally depend on client preference. Most CROs are able to use a variety of systems. The client should ensure they are satisfied that codes are allocated by an appropriately qualified person.

• The client should ensure that the procedures for verification of data meet their requirements.
14.7 Data Validation

- The client may specify any or all validation checks or require that the CRO proposes checks based on the Monitoring/Validation Guidelines. These should then be approved by the client.

- Discrepancies found may be resolved according to agreed conventions or transferred to query forms for client or monitor or investigator resolution.

14.8 Query Handling

- A CRO will generally have their own format of query forms which will be used unless the client specifies that their own are required.

- A query tracking system should be used.

- Details of query distribution should be agreed at the outset. Queries may go to one contact at the client company or to individual monitors or direct to investigators.

- The query database should be monitored to check whether there are any particular items which generate a lot of queries or unusual trends. The client should be made aware of these problems and corrective action be agreed with the CRO.

- The timing of monitoring visits should be communicated so that queries can be prepared in advance.

14.9 Quality Check/Audit

- It is assumed that quality assurance will be dealt with by an independent individual/department, and is therefore not covered by this document.

- Data quality checks and auditing procedures should be agreed and documented.

- The acceptable quality limit should be agreed between client and CRO and actions documented, if the limits are not met.

14.10 Data Transfer And Archiving

- The client’s data transfer requirements for the end of the study should be stated.

- The archiving policies should be agreed and documented in the DHP.
APPENDIX 2

GLOSSARY OF TERMS

Clinical monitoring checks - manual checks performed at the study site normally by Clinical Research staff
Data validation checks - computerised or manual checks performed inhouse, normally by Data Management staff.
Verification - procedure to confirm accuracy of initial data entry, e.g. 100% manual check, independent second entry with file comparison
Case Report Form (CRF) - may imply CRF page, booklet or the whole CRF
CRO - Contract Research Organisation